The genetic etiology of familial breast cancer: Assessing the role of rare genetic variation using next generation sequencing

Hilbers, F.S.M.

Chapter 1

General introduction

1 Clinical aspects of familial breast cancer

1.1 Breast cancer tumorigenesis

Healthy breast tissue consists of a broad range of cell types meticulously arranged into highly organized structures of branched ducts and lobules supported by fatty and fibrous connective tissue. Homeostasis is maintained by tight regulation of the morphology and proliferation of each cell via both intra- and extracellular signals. These signals comprise multiple barriers which a cell has to overcome in order to become malignant i.e. to be able to survive and proliferate, and potentially invade and metastasize, regardless of external signals. To achieve this, a cell has to acquire, among others, the capabilities to ensure continued growth signaling, ignore growth inhibition and apoptotic signals, adjust its energy metabolism and stimulate angiogenesis.[1] Traditionally this malignant behavior was thought to be the result of the accumulation of somatic genetic changes within a single cell: the so-called somatic mutation theory. Nowadays, it is appreciated that the process of tumorigenesis is much more complex and involves changes at various organization levels and a dynamic reciprocity between them, including genetic mutations, epigenetic changes and alteration of the microenvironment.[2–4]

Although no longer thought to be the only facet in tumorigenesis, somatic genetic alterations that critically contribute to malignancy, the so-called driver mutations, remain one of the most well studied aspects of breast tumors. Genes frequently affected by mutations or structural variants include *TP53*, *PIK3CA* and *GATA3*.[5] However, the mutational landscape of breast tumors is far from homogeneous; a large diversity in the combination of alterations in over 50 significantly mutated genes has been found.[6–11] Besides genetic alterations, it has become clear that epigenetic changes play an important role in breast tumor development as well. For many cancers, including those of the breast, a global increase in CpG-island hypermethylation has been observed. Methylation of the CpG-islands in a promoter region generally leads to reduced expression of the corresponding gene. Interestingly, hypermethylation of promoter regions is especially frequently seen in tumor suppressor genes, including *CDH1*, *CDKN2A*, *PTEN* and *BRCA1* (reviewed in [12,13]). Nonetheless, with regard to epigenetic alterations, substantial heterogeneity has been observed between tumors: a large study analyzing genome-wide methylation data found that breast tumors can be classified into at least five different methylation groups by unsupervised clustering.[7]

Another important factor in breast cancer tumorigenesis is the microenvironment. The tumor microenvironment comprises the extracellular matrix and multiple cell types such as fibroblasts and immune cells. Moreover, it acts as the medium for many important soluble factors such as cytokines, growth factors and enzymes. The microenvironment can both inhibit and facilitate the malignant behavior of the tumor cells.[4] For example, while normal myoepithelial cells in the tumor microenvironment may inhibit the growth of breast tumor cells, cancer-associated fibroblasts can promote growth and invasion.[14]

1.2 Breast cancer subtypes

Although all tumors in the end acquire approximately the same set of capabilities, the examples described above illustrate the tremendous variation in the way they achieve this. Because of this heterogeneity, breast tumors are often classified in subtypes. The large majority of breast tumors, roughly 95%, are adenocarcinomas arising in the breast epithelium, while a small percentage consists of sarcomas originating from the stromal cells of the connective tissue of the breast. Adenocarcinomas are further classified according to their morphological and cytological patterns. By far the largest group are invasive ductal carcinomas of "no special

type", which is a diagnosis of exclusion used for tumors that do not possess characteristics that would classify them as one of the special subtypes. These special subtypes comprise ~25% of all breast tumors and consist, among others, of lobular, tubular, cribiform and metaplastic carcinomas.[15] Histological subtypes can be associated with specific molecular characteristics, for example, while lobular carcinomas have fewer somatic genetic aberrations overall, they almost always lose E-cadherin function, through inactivating mutations or promotor hypermethylation of *CDH1* (the gene encoding E-Cadherin) or impaired integrity of the E-cadherin-catenin complex.[16]

Another important factor used in defining breast cancer subtypes is whether or not they express the estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and proliferation marker Ki67. The presence of these proteins is routinely assessed in clinical practice by immunohistochemistry and tumors are subsequently classified into four subtypes: Luminal A (ER+/PR+/HER2-/Ki67 low), Luminal B (ER+/PR+/HER2-/Ki67 high), HER2-overexpressing (HER2+, any ER/PR/Ki67) and basal-like (ER-/PR-/HER2-, any Ki67).[17] The presence of the three receptors indicate the pathways through which the cells receive growth signals and, more importantly, offer opportunities for therapeutic targeting. While HER2 overexpression in breast tumors was once associated with a poor prognosis, this drastically changed when the first HER2-targeting agent, trastuzumab, became available.[18] Also the dependence of ER-positive/PR-positive breast cancer on estrogen signalling can be targeted by several drugs, such as selective estrogen receptor modulators (e.g. tamoxifen) and aromatase inhibitors (e.g. letrozole). These therapies are not indicated for breast cancers lacking all three receptors, known as "triple negative" breast cancer, which are associated with poor prognosis. For this group of tumors, chemotherapy has long been the only available treatment option. However, recently clinical trials have started exploring immunotherapy as a treatment option for triple negative breast cancer as these tumors have been shown to be more immunogenic, with higher PD-L1 expression and T-cell infiltration than other subtypes.[19]

Besides these subtypes commonly used in clinical practice, further classification of breast tumors has been proposed based on gene expression analysis and mutational signatures. Gene expression analysis has resulted in a similar but more detailed range of subtypes, including luminal A, luminal B, HER2-enriched, basal-like, normal-like, Claudin-low and apocrine (reviewed in [20]). These subtypes are often called the "intrinsic subtypes". The luminal A, luminal B, HER2-enriched and basal-like subtypes largely overlap with their immunohistochemistry counterparts. The claudin-low subtype is thought to represent a group of tumors characterized by epithelial-to-mesenchymal transition and stem cell like features and which are usually negative for ER, PR and HER2.[21] The apocrine expression subtype represents another subset of triple negative breast cancers which express the androgen receptor.[22] Mutational signatures describe the kind of nucleotide substitutions and structural variations that are overrepresented in a tumor. These signatures are thought to provide insight into the mutagenic processes and possible DNA repair deficiencies that have shaped the tumor's genome. For breast cancer twelve different nucleotide substitution and six rearrangement signatures have been described, some of which have been linked to specific tumor attributes, such as deficiency in homologous recombination or mismatch repair.[6,23]

