From the Department of Oncology-Pathology
Karolinska Institutet and the Department of Oncology,
Södersjukhuset, Stockholm, Sweden

# GENETIC STUDIES IN FAMILIAL NON-BRCA BREAST CANCER

Camilla Wendt

Karolinska
Institutet

Stockholm 2017

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by E-Print AB 2017
© Camilla Wendt, 2017
ISBN 978-91-7676-860-0

# GENETIC STUDIES IN FAMILIAL NON-BRCA BREAST CANCER
# THESIS FOR DOCTORAL DEGREE (Ph.D.)

By


## Camilla Wendt

*Principal Supervisor:*
Professor Annika Lindblom
Karolinska Institutet
Department of Molecular Medicine and Surgery

*Co-supervisor:*
M.D, Ph.D. Sara Margolin
Karolinska Institutet
Department of Clinical Science and Education, Södersjukhuset

*Opponent:*
Professor Olle Stål
University of Linköping
Department of Clinical and Experimental Medicine

*Examination Board:*
Professor Gunilla Enblad
University of Uppsala
Department of Immunology, Genetics, Pathology, Experimental and Clinical Oncology

M.D, Ph.D. Stefan Söderhäll
Karolinska Institutet
Department of Women´s and Children´s Health

Professor Johan Hansson
Karolinska Institutet
Department of Oncology-Pathology

Till min familj

# ABSTRACT

Family history is an important risk factor for breast cancer, the presence of breast cancer in a first degree relative in general nearly doubles the risk and the risk increases with the number of affected relatives. Pathogenic mutations in *BRCA1*, *BRCA2* and other high- and moderate risk-genes account for 25% of the familial risk for breast cancer. About 180 low-risk variants explain an additional 18% of the excess familial risk. The remainder of the genetic contribution to familial breast cancer is unexplained. A polygenic model, where pathogenic mutations with differential impact together confer an increased risk for breast cancer, has been suggested. The aim of this thesis has been to study and better understand how breast cancer is inherited and to identify underlying genetic factors that contribute to the risk in familial breast cancer without pathogenic mutations in *BRCA1* or *BRCA2 (*non-*BRCA* families*)*.

In paper I tumour spectrum was investigated in our cohort of non-*BRCA* families with at least 2 cases of breast cancer and one case of other tumour type in first-, second degree relatives or first cousins. Distribution of tumour types, other than breast cancer, was compared with the distribution in Sweden in two reference years. We found an overrepresentation of endometrial cancer in the non-*BRCA* families with a 6.36 % proportion (CI 4.67–8.2) compared to the proportion in the general population in the reference years 1970 (3.07 %) and 2010 (2.64 %). The main finding of the study was the strong support for a breast- and endometrial cancer syndrome, which is a first step towards detecting new susceptibility variants.

In paper II we investigated if breast cancer prognosis is affected by parent-of-origin in our cohort of non-*BRCA* families. A difference in prognosis may indicate an influence of a genetic mechanism that produces inter-lineage effects, such as genomic imprinting.
No significant difference in overall or recurrence-free survival between maternal and paternal inheritance of breast cancer was observed with HRs of 0.99 (95% CI=0.54 to 1.80) and 1.22 (95% CI=0.78 to 1.92) respectively. An interesting finding in paper II was the predominance of maternally inherited cases, which indicates that parent-of-origin may not have an effect on breast cancer prognosis, but rather the risk of being affected.

The protein truncating mutation *CHEK2*1100delC* is a moderate-risk variant associated with a 2-3 fold increased risk of developing breast cancer, but the risk is considerably higher in carriers with a family history. The individual risk for breast cancer in carriers of *CHEK2*1100delC* is thereby difficult to predict. In paper III we performed whole-exome sequencing in cases of *CHEK2*1100delC* carriers in search of genetic variants that may modify breast cancer risk in this patient group. All non-synonymous mutations were evaluated and 11 candidate alleles were selected and tested in a validation. No *CHEK2* specific modifier could be identified though, as none of the variants showed significant difference in allele frequency in *CHEK2*1100delC* carriers compared to controls. Continuous studies of genetic modifiers are of importance to improve breast cancer risk prediction for *CHEK2*1100delC* carriers.