1.3 Breast cancer risk and familial clustering of breast cancer

Breast cancer is the most common cancer diagnosis and represents the most common cancer-associated cause of mortality in women worldwide.[24] In the Netherlands, approximately one in eight women will develop breast cancer at some point during her life.[25] The identification of risk factors is important for both prevention and the development of efficient screening programs aimed at early detection. However, the etiology of breast cancer is complex; among others, genetics, physical characteristics, lifestyle and reproductive factors are known to affect disease risk.[26] Gender is an obvious factor in breast cancer; the life time risk to develop breast cancer is approximately 13% for females, while for males it is only 0.11%.[25] Geographical location is another important factor associated with the risk of developing breast cancer. The incidence of breast cancer is remarkably higher in "western" countries (Europe, North America, Australia and New Zealand) compared to the rest of the world.[24] The incidence of breast cancer, alike that of many other cancers, also increases with age.[25] This association between cancer and age is generally attributed to the time needed for the stochastic process of accumulating the tumorigenic capabilities described above, through somatic mutations, epigenetic changes, telomere shortening, declined DNA repair efficiency and changes in the microenvironment.[27] In addition, early menarche and late menopause are associated with an increased risk of breast cancer.[28] The estrogen and progesterone fluctuation during menstrual cycles induces repetitive phases of mammary epithelium proliferation and regression, which cause an increased chance of genetic errors.[29] At the same time, reproductive factors, such as a first-full term pregnancy at relatively young age, a higher number of full-term pregnancies and total duration of breast feeding are associated with an decreased risk of breast cancer.[30] Pregnancy and lactation are thought to induce several, long-term, systemic and local changes that could explain their associated with a decreased breast cancer risk, which include changes in circulating hormone levels, estrogen responsiveness, number of mammary stem cells and differentiation status.[31] Furthermore, several lifestyle factors such as alcohol consumption, lack of physical activity, post-menopausal obesity and exposure to exogenous estrogen via oral contraception or hormone replacement therapy increase the risk of breast cancer (reviewed in [26]).

Family history is another well-established risk factor for breast cancer. Women who have one first-degree relative diagnosed with breast cancer, have a relative risk (RR) of approximately 1.8 to develop breast cancer themselves. Having three or more affected first degree relatives is associated with a RR of 3.9.[32,33] Moreover, a lower age of onset is associated with a higher risk in first-degree relatives compared to late onset breast cancer.[32,33] This relative risk associated with a family history is known as the familial relative risk (FRR). Several algorithms have been developed to calculate an individual's risk of breast cancer based on a specific family history, including BOADICEA,[34] BRCAPRO,[35] the Tyrer-Cuzick model,[36] the Claus model[37] and Gail model.[38] Studies in monozygotic and dizygotic twins have determined that genetic factors account for approximately 27% of the variance in breast cancer susceptibility. Although genetic factors likely play at least some role the etiology of every breast cancer case, only ~15% of cases are considered "familial". Of these familial breast cancer cases, approximately 5-10% carry a mutation in a known high-risk breast cancer susceptibility gene. Breast cancer susceptibility loci associated with a moderate or small risk increase have also been identified (see sections 2.1, 2.2 and 2.3 below). However, most familial breast cancer cases are not yet tested for these moderate and low-risk loci in diagnostic settings in the Netherlands and most EU countries. Moreover, all known genetic risk factors jointly still explain less than half of the familial relative risk.

1.4 Clinical management of hereditary and familial breast cancer

When a woman has multiple first or second-degree family members affected with breast cancer, genetic testing for mutations in breast cancer susceptibility genes is indicated. Moreover, if a woman herself has been diagnosed with breast cancer at a particularly young age, or with both breast and ovarian cancer or with bilateral disease, referral for genetic counseling might be appropriate. (See [39–41] for the exact indications for genetic counseling in the Netherlands.) When genetic testing is warranted, at a minimum *BRCA1* and *BRCA2* will be tested. Nowadays, however, gene panels containing additional moderate and high-risk susceptibility genes (see section 2.1 and 2.2) are commonly assessed using next generation sequencing (NGS, also known as massively parallel sequencing (MPS)). The results of these genetic tests can be classified in three categories. In the far majority, no potentially causal mutation will be detected. On the other hand, there are cases in which a mutation is found that is clearly associated with an increased risk of breast cancer. Besides these straightforward results, there is a third category of cases in which a genetic variant is identified for which the association with breast cancer risk is unclear. These variants are referred to as variants of uncertain significance (VUS). For *BRCA1* and *BRCA2* testing alone, in 5-10% of cases a clearly pathogenic mutation is found, while in approximately 15% a VUS detected.

With regard to clinical surveillance, a distinction is made between "hereditary" breast cancer, when a causal mutation has been found, and "familial" breast cancer, when either no mutation or a VUS has been detected. Guidelines for the clinical management of familial and hereditary breast cancer describe screening strategies including age to start screening, the frequency and the methods to be employed.[39–42] In the Netherlands, there are separate guidelines for families with *BRCA1*, *BRCA2*, *PALB2*, and *CHEK2* mutations in addition to guidelines for rare syndromes associated with an increased breast cancer risk such as Li-Fraumeni syndrome. Women from breast cancer families in which no clearly pathogenic mutation has been detected are classified based on family history in either moderate high (RR 2-3) or high risk (RR 3-4). Women at higher risk start screening at younger age and at higher frequency compared with women in lower risk categories. In addition, for women at a very high risk, risk reduction via prophylactic mastectomy or salpingo-oophorectomy is an option. Woman who carry a high-risk mutation might also opt for pre-implantation genetic diagnosis (PGD) also called embryo selection, to ensure that the predisposing mutation is not passed on to their children.

Given these options for prevention and early detection, it can make a dramatic difference for women from breast cancer families to know their mutational status, as women from families in which no causal mutation has been detected face much larger uncertainties. While in a hereditary breast cancer family mutation testing can identify the family members carrying the high risk variant, sisters within a familial breast cancer family are all assumed to have the same risk and all receive the same screening advice. In addition, the magnitude of the increased risk might be more uncertain. Risks for *BRCA1* and *BRCA2* mutation carriers have been determined based on studies with large numbers of carriers.[43–45] Although these risk estimates are likely to have been biased by the enrichement of carriers with a familial in these studies. In familial cancer clinics, risk in women from familial breast cancer families is presently estimated on the basis of family history alone; these estimates derive from large case-control analyses[32,33] and have been incorporated into models mentioned in paragraph 1.3. As the number of affected female relatives is an important variable in all commonly used risk prediction models, small families or families with few women are less informative and risk estimates might be less precise, and possibly underestimated in these cases. Consequently,

decisions on preventive measures are much more complicated in the presence of this uncertainty about risk.