# LIST OF SCIENTIFIC PAPERS

I. Wendt, C, Lindblom, A, Arver, B, von Wachenfeldt, A, Margolin, S.
*Tumour spectrum in non-BRCA hereditary breast cancer families in Sweden*
Hered Cancer Clin Pract, 2015. **13**(1): p. 15.

Wendt C, Margolin S.
*A Breast and endometrial cancer syndrome*,
Maturitas. 2016 May;87:3-4. doi: 10.1016/j.maturitas.2016.01.011. Epub 2016 Jan 22.

II. Wendt, C, Lindblom, A, Arver, B, von Wachenfeldt, A, Margolin, S.
*Parent of Origin and Prognosis in Familial Breast Cancer in Sweden.*
Anticancer Res, 2017. **37**(3): p. 1257-1262

III. Wendt, C, Thutkawkorapin, J, Jiao, X, Ehrencrona, H, Tham, E, Kvist, A, Arver, B, Melin, B, Kuchinskaya, E, Stenmark Askmalm, M, Paulsson-Karlsson, Y, Einbeigi, Z, von Wachenfeldt Väppling, A, Borg, Å, Lindblom, A *Exome sequencing in Swedish CHEK2*1100delC carriers*. Manuscript

## ADDITIONAL PAPERS

Jiao, Aravidis, Marikkannu, Rantala, Picelli, Adamovic, Liu, Maguire, Kremeyer, Luo, von Holst, Kontham, Thutkawkorapin, Margolin, Du, Lundin, Michalidou, Bolla, Wang, Dennis, Lush, Ambrosone, Andrulis, Anton-Culver, Antonenkoa, Arndt, Beckmann, Blomqvist, Blot, Boeckx, Bojesen, Bonanni, Brand, Brauch, Brenner, Broeks, Bruning, Burwinkel, Cai, Chang-Claude, NBCS Collaborators, Couch, Cox, Cross, Deming-Halveron, Devilee, dos-Santos-Silva, Dörk, Eriksson, Fashing, Figuroa, Flesch-Janys, Flyger, Gabrielson, Garcia-Closas, Giles, Gonzalez-Neira, Guenel, Guo, Gundert, Haiman, Hallberg, Hamann, Harrington, Hooning, Hopper, Huang, Jakubrowska, Jones, Kerin, Kosma, Kristensen, Lambrechts, Marchand, Lubinski, Mannermaa, Martens, Meindl, Milne, Mulligan, Neuhausen, Nevanlinna, Peto, Pylkäs, Radice, Rhenius, Sawyer, Schmidt, Schmutzler, Seynaeve, Shah, Simard, Southey, Swerdlow, Truoung, Wendt, Winqvist, Zheng, kConFab/AOCS Investigators, Benitez, Dunning, Pharoah, Easton, Czene, Hall, Lindblom.
*PHIH – a novel candidate breast cancer susceptibility locus on 6q14.1,* Oncotarget. Received: July 06, 2017, Accepted: August 31 2017, Published: October 12, 2017.

Michailidou, K et al. Association analysis identifies 65 new breast cancer risk loci. Nature; 23 Oct 2017; DOI: 10.1038/nature24284
Breast cancer risk is influenced by rare coding variants in susceptibility genes, such as BRCA1, and many common, mostly non-coding variants. However, much of the genetic contribution to breast cancer risk remains unknown. Here we report the results of a genome-wide association study of breast cancer in 122,977 cases and 105,974 controls of European ancestry and 14,068 cases and 13,104 controls of East Asian ancestry. We identified 65 new loci that are associated with overall breast cancer risk at $P < 5 \times 10^{-8}$. The majority of credible risk single-nucleotide polymorphisms in these loci fall in distal regulatory elements, and by integrating in silico data to predict target genes in breast cells at each locus, we demonstrate a strong overlap between candidate target genes and somatic driver genes in breast tumours. We also find that heritability of breast cancer due to all single-nucleotide polymorphisms in regulatory features was 2-5 fold enriched relative to the genome-wide average, with strong enrichment for particular transcription factor binding sites. These results proved further insight into genetic susceptibility to breast cancer and will improve the use of genetic risk score for individualised screening and prevention.