2 The genetic landscape of breast cancer

2.1 High risk breast cancer susceptibility genes

Genetic variants associated with breast cancer risk are often subdivided into categories according to the magnitude of the risk increase and their population allele frequency. Although there is no generally accepted cut-off, genetic variants that are associated with a risk that is more than three times that of the general population are often considered high risk in clinical guidelines in the EU and the USA. This generally translates to a lifetime risk of higher than 30%. The two most well-known high-risk breast cancer susceptibility genes, *BRCA1* and *BRCA2*, were discovered after linkage analysis and subsequent sequencing in multi-case breast cancer families. Mutations in *BRCA1* and *BRCA2* are associated with a risk to develop breast cancer before the age of 70 of approximately 60% and 55%, respectively.[43–45] Both genes also give an increased risk of ovarian cancer with a 59% and 16% risk by age 70 for *BRCA1* and *BRCA2* respectively. In addition mutations in *BRCA1* and *BRCA2* have been linked to pancreatic[46–48] and prostate cancer.[49–51] The chance for a male mutation carrier to develop breast cancer before age 70 is approximately 1.2% for *BRCA1* and 6.8% for *BRCA2*.[52] Interestingly, while tumors arising in *BRCA1* mutation carriers are strongly enriched for the "triple negative" (lacking the receptors ER, PR and HER2) phenotype as well as basal-like molecular subtype expression profiles, *BRCA2* associated tumors are much more heterogeneous[53] and akin to "sporadic" breast cancer. Yet, both *BRCA1* and *BRCA2* play a crucial role in repair of DNA double-strand breaks via homologous recombination. In addition, *BRCA1* is involved in several cell cycle checkpoints that prevent a cell with DNA damage from entering mitosis.[54] Consequently, mutations in *BRCA1* and *BRCA2* are thought to contribute to tumorigenesis via accelerated accumulation of somatic mutations.

More recently, *PALB2* has been established as a breast cancer susceptibility gene.[55,56] The risk of developing breast cancer before the age of 70 for a carrier of a protein-truncating PALB2 mutation is approximately 35%. Moreover, in the context of a family history with two affected first-degree relatives the life-time risk for breast cancer is increased to 58%. In addition to increasing the risk of breast cancer, protein-truncating variants (PTV) in *PALB2* are associated with an increased pancreatic cancer risk.[57] Like mutations in *BRCA2*[58] and specific, more moderate-risk, mutations in *BRCA1*,[59] homozygous or compound-heterozygous mutations in *PALB2* cause Fanconi Anaemia,[60,61] a disease characterized by bone marrow failure, congenital anomalies and predisposition to several malignancies. Interestingly, PALB2 binds directly to both BRCA1 and BRCA2[62] and, like its binding partners, is involved in homologous recombination.

In addition to variants in these three genes, variants associated with a high risk for developing breast cancer risk can be found in *CDH1*,[63–65] *PTEN*,[66] *STK11*[67,68] and *TP53*.[69] In contrast to the variants in *BRCA1*, *BRCA2* and *PALB2*, variants in these genes cause cancer syndromes, defined by an increased risk for multiple types of cancer and sometimes other phenotypical features. It is therefore rare to find mutations in these genes in families with only an increased occurrence of breast cancer.[70–74] Due to the relative low numbers of patients with these syndromes, risk estimates vary widely, but all of these genes are generally regarded as high-risk breast cancer susceptibility genes. Biologically, these genes play very diverse roles in the cell. *TP53* is the most frequently somatically mutated gene in human cancer.

Via the regulation of expression of several genes it can, among others, activate DNA repair, arrest the cell cycle or induce apoptosis upon DNA damage.[75] PTEN is a negative regulator survival and proliferation via inhibition of the PI3K/AKT pathway.[76] While STK11 and CDH1 are important for cell polarity, cell-cell adhesion and energy metabolism, thereby suppressing cell proliferation and migration.[77–80]

2.2 Moderate risk breast cancer susceptibility genes

A second category of risk alleles consists of variants in moderate risk genes associated with a two- to four-fold increased risk. Well established moderate risk variants have been found in *ATM*,[81] *CHEK2*,[82] *BARD1*,[83–85] *FANCM*,[86,87] *NBS1*[88] and *RECQL*.[89,90] Variants in these genes can be relatively common in the general population. This is for example the case for the c.1100delC variant in *CHEK2*, which has an allele frequency of ~1% in north-western Europe.[82] However, for most of these genes pathogenic variants are still very rare, which makes estimating risks challenging. Moreover, it has been shown that breast cancer risks can vary between variants within the same gene. For example, while most variants in *ATM* are thought to be associated with an approximately 2- to 3-fold increased risk, a specific missense variant, c.7271T>G (p.V2424G), has been associated with risks 8-10 times as high as in the general population.[91–93] This strongly increased risk is probably caused by a dominant-negative effect of this specific variant on ATM protein function. [94]

The relatively moderate increase in risk associated with genetic variants in these genes, also makes the link with a family history of breast cancer less strong. Contrarily to the high-risk genes discussed above, most of these moderate risk genes have therefore not been identified via family-based linkage analysis. *ATM* and *NBS1*, which cause the rare recessive disorders ataxia-telangiectasia and Nijmegen breakage syndrome respectively, have been identified as breast cancer susceptibility genes because first-degree relatives of patients diagnosed with these syndromes had a markedly increased incidence of breast cancer. For the other moderate risk genes an increased breast cancer risk was observed after "candidate-gene" re-sequencing, meaning that genes with a function similar to that of *BRCA1* and *BRCA2* have selectively been tested for an association in either family-based or case-control studies. It is therefore not surprising that all of these genes have a function in DNA damage response because this strongly determined their candidacy. In addition to the genes mentioned above, there are a number of genes for which a link with breast cancer has been suggested, but not yet convincingly established. These include: *FAM175A*,[95] *MEN1*,[96] *MRE11A*,[97,98] *MSH6*,[85,99,100] *NF1*,[85,101] *RAD50*,[98,102–104] *RAD51C*,[85,99,105–110] *RAD51D*,[85,99,111,112] *RINT1*[113] and *XRCC2*.[114–116]

Until recently, moderate-risk genes were not routinely assessed in clinical practice. Therefore, the exact contribution of variants in these genes to familial breast cancer remained uncertain. Lately, two large studies reporing the results of gene panel testing of (familial) breast cancer cases found that mutations in these moderate risk genes are present in approximately 4% of the tested breast cancer patients.[85,117] Due to the relative rarity of these variants no specific guidelines for clinical management of carriers exist for the moderate risk genes other than *CHEK2*. Therefore, the breast cancer risk for an individual carrier has to be estimated from the family history and the risk associated with the mutation (for *ATM*, *CHEK2* and *PALB2* this can now also be done using BOADICEA[118]). After this, the guidelines for familial breast cancer can be followed.

2.3 Low risk breast cancer susceptibility alleles

A last category of breast cancer susceptibility alleles is formed by those associated with low RRs, usually between 0.7 and 1.3. Currently there are more than 300 single nucleotide polymorphisms (SNPs) found to be associated with breast cancer overall (i.e., irrespective of subtype), while a handful of variants are associated with specific subtypes of breast cancer.[119] This type of susceptibility alleles can be very common, with most risk alleles being present in more than five percent of the general population. Virtually all low risk genetic variants have been identified with so-called genome wide association studies (GWAS) studies. These are large case-control studies in which hundreds of thousands of single nucleotide polymorphisms are genotyped in each individual. Associations with breast cancer are calculated not only for the directly genotyped SNPs, but also for SNPs in close proximity that are "tagged" due to linkage disequilibrium and can therefore be imputed on the basis of a haplotype reference panel. This linkage disequilibrium only exists between SNPs with similar allele frequencies. Most SNP arrays are therefore not suitable to detect the risk associated with rare variants with allele frequencies <5%.

The linkage disequilibrium between SNPs also makes it challenging to pinpoint the causal variant in a region with multiple associated SNPs. So-called "fine-mapping" studies, often combined with functional in vitro assays of the variant, try to discover the causal variant by genotyping a larger number of variants within a susceptibility locus in more ethnically diverse populations.[120] Interestingly, most low risk genetic variants known today are not located in the protein-encoding regions of the genome. Low-risk variants for which a mechanism has been unraveled are often present in regulatory regions and affect gene expression. Such effects have for example been shown for a locus on 11q13. In this region two functional SNPs were uncovered: one reducing the binding of transcription factor *ELK4* to an enhancer region, the other increasing binding of GATA3 to a silencer. The risk alleles of both SNPs reduced transcriptional activity of *CCND1*.[121] Although *CCND1* has traditionally been considered to be a oncogene, it also promotes the recruitment of RAD51 to double strand breaks and reduction of CCND1 levels inhibits homologous recombination.[122] More complex mechanisms are likely to be revealed, as single susceptibility loci can affect multiple genes in both cis and trans, and not only in the tissue from which the tumor arises.[123] This is in line with the complex interactions between cell types within a tumor mass during tumorigenesis (see 1.1).

Due to the large number of low-risk variants and the small risks associated with them, the genotype of an individual for a single SNP has very little positive prediction power for the occurrence of breast cancer. Currently, these low-risk variants do not have any clinical implications. However, several studies have shown that these SNPs can be combined multiplicatively into a single risk score which predicts risk in a much more discriminatory way.[124,125] A recent modeling study has suggested that tailoring population-based screening programs for breast cancer based on polygenic risk scores could improve the cost-effectiveness and benefit-to-harm ratio of such programs.[126] Several ongoing clinical trials are assessing if screening strategies can be improved by considering an individual's genetic risk (e.g. the WISDOM study (NCT02620852) and the MyPeBS study (NCT03672331)). In general, it would be valuable to integrate the low-risk variants with the rare moderate and high risk variants and non-genetic factors into risk-prediction models so that their accuracy to predict breast cancer increases. It has already been shown that low-risk variants can be combined in to a model with moderate risk gene *CHEK2*.[127] The effect of these low risk variants in *BRCA1* or *BRCA2* mutation carriers has also been studied extensively. Many of the genetic variants that

are associated with small increases in breast cancer in the general population, have similar effects in carriers.[128] Moreover, several polymorphism have been identified that specifically modify risk for *BRCA1* and *BRCA2* mutation carriers.[129]

3. The missing heritability of breast cancer and the potential role of next generation sequencing

3.1 What type of genetic variants could explain the missing heritability

Despite countless studies in very large numbers of familial breast cancer cases and controls, more than half of the FRR remains unexplained. Of note, the techniques employed for the discovery of breast cancer susceptibility alleles have had a substantial impact on the types of genetic variants that studies have been able to discover. Genetic linkage studies, which statistically weigh co-segregation of genomic regions with disease within families, are only able to detect regions harboring high-risk breast cancer alleles in extended "informative" pedigrees. GWAS on the other hand, by nature of their design are limited to susceptibility alleles that are relatively common but can detect weak associations of small increased risk. They perform very well in situations of genetic heterogeneity. Most SNP arrays genotype and tag SNPs with a minor allele frequency of 5% or higher in the general population. The number of cases and controls in a study, together with the MAF of a SNP dictate the effect sizes and allele frequencies that can be detected. The largest GWAS study to date has included 122,977 cases and 105,974 controls and was able to detect significant ORs as low as 1.03 and associations with SNPs with a MAF of 1%.[130] However, the MAF of variants associated with a larger increase in risk (OR >2) are typically much lower than 1% and can therefore not be detected with the current GWAS and imputation strategies. Candidate gene re-sequencing studies are able to detect moderate and high-risk breast cancer alleles. The statistical power of this approach is strongly determined by the number of genes tested and the sample-size. Selecting familial cases, in which breast cancer risk alleles are thought to be enriched, and comparing the allele frequency in this population with that in population controls, allows for the dection of moderate and high risk alleles with population frequencies of less than 0.5%. Moreover, the focus on a single gene reduces issues with multiple testing and lowered p-value cut-offs to correct for that. In this way, moderate risks genes can be detected, without the need for very large studies. A clear downside, however, is that selecting candidates depends heavily on our assumptions on the pathways and genes involved in breast cancer susceptibility.

After the eras of linkage studies, candidate gene re-sequencing and GWAS, two main unexplored areas of genetic variation remained in which the missing heritability of breast cancer might reside. First of all, additional risk alleles might be found among the very rare genetic variants. The allele frequency cut-off for variants not detectable by the studies published to date, ranges from smaller than ~1% for low-risk polymorphisms to smaller than ~0.01% combined frequency in areas with high-risk alleles. In order to explore this area, larger case-control studies are needed. In addition, selecting phenotypically more homogenous groups of (familial) cases, might increase the efficiency of the study by also selecting for a more homogeneous genetic etiology. The second main unexplored area consists of moderate-risk variants in regions currently not linked with pathways involved in DNA damage repair. To explore this area, we need to be able to detect relatively rare (>0.1%) and potentially novel variants in all gene-coding regions of the genome. Moreover, this needs to be done fast and cost-effectively enough to be able to assess multiple cases from multiple unexplained breast

cancer families. Also for these variants, it will be important to select phenotypically more homogenous groups of cases to increase the efficiency of the study.

3.2 Challenges in the use of NGS to discover novel breast cancer risk alleles

By sequencing millions of short DNA fragments in parallel, NGS makes it possible to analyze large parts of the genome in a timely manner. However, although costs of NGS continue to decline, they still warrant carefully designed studies to increase the cost-effectiveness. One common choice is to focus on the protein coding regions of the genome, also known as the exome. This still enables searching for new breast cancer susceptibility genes without first narrowing down the list of genes of interest, allowing for a more agnostic view on which genes might be involved in breast cancer susceptibility. However, by focusing on the protein encoding regions it becomes more difficult to accurately detect copy number and structural variation, because of this many studies limit themselves to single nucleotide variants and small insertions and deletions. Moreover, it is usually not possible to assess the association of genetic variants with breast cancer within the limited number of individuals that have been exome sequenced. Therefore, many studies opt for a two stage design, where potentially interesting variants are discovered in a relatively small number of families using NGS, after which an association with breast cancer is assessed in a much larger group of familial cases and controls employing different genotyping techniques. However, the relatively small number of familial cases available and the need to control for multiple testing in the statistical analysis, make that only a hand full of variants can be tested with sufficient power in the second phase of the study. Given that, on average, an exome of a person from European descent contains ~12,000 non-synonymous variants,[131] selection of potentially interesting variants is not a trivial task and will often depend on assumptions about the features of a causal variant.