Milne, R et al*, Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer,* Nature Genetics. Received 30 May 2016; accepted 11 January 2017; published online 23 October 2017.
Most common breast cancer susceptibility variants have been identified through genome-wide association studies (GWAS) of predominantly estrogen receptor (ER)-positive disease. We conducted a GWAS using 21,468 ER-negative cases and 100,594 controls combined with 18,908 BRCA1 mutation carriers (9,414 with breast cancer), all of European origin. We identified independent associations at $P < 5 \times 10^{-8}$ with ten variants at nine new loci. At $P < 0.05$, we replicated associations with 10 of 11 variants previously reported in ER-negative disease or BRCA1 mutation carrier GWAS and observed consistent associations with ER-negative disease for 105 susceptibility variants identified by other studies. These 125 variants explain approximately 16% of the familial risk of this breast cancer subtype. There was high genetic correlation (0.72) between risk of ER-negative breast cancer and breast cancer risk for BRCA1 mutation carriers. These findings may lead to improved risk prediction and inform further fine-mapping and functional work to better understand the biological basis of ER-negative breast cancer.

# CONTENTS

1 INTRODUCTION ..... 1
1.1 EPIDEMIOLOGICAL STUDIES ..... 3
1.2 IDENTIFYING BREAST CANCER SUSCEPTIBILITY GENES ..... 3
1.2.1 High-risk variants ..... 3
1.2.2 Moderate-risk variants ..... 4
1.2.3 Common low-risk variants ..... 5
1.3 EVIDENCE OF ASSOCIATION FOR SPECIFIC GENES AND VARIANTS ..... 6
1.3.1 High-risk genes ..... 7
1.3.2 Moderate-risk genes ..... 9
1.3.3 Other breast cancer gene susceptibility candidates ..... 12
1.3.4 Low penetrance variants ..... 12
1.4 GENETIC CONTRIBUTION IN BREAST CANCER ..... 13
1.5 GENETIC MODIFIERS ..... 14
1.6 OTHER GENETIC RISK FACTORS ..... 15
1.6.1 Epigenetic regulation ..... 15
1.6.2 Copy Number Variation ..... 17
1.7 BREAST CANCER SYNDROMES ..... 17
1.8 BREAST CANCER PROGNOSIS ..... 18
1.9 CLINICAL ASPECTS ..... 19
2 AIMS ..... 21
3 MATERIAL AND METHODS ..... 23
3.1 MATERIAL ..... 23
3.1.1 Patients from Cancer Counselling Clinics, Karolinska University Hospital (paper I-III) ..... 23
3.1.2 Patients from other cohorts ..... 25
3.1.3 Paper I ..... 26
3.1.4 Paper II ..... 27
3.1.5 Paper III ..... 27
3.2 METHODS ..... 28
3.2.1 Paper I ..... 28
3.2.2 Paper II ..... 29
3.2.3 Paper III ..... 29
4 RESULTS AND DISCUSSION ..... 33
4.1 PAPER I ..... 33
4.2 PAPER II ..... 35
4.3 PAPER III ..... 36
5 CONCLUSIONS AND CLINICAL IMPLICATIONS ..... 38
6 ACKNOWLEDGEMENTS ..... 39
7 References ..... 41