When narrowing down the list of potentially interesting genetic variants, there are several characteristics that can be taken into account. A common selection factor is to assess the severity of the variant, assuming that variants leading to a completely inactivated protein are more likely to cause disease than variants that only partially impair protein function, and that these variants are in turn more likely to cause disease than synonymously coding variants. Thus PTVs are generally considered pathogenic whereas the severity of missense variants is commonly assessed by in silico prediction tools, which take into account information on factors such as evolutionary conservation, known functional protein-domains, three-dimensional structure and the characteristic of the changed amino acid. These tools, however, have various limitations. For example, a variant that truly affects an exonic splice enhancer is likely to be wrongly classified as "benign" because of our limited abilities to predict this effect. Another commonly assessed variant characteristic is the allele frequency of a variant in reference datasets containing data from the general population. For example, if a variant is relatively frequent in the general population it cannot be associated with a high risk of breast cancer. Contrarily, if a variant is detected in multiple families within a study, while being rare or absent in the general population, this can be an indication to select a variant for further analysis. After the selection of a list of potentially causal variants, the next step depends on the availability of additional DNA samples. If the family of the index case has been extensively sampled, co-segregation analysis is a powerful way to assess the association with breast cancer. However, often only a limited number of family members can be assessed causing this analysis to be inconclusive. In this case, additional (familial) cases and controls will need to be genotyped.

4 Scope and outline of this thesis

This thesis aims to contribute to our understanding of the genetic etiology of breast cancer with the help of next generation sequencing. It will focus on families with a clear clustering of breast cancer, but in which no mutations in *BRCA1* or *BRCA2* have been detected. This thesis intends to give new insights into the genetic factors that are responsible for the clustering of breast cancer and provide clues for a better risk prediction in these families.

Chapter 2 of this thesis describes an exome sequencing effort in six breast cancer families. In order to select a potentially more homogeneous set of breast cancer cases, families were selected in which multiple women had tumor that showed a specific array CGH profile.

Chapter 3 reports on a large international case-control study that aimed to validate *XRCC2* as a new breast cancer susceptibility gene. This gene had recently been discovered using exome sequencing.

Chapter 4 reports on the functional analysis of missense variants detected in *XRCC2*. By selecting those variants that affect XRCC2 function, a more accurate burden analysis could be performed on the data of the case-control study described in chapter 3.

Chapter 5 describes an exome sequencing effort in families with an potential recessive mode of inheritance. All families selected for this study had at least three siblings affected with breast cancer and no breast cancer cases in first degree relatives from the previous of following generation. This study combined the exome sequencing results with haplotype-sharing data to more efficiently filter the genetic variants.

Chapter 6 is a review of the literature on the methods to determine the role of extremely rare genetic variants in familial breast cancer.

Chapter 7 discusses the main findings of this thesis, their potential consequences, clinical implications and future directions.

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-674. doi:10.1016/j.cell.2011.02.013
2. Vineis P, Schatzkin A, Potter JD. Models of carcinogenesis: an overview. Carcinogenesis. 2010;31(10):1703-1709. doi:10.1093/carcin/bgq087
3. Shen H, Laird PW. Interplay between the cancer genome and epigenome. Cell. 2013;153(1):38-55. doi:10.1016/j.cell.2013.03.008
4. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: biology and implications for treatment. Breast Cancer Res BCR. 2011;13(6):227. doi:10.1186/bcr2912
5. Forbes SA, Beare D, Gunasekaran P, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. 2015;43(Database issue):D805-811. doi:10.1093/nar/gku1075
6. Stephens PJ, Tarpey PS, Davies H, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012;486(7403):400-404. doi:10.1038/nature11017
7. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61-70. doi:10.1038/nature11412
8. Banerji S, Cibulskis K, Rangel-Escareno C, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature. 2012;486(7403):405-409. doi:10.1038/nature11154
9. Dorman SN, Viner C, Rogan PK. Splicing mutation analysis reveals previously unrecognized pathways in lymph node-invasive breast cancer. Sci Rep. 2014;4:7063. doi:10.1038/srep07063
10. Ellis MJ, Ding L, Shen D, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature. 2012;486(7403):353-360. doi:10.1038/nature11143
11. Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature. 2012;486(7403):395-399. doi:10.1038/nature10933
12. Stefansson OA, Esteller M. Epigenetic modifications in breast cancer and their role in personalized medicine. Am J Pathol. 2013;183(4):1052-1063. doi:10.1016/j.ajpath.2013.04.033
13. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007;128(4):683-692. doi:10.1016/j.cell.2007.01.029
14. Hu M, Yao J, Carroll DK, et al. Regulation of in situ to invasive breast carcinoma transition. Cancer Cell. 2008;13(5):394-406. doi:10.1016/j.ccr.2008.03.007
15. SR L, IO E, SJ S, PH T, MJ van de V. WHO Classification of Tumours of the Breast. http://publications.iarc.fr/Book-And-Report-Series/Who-Iarc-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Breast-2012. Accessed September 24, 2018.
16. Rakha EA, Patel A, Powe DG, et al. Clinical and biological significance of E-cadherin protein expression in invasive lobular carcinoma of the breast. Am J Surg Pathol. 2010;34(10):1472-1479. doi:10.1097/PAS.0b013e3181f01916
17. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn H-J. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol. 2011;22(8):1736-1747. doi:10.1093/annonc/mdr304
18. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 2005;353(16):1659-1672. doi:10.1056/NEJMoa052306
19. Katz H, Alsharedi M. Immunotherapy in triple-negative breast cancer. Med Oncol Northwood Lond Engl. 2017;35(1):13. doi:10.1007/s12032-017-1071-6
20. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. Mol Oncol. 2011;5(1):5-23. doi:10.1016/j.molonc.2010.11.003
21. Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res BCR. 2010;12(5):R68.