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ATM | Ataxia telangiectasia mutated |
| BCAC | Breast Cancer Association Consortium |
| BIC | Breast Cancer Information Core |
| BRCA | Breast cancer gene |
| CDH1 | Cadherin 1 |
| CHEK2 | Chek point kinase 2 |
| CI | Confidence Interval |
| CNV | Copy number variant |
| DHPLC | Denaturing high performance liquid chromatography |
| DNA | Deoxyribonucleic acid |
| DSB | Double-strand breaks |
| GWAS | Genome-wide Association Study |
| ER | Oestrogen receptor |
| FA | Fanconi Anemia |
| HER2 | Human Epidermal Growth Factor Receptor 2 |
| HDGC | Hereditary diffuse gastric cancer |
| HNPCC | Hereditary non-polyposis colorectal cancer |
| HRT | Hormonal replacement therapy |
| LD | Linkage disequilibrium |
| LOD | Logarithm of odds |
| MLPA | Multiple ligation-dependent probe amplification |
| NBN | Nibrin |
| NF1 | Neurofibromatosis 1 |
| NGS | Next Generation Sequencing |
| PALB2 | Partner and localizer of BRCA2 |
| PgR | Progesterone receptor |
| PCR | Polymerase chain reaction |
| PHTS | PTEN Hamartoma Tumour Syndrome |
| PRS | Polygenic risk score |
| PTEN | Phosphatase and Tensin homolog |
| RNA | Ribonucleic acid |
| RR | Relative risk |
| RECQL | RecQ like helicase |
| SNP | Single nucleotide polymorphism |
| STK11 | Serine/Threonine kinase 11 |

| | |
|---|---|
| TP53 | Tumor protein P53 |
| VUS | Variant of unknown significance |

# 1 INTRODUCTION

Breast cancer is the most common cancer type affecting women. In Sweden the life time risk is of being affected is estimated to 11%. Many risk factors for breast cancer have been identified including hormonal and lifestyle factors. Furthermore, the individual risk increases proportionally with the number of close relatives and early age at onset. Hereditable factors are estimated to contribute to 27% of the total risk of breast cancer (Table 1) [1, 2]. Clustering of breast cancer in families has been investigated and supported by epidemiological studies since the late 1940s [3]. The search and identification of risk genes for hereditary breast cancer progressed in the early nineties when linkage analyses in affected families lead to the discovery of *BRCA1*, *BRCA2*. A candidate gene approach using affected family members in cancer families led to the discovery of TP53 as causal factor behind Li-Fraumeni syndrome. These high penetrance traits and identification of the breast cancer predisposing genes *STK11, CDH1* and *PTEN*, that followed, have in common the occurrence of breast cancer as a part of a cancer syndrome, with several types of cancer involved. Further linkage and candidate gene studies have been unsuccessful in finding more high-penetrance susceptibility genes despite the fact that the already identified ones only represent a small part of the estimated heritable fraction. It has become clear that the remaining fraction of hereditability is not explained by variants in high-risk breast cancer genes. Evidence has accumulated that breast cancer is inherited as a polygenic disease where pathogenic mutations with differential impact together confers an increased risk. Focus has shifted towards research for genetic risk factors that confers a low to moderate risk of breast cancer. The first identified moderate risk gene, cell cycle checkpoint kinase 2 (*CHEK2*) identified in 2002, was found to cause a twofold risk of breast cancer and an up to fourfold risk in familial breast cancer cases [4, 5]. Advances in genomic technologies have led to large scale sequencing in genome-wide associations studies that has identified a large amount of common low penetrance loci. Great progress has been made, but a substantial portion of the genetic background in familial breast cancer is yet unexplained. In addition, the functional and clinical implications of the risk-genes already identified are limited since their polygenic context is unclear.