doi:10.1186/bcr2635
22. Farmer P, Bonnefoi H, Becette V, et al. Identification of molecular apocrine breast tumours by microarray analysis. Oncogene. 2005;24(29):4660-4671. doi:10.1038/sj.onc.1208561
23. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature. 2016;534(7605):47-54. doi:10.1038/nature17676
24. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424. doi:10.3322/caac.21492
25. IKNL. Cijfers over kanker, Nederlandse Kankerregistratie. http://www.cijfersoverkanker.nl/. Accessed July 23, 2017.
26. Dumitrescu RG, Cotarla I. Understanding breast cancer risk -- where do we stand in 2005? J Cell Mol Med. 2005;9(1):208-221.
27. DePinho RA. The age of cancer. Nature. 2000;408(6809):248-254. doi:10.1038/35041694
28. Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. Lancet Oncol. 2012;13(11):1141-1151. doi:10.1016/S1470-2045(12)70425-4
29. Atashgaran V, Wrin J, Barry SC, Dasari P, Ingman WV. Dissecting the Biology of Menstrual Cycle-Associated Breast Cancer Risk. Front Oncol. 2016;6:267. doi:10.3389/fonc.2016.00267
30. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. Lancet Lond Engl. 2002;360(9328):187-195. doi:10.1016/S0140-6736(02)09454-0
31. Dall G, Risbridger G, Britt K. Mammary stem cells and parity-induced breast cancer protection- new insights. J Steroid Biochem Mol Biol. 2017;170:54-60. doi:10.1016/j.jsbmb.2016.02.018
32. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer. 1997;71(5):800-809.
33. Collaborative Group on Hormonal Factors in Breast Cancer. 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 Lond Engl. 2001;358(9291):1389-1399. doi:10.1016/S0140-6736(01)06524-2
34. Antoniou AC, Pharoah PDP, McMullan G, et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. Br J Cancer. 2002;86(1):76-83. doi:10.1038/sj.bjc.6600008
35. Parmigiani G, Berry D, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. Am J Hum Genet. 1998;62(1):145-158.
36. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med. 2004;23(7):1111-1130. doi:10.1002/sim.1668
37. Claus EB, Risch N, Thompson WD. Genetic analysis of breast cancer in the cancer and steroid hormone study. Am J Hum Genet. 1991;48(2):232-242.
38. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst. 1989;81(24):1879-1886.
39. STOET, VKGN. Landelijke Richtlijn Familiair Mamma/Ovariumcarcinoom.; 2010. http://www.oncoline.nl/familiair-mamma-ovariumcarcinoom.
40. STOET. Erfelijke En Familiaire Tumoren: Richtlijnen Voor Diagnostiek En Preventie.; 2017. https://www.stoet.nl/wp-content/uploads/2017/04/STOET-Richtlijnenboekje-april2017_DEF.pdf.
41. NABON. Landelijke Richtlijn Mammacarcinoom.; 2017. http://www.oncoline.nl/mammacarcinoom.
42. STOET, VKGN. Landelijke Richtlijn Hereditair Mamma/Ovariumcarcinoom.; 2010. http://www.

oncoline.nl/hereditair-mamma-ovariumcarcinoom.
43. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. J Clin Oncol Off J Am Soc Clin Oncol. 2007;25(11):1329-1333. doi:10.1200/JCO.2006.09.1066
44. Mavaddat N, Peock S, Frost D, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers: results from prospective analysis of EMBRACE. J Natl Cancer Inst. 2013;105(11):812-822. doi:10.1093/jnci/djt095
45. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA. 2017;317(23):2402-2416. doi:10.1001/jama.2017.7112
46. Murphy KM, Brune KA, Griffin C, et al. Evaluation of candidate genes MAP2K4, MADH4, ACVR1B, and BRCA2 in familial pancreatic cancer: deleterious BRCA2 mutations in 17%. Cancer Res. 2002;62(13):3789-3793.
47. Hahn SA, Greenhalf B, Ellis I, et al. BRCA2 germline mutations in familial pancreatic carcinoma. J Natl Cancer Inst. 2003;95(3):214-221.
48. Couch FJ, Johnson MR, Rabe KG, et al. The prevalence of BRCA2 mutations in familial pancreatic cancer. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2007;16(2):342-346. doi:10.1158/1055-9965.EPI-06-0783
49. Lecarpentier J, Silvestri V, Kuchenbaecker KB, et al. Prediction of Breast and Prostate Cancer Risks in Male BRCA1 and BRCA2 Mutation Carriers Using Polygenic Risk Scores. J Clin Oncol Off J Am Soc Clin Oncol. 2017;35(20):2240-2250. doi:10.1200/JCO.2016.69.4935
50. Thompson D, Easton DF, Breast Cancer Linkage Consortium. Cancer Incidence in BRCA1 mutation carriers. J Natl Cancer Inst. 2002;94(18):1358-1365.
51. Breast Cancer Linkage Consortium. Cancer risks in BRCA2 mutation carriers. J Natl Cancer Inst. 1999;91(15):1310-1316.
52. Tai YC, Domchek S, Parmigiani G, Chen S. Breast cancer risk among male BRCA1 and BRCA2 mutation carriers. J Natl Cancer Inst. 2007;99(23):1811-1814. doi:10.1093/jnci/djm203
53. Mavaddat N, Barrowdale D, Andrulis IL, et al. Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2012;21(1):134-147. doi:10.1158/1055-9965.EPI-11-0775
54. Roy R, Chun J, Powell SN. BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat Rev Cancer. 2011;12(1):68-78. doi:10.1038/nrc3181
55. Rahman N, Seal S, Thompson D, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat Genet. 2007;39(2):165-167. doi:10.1038/ng1959
56. Antoniou AC, Casadei S, Heikkinen T, et al. Breast-cancer risk in families with mutations in PALB2. N Engl J Med. 2014;371(6):497-506. doi:10.1056/NEJMoa1400382
57. Jones S, Hruban RH, Kamiyama M, et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. Science. 2009;324(5924):217. doi:10.1126/science.1171202
58. Howlett NG, Taniguchi T, Olson S, et al. Biallelic inactivation of BRCA2 in Fanconi anemia. Science. 2002;297(5581):606-609. doi:10.1126/science.1073834
59. Sawyer SL, Tian L, Kähkönen M, et al. Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. Cancer Discov. 2015;5(2):135-142. doi:10.1158/2159-8290.CD-14-1156
60. Xia B, Dorsman JC, Ameziane N, et al. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. Nat Genet. 2007;39(2):159-161. doi:10.1038/ng1942
61. Reid S, Schindler D, Hanenberg H, et al. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. Nat Genet. 2007;39(2):162-164. doi:10.1038/ng1947
62. Sy SMH, Huen MSY, Chen J. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. Proc Natl Acad Sci U S A. 2009;106(17):7155-7160. doi:10.1073/pnas.0811159106
63. Guilford P, Hopkins J, Harraway J, et al. E-cadherin germline mutations in familial gastric cancer. Nature. 1998;392(6674):402-405. doi:10.1038/32918