The aim of this thesis has been to study and better understand how breast cancer is inherited and to identify underlying genetic factors that contribute to the risk in familial breast cancer. We decided to focus on families without pathogenic mutations in *BRCA1* or *BRCA2* (non-

*BRCA* families). In these families the susceptibility to inherited breast cancer varies from moderate, with few cases of breast cancer in close relatives, to high with many cases in every generation and often with low age at onset. A better understanding of inheritance patterns is warranted in search of new genetic candidate variants. Identifying new susceptibility variants will improve prevention programs and individual counselling in affected families.

| | **Factor** | **Risk group** | **Effect on risk (RR)** | **References** |
|---|---|---|---|---|
| Demographic factors | Geographic area | US and Northern Europe vs Africa, Asia | 3-4 | Globocan 2012 |
| Family history | On close relative with breast cancer | Yes vs No | 1.1-2 | Pharoah [2], Claus [6] |
| | Two first degree relatives | Yes vs No | 3,6 | Pharoah [2] |
| Reproductive factors | Age at menarche | Early vs Late | 1.1-2.0 | Hulka [7] |
| | Age at menopause | Late vs Early | 1.1-2.0 | Hulka [7] |
| | Age at first birth | >30 vs < 20 years | 1.1-2.0 | Hulka [7] |
| | Parity | Nulliparous vs Multiparous | 1.1-2.0 | Hulka [7] |
| | Breastfeeding | None vs >1 year | 1.1-2.0 | Lancet [8], Hulka [7] |
| Other Hormonal Factors | BMI (postmenopausal) | High vs Low | 1.1-2.0 | van den Brandt[9] |
| | BMI (premenopausal) | Low vs High | 1.1-2.0 | van den Brandt[9] |
| | Oral contraceptives | Ever vs Never use | 1.07 | Lancet [10] |
| | Oral contraceptives | Current use vs None | 1.24 | Lancet [10] |
| | HRT | Current use vs None | 1.1-2.0 | Lancet [11] |
| | Serum estradiol | High vs Low | 1.1-2.0 | Thomas [12], Hankinson [13] |
| Breast Properties | Mammographic density | Dense vs Not | 4-5 | McCormack [14] |
| | Atypical Ductal Hyperplasia | Yes vs No | 1.1-2.0 | Byrne [15] Page [16] |
| Other | Ionizing radiation to the chest at young age | Yes vs No | 1.1-2-0 | Hancock [17] |
| | Alcohol | Daily intake vs None | 1.2-2.0 | Schatzkin [18], Smith-Warner [19] |

Table 1. Risk factors in breast breast cancer (Adapted from Hulka et al Lancet [7]). RR = relative risk.

## 1.1 EPIDEMIOLOGICAL STUDIES

A systematic review of 74 studies, from 1935-1995, stated a pooled estimate of relative risk of 1.9 for having any relative affected by breast cancer, 2.1 for having one first degree relative and 3.6 when both mother and a sister were affected. [2]. One of the studies in the review was a Swedish population based study of 2660 women that found a relative risk of 1.7 with having one or more affected first degree relatives [20]. Besides the number of affected relatives, it was shown that presence of bilateral breast cancer and/or low age at diagnosis in a relative also increases the individual risk [21, 22]. Through pedigree studies it became clear that inheritance patterns of breast cancer showed a spectrum of risk levels. Some families had a picture of a strongly inherited breast cancer with many cases across the generations. In addition, these families had higher incidence of bilateral disease and women were younger at diagnosis compared to sporadic breast cancer [23]. Several studies with segregation analysis followed that suggested an autosomal dominant disease and also an association with ovarian cancer in some affected families [24-27].) In addition, twin studies indicated a higher risk of breast cancer in monozygotic twins compared to dizygotic twins, supporting a strong inherited component behind familial clustering of breast cancer [1, 28].

## 1.2 IDENTIFYING BREAST CANCER SUSCEPTIBILITY GENES

Various approaches have led to identification of breast cancer susceptibility genes that in general are divided into three groups depending on their allele frequency and their conferred risk. The risk alleles mostly show an inversely proportional impact, from very rare mutations with high penetrance to the common low risk polymorphisms with allele frequency of up to 50%:

*High-risk variants:* Very rare in the population with a minor allele frequency of <0.005. The conferred risk of pathogenic mutations can be ten-fold.

*Moderate-risk variants:* Minor allele frequency in population of 0.005-0.01%. Pathogenic variants confer a two- to fourfold increased risk of breast cancer.