64. Keller G, Vogelsang H, Becker I, et al. Diffuse type gastric and lobular breast carcinoma in a familial gastric cancer patient with an E-cadherin germline mutation. Am J Pathol. 1999;155(2):337-342. doi:10.1016/S0002-9440(10)65129-2
65. Pharoah PD, Guilford P, Caldas C, International Gastric Cancer Linkage Consortium. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. Gastroenterology. 2001;121(6):1348-1353.
66. Nelen MR, van Staveren WC, Peeters EA, et al. Germline mutations in the PTEN/MMAC1 gene in patients with Cowden disease. Hum Mol Genet. 1997;6(8):1383-1387.
67. Hemminki A, Markie D, Tomlinson I, et al. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. Nature. 1998;391(6663):184-187. doi:10.1038/34432
68. Jenne DE, Reimann H, Nezu J, et al. Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. Nat Genet. 1998;18(1):38-43. doi:10.1038/ng0198-38
69. Malkin D, Li FP, Strong LC, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science. 1990;250(4985):1233-1238.
70. FitzGerald MG, Marsh DJ, Wahrer D, et al. Germline mutations in PTEN are an infrequent cause of genetic predisposition to breast cancer. Oncogene. 1998;17(6):727-731. doi:10.1038/sj.onc.1201984
71. Ginsburg OM, Akbari MR, Aziz Z, et al. The prevalence of germ-line TP53 mutations in women diagnosed with breast cancer before age 30. Fam Cancer. 2009;8(4):563-567. doi:10.1007/s10689-009-9287-z
72. Guénard F, Pedneault CS-L, Ouellette G, et al. Evaluation of the contribution of the three breast cancer susceptibility genes CHEK2, STK11, and PALB2 in non-BRCA1/2 French Canadian families with high risk of breast cancer. Genet Test Mol Biomark. 2010;14(4):515-526. doi:10.1089/gtmb.2010.0027
73. Rapakko K, Allinen M, Syrjäkoski K, et al. Germline TP53 alterations in Finnish breast cancer families are rare and occur at conserved mutation-prone sites. Br J Cancer. 2001;84(1):116-119. doi:10.1054/bjoc.2000.1530
74. Schrader KA, Masciari S, Boyd N, et al. Germline mutations in CDH1 are infrequent in women with early-onset or familial lobular breast cancers. J Med Genet. 2011;48(1):64-68. doi:10.1136/jmg.2010.079814
75. Bieging KT, Mello SS, Attardi LD. Unravelling mechanisms of p53-mediated tumour suppression. Nat Rev Cancer. 2014;14(5):359-370. doi:10.1038/nrc3711
76. Song MS, Salmena L, Pandolfi PP. The functions and regulation of the PTEN tumour suppressor. Nat Rev Mol Cell Biol. 2012;13(5):283-296. doi:10.1038/nrm3330
77. Shackelford DB, Shaw RJ. The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. Nat Rev Cancer. 2009;9(8):563-575. doi:10.1038/nrc2676
78. Li J, Liu J, Li P, et al. Loss of LKB1 disrupts breast epithelial cell polarity and promotes breast cancer metastasis and invasion. J Exp Clin Cancer Res CR. 2014;33:70. doi:10.1186/s13046-014-0070-0
79. Bays JL, Campbell HK, Heidema C, Sebbagh M, DeMali KA. Linking E-cadherin mechanotransduction to cell metabolism through force-mediated activation of AMPK. Nat Cell Biol. 2017;19(6):724-731. doi:10.1038/ncb3537
80. Chen A, Beetham H, Black MA, et al. E-cadherin loss alters cytoskeletal organization and adhesion in non-malignant breast cells but is insufficient to induce an epithelial-mesenchymal transition. BMC Cancer. 2014;14:552. doi:10.1186/1471-2407-14-552
81. Renwick A, Thompson D, Seal S, et al. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. Nat Genet. 2006;38(8):873-875. doi:10.1038/ng1837
82. Meijers-Heijboer H, van den Ouweland A, Klijn J, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet. 2002;31(1):55-59. doi:10.1038/ng879
83. De Brakeleer S, De Grève J, Loris R, et al. Cancer predisposing missense and protein truncating BARD1 mutations in non-BRCA1 or BRCA2 breast cancer families. Hum Mutat.

2010;31(3):E1175-1185. doi:10.1002/humu.21200
84. Ratajska M, Antoszewska E, Piskorz A, et al. Cancer predisposing BARD1 mutations in breast-ovarian cancer families. Breast Cancer Res Treat. 2012;131(1):89-97. doi:10.1007/s10549-011-1403-8
85. Couch FJ, Shimelis H, Hu C, et al. Associations Between Cancer Predisposition Testing Panel Genes and Breast Cancer. JAMA Oncol. 2017;3(9):1190-1196. doi:10.1001/jamaoncol.2017.0424
86. Kiiski JI, Pelttari LM, Khan S, et al. Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. Proc Natl Acad Sci U S A. 2014;111(42):15172-15177. doi:10.1073/pnas.1407909111
87. Peterlongo P, Catucci I, Colombo M, et al. FANCM c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity and is a familial breast cancer risk factor. Hum Mol Genet. 2015;24(18):5345-5355. doi:10.1093/hmg/ddv251
88. Steffen J, Nowakowska D, Niwińska A, et al. Germline mutations 657del5 of the NBS1 gene contribute significantly to the incidence of breast cancer in Central Poland. Int J Cancer. 2006;119(2):472-475. doi:10.1002/ijc.21853
89. Cybulski C, Carrot-Zhang J, Kluźniak W, et al. Germline RECQL mutations are associated with breast cancer susceptibility. Nat Genet. 2015;47(6):643-646. doi:10.1038/ng.3284
90. Sun J, Wang Y, Xia Y, et al. Mutations in RECQL Gene Are Associated with Predisposition to Breast Cancer. PLoS Genet. 2015;11(5):e1005228. doi:10.1371/journal.pgen.1005228
91. Southey MC, Goldgar DE, Winqvist R, et al. PALB2, CHEK2 and ATM rare variants and cancer risk: data from COGS. J Med Genet. 2016;53(12):800-811. doi:10.1136/jmedgenet-2016-103839
92. Bernstein JL, Teraoka S, Southey MC, et al. Population-based estimates of breast cancer risks associated with ATM gene variants c.7271T>G and c.1066-6T>G (IVS10-6T>G) from the Breast Cancer Family Registry. Hum Mutat. 2006;27(11):1122-1128. doi:10.1002/humu.20415
93. Goldgar DE, Healey S, Dowty JG, et al. Rare variants in the ATM gene and risk of breast cancer. Breast Cancer Res BCR. 2011;13(4):R73. doi:10.1186/bcr2919
94. Waddell N, Jonnalagadda J, Marsh A, et al. Characterization of the breast cancer associated ATM 7271T>G (V2424G) mutation by gene expression profiling. Genes Chromosomes Cancer. 2006;45(12):1169-1181. doi:10.1002/gcc.20381
95. Solyom S, Aressy B, Pylkäs K, et al. Breast cancer-associated Abraxas mutation disrupts nuclear localization and DNA damage response functions. Sci Transl Med. 2012;4(122):122ra23. doi:10.1126/scitranslmed.3003223
96. Dreijerink KMA, Goudet P, Burgess JR, Valk GD. Breast-Cancer Predisposition in Multiple Endocrine Neoplasia Type 1. N Engl J Med. 2014;371(6):583-584. doi:10.1056/NEJMc1406028
97. Bartkova J, Tommiska J, Oplustilova L, et al. Aberrations of the MRE11-RAD50-NBS1 DNA damage sensor complex in human breast cancer: MRE11 as a candidate familial cancer-predisposing gene. Mol Oncol. 2008;2(4):296-316. doi:10.1016/j.molonc.2008.09.007
98. Damiola F, Pertesi M, Oliver J, et al. Rare key functional domain missense substitutions in MRE11A, RAD50, and NBN contribute to breast cancer susceptibility: results from a Breast Cancer Family Registry case-control mutation-screening study. Breast Cancer Res BCR. 2014;16(3):R58. doi:10.1186/bcr3669
99. Lu H-M, Li S, Black MH, et al. Association of Breast and Ovarian Cancers With Predisposition Genes Identified by Large-Scale Sequencing. JAMA Oncol. August 2018. doi:10.1001/jamaoncol.2018.2956
100. Win AK, Lindor NM, Jenkins MA. Risk of breast cancer in Lynch syndrome: a systematic review. Breast Cancer Res BCR. 2013;15(2):R27. doi:10.1186/bcr3405
101. Seminog OO, Goldacre MJ. Age-specific risk of breast cancer in women with neurofibromatosis type 1. Br J Cancer. 2015;112(9):1546-1548. doi:10.1038/bjc.2015.78
102. Heikkinen K, Rapakko K, Karppinen S-M, et al. RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. Carcinogenesis. 2006;27(8):1593-