*Low-risk variants:* Minor allele frequency > 0.05 and conferred risk of breast cancer of less than 1.5-fold [29].

### 1.2.1 High-risk variants

The genes underlying high-risk syndromes like *BRCA1* and *BRCA2* have mainly been identified through linkage analysis and positional cloning. In 1990, after 17 years of research,

Mary-Claire King and her team had identified a high-risk gene assigned to chromosome 17q21 by linkage analysis in families with early breast cancer [30]. The locus on chromosome 17 was supported again by Steven Narod and his team in 1991 and, in addition, an association with ovarian cancer was established [31]. Subsequently, in 1994, the *BRCA1* gene was identified through positional cloning and truncating mutations in the protein coding sequence were found to segregate with disease [32]. In 1994, an international team lead by Mike Stratton analysed 15 early-onset breast cancer families that were not related to *BRCA1* to search for a second high-risk gene that was localised to chromosome 13q12-13 [33]. Shortly thereafter, BRCA2 was cloned and, when tumour samples were examined for loss of heterozygosity, association with defect BRCA2 was confirmed [34, 35]. Li-Fraumeni syndrome was first described in 1969 as a rare clinical syndrome with a high penetrance of premenopausal breast cancer and sarcomas, brain tumours and adrenal cortical carcinomas at very young age [36]. The rarity and high mortality of the syndrome made linkage analysis difficult. Instead a candidate gene approach was used to identify germline mutations of the tumour suppressor gene *TP53,* located on chromosome 17p13.1, as the causative factor. *TP53* was selected as a plausible gene because it was identified as the most common inactivated gene in sporadic forms of cancers associated with Li-Fraumeni [37, 38].

### 1.2.2 Moderate-risk variants

Also the first identified moderate risk gene, *CHEK2* (1100delC), was localized through linkage-analysis of a large non-*BRCA* family [4]. To assess the conferred risk, linkage analysis was combined with an association study since moderate risk genes display incomplete co-segregation due to their intermediate penetrance. For that reason, candidate gene approach is more applicable when identifying moderate penetrance genes. This search has focused on candidate genes, which like *BRCA1* and *BRCA2,* are involved in DNA damage response or other pathways important for breast cancer biology. By resequencing of genes in cases from highly penetrant families moderate susceptibility genes have been identified and thereafter evaluated in case-control association studies. Breast cancer genes identified with this approach includes pathogenic variants in moderate risk genes *ATM* [39-41] and *PALB2* [42, 43]. Since the high-throughput technique Next Generation Sequencing (NGS) now is more affordable whole-exome sequencing has become a common approach to identify rare variants. When searching for rare risk alleles reduced statistical power demands sequencing of enormous numbers of unselected cases and controls. One strategy to reduce sample size is by performing a staged study starting with sequencing in a small cohort of

cases with familial aggregation of breast cancer, assuming that genetic contribution is higher. The cases can also be chosen due to certain criteria such as being negative for *BRCA1/2* mutations (as in paper I), tumour phenotype or sharing a previously identified susceptibility gene (as in paper III). Identified risk variants are thereafter genotyped in larger cohorts. Pathogenic variants in a new breast cancer susceptibility gene, *RECQL*, were identified through whole-exome sequencing in two cohorts consisting of 144 Polish and 51 French-Canadian women with familial breast cancer that all were negative for *BRCA1/2* founder mutations. In the discovery phase, candidate genes for further study were selected by focusing on highly deleterious mutations in genes with a function previously associated to cancer. Several truncating mutations in the gene *RECQL*, a DNA-helicase gene previously related to cancer were identified. Pathogenic *RECQL* variants were then validated in larger cohorts containing patients with sporadic and familial breast cancer as well as a control group [44].

### 1.2.3 Common low-risk variants

The results of search for the third category of susceptibility genes, common low risk alleles, have mainly been achieved through genome-wide association studies (GWAS). This approach scans most of the genome for causal genetic variants without any assumption of biological function or location. The genome of two individuals