1599. doi:10.1093/carcin/bgi360
103. Tommiska J, Seal S, Renwick A, et al. Evaluation of RAD50 in familial breast cancer predisposition. Int J Cancer. 2006;118(11):2911-2916. doi:10.1002/ijc.21738
104. Uhrhammer N, Delort L, Bignon Y-J. Rad50 c.687delT does not contribute significantly to familial breast cancer in a French population. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2009;18(2):684-685. doi:10.1158/1055-9965.EPI-08-0971
105. Meindl A, Hellebrand H, Wiek C, et al. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nat Genet. 2010;42(5):410-414. doi:10.1038/ng.569
106. Clague J, Wilhoite G, Adamson A, Bailis A, Weitzel JN, Neuhausen SL. RAD51C germline mutations in breast and ovarian cancer cases from high-risk families. PloS One. 2011;6(9):e25632. doi:10.1371/journal.pone.0025632
107. De Leeneer K, Van Bockstal M, De Brouwer S, et al. Evaluation of RAD51C as cancer susceptibility gene in a large breast-ovarian cancer patient population referred for genetic testing. Breast Cancer Res Treat. 2012;133(1):393-398. doi:10.1007/s10549-012-1998-4
108. Loveday C, Turnbull C, Ruark E, et al. Germline RAD51C mutations confer susceptibility to ovarian cancer. Nat Genet. 2012;44(5):475-476; author reply 476. doi:10.1038/ng.2224
109. Thompson ER, Boyle SE, Johnson J, et al. Analysis of RAD51C germline mutations in high-risk breast and ovarian cancer families and ovarian cancer patients. Hum Mutat. 2012;33(1):95-99. doi:10.1002/humu.21625
110. Blanco A, Gutiérrez-Enríquez S, Santamariña M, et al. RAD51C germline mutations found in Spanish site-specific breast cancer and breast-ovarian cancer families. Breast Cancer Res Treat. 2014;147(1):133-143. doi:10.1007/s10549-014-3078-4
111. Chen X, Li Y, Ouyang T, et al. Associations between RAD51D germline mutations and breast cancer risk and survival in BRCA1/2-negative breast cancers. Ann Oncol Off J Eur Soc Med Oncol. August 2018. doi:10.1093/annonc/mdy338
112. Loveday C, Turnbull C, Ramsay E, et al. Germline mutations in RAD51D confer susceptibility to ovarian cancer. Nat Genet. 2011;43(9):879-882. doi:10.1038/ng.893
113. Park DJ, Tao K, Le Calvez-Kelm F, et al. Rare mutations in RINT1 predispose carriers to breast and Lynch syndrome-spectrum cancers. Cancer Discov. 2014;4(7):804-815. doi:10.1158/2159-8290.CD-14-0212
114. Park DJ, Lesueur F, Nguyen-Dumont T, et al. Rare mutations in XRCC2 increase the risk of breast cancer. Am J Hum Genet. 2012;90(4):734-739. doi:10.1016/j.ajhg.2012.02.027
115. Hilbers FS, Wijnen JT, Hoogerbrugge N, et al. Rare variants in XRCC2 as breast cancer susceptibility alleles. J Med Genet. 2012;49(10):618-620. doi:10.1136/jmedgenet-2012-101191
116. Hilbers FS, Luijsterburg MS, Wiegant WW, et al. Functional Analysis of Missense Variants in the Putative Breast Cancer Susceptibility Gene XRCC2. Hum Mutat. 2016;37(9):914-925. doi:10.1002/humu.23019
117. Kurian AW, Ward KC, Howlader N, et al. Genetic Testing and Results in a Population-Based Cohort of Breast Cancer Patients and Ovarian Cancer Patients. J Clin Oncol Off J Am Soc Clin Oncol. 2019;37(15):1305-1315. doi:10.1200/JCO.18.01854
118. Lee AJ, Cunningham AP, Tischkowitz M, et al. Incorporating truncating variants in PALB2, CHEK2, and ATM into the BOADICEA breast cancer risk model. Genet Med Off J Am Coll Med Genet. 2016;18(12):1190-1198. doi:10.1038/gim.2016.31
119. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. Am J Hum Genet. 2019;104(1):21-34. doi:10.1016/j.ajhg.2018.11.002
120. Rivandi M, Martens JWM, Hollestelle A. Elucidating the Underlying Functional Mechanisms of Breast Cancer Susceptibility Through Post-GWAS Analyses. Front Genet. 2018;9:280. doi:10.3389/fgene.2018.00280

121. French JD, Ghoussaini M, Edwards SL, et al. Functional variants at the 11q13 risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers. Am J Hum Genet. 2013;92(4):489-503. doi:10.1016/j.ajhg.2013.01.002
122. Jirawatnotai S, Hu Y, Michowski W, et al. A function for cyclin D1 in DNA repair uncovered by protein interactome analyses in human cancers. Nature. 2011;474(7350):230-234. doi:10.1038/nature10155
123. Ferreira MA, Gamazon ER, Al-Ejeh F, et al. Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer. Nat Commun. 2019;10. doi:10.1038/s41467-018-08053-5
124. Mavaddat N, Pharoah PDP, Michailidou K, et al. Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst. 2015;107(5). doi:10.1093/jnci/djv036
125. Muranen TA, Mavaddat N, Khan S, et al. Polygenic risk score is associated with increased disease risk in 52 Finnish breast cancer families. Breast Cancer Res Treat. 2016;158(3):463-469. doi:10.1007/s10549-016-3897-6
126. Pashayan N, Morris S, Gilbert FJ, Pharoah PDP. Cost-effectiveness and Benefit-to-Harm Ratio of Risk-Stratified Screening for Breast Cancer: A Life-Table Model. JAMA Oncol. July 2018. doi:10.1001/jamaoncol.2018.1901
127. Muranen TA, Greco D, Blomqvist C, et al. Genetic modifiers of CHEK2*1100delC-associated breast cancer risk. Genet Med Off J Am Coll Med Genet. October 2016. doi:10.1038/gim.2016.147
128. Kuchenbaecker KB, Neuhausen SL, Robson M, et al. Associations of common breast cancer susceptibility alleles with risk of breast cancer subtypes in BRCA1 and BRCA2 mutation carriers. Breast Cancer Res BCR. 2014;16(6):3416. doi:10.1186/s13058-014-0492-9
129. Couch FJ, Wang X, McGuffog L, et al. Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. PLoS Genet. 2013;9(3):e1003212. doi:10.1371/journal.pgen.1003212
130. Michailidou K, Lindström S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. Nature. October 2017. doi:10.1038/nature24284
131. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-291. doi:10.1038/nature19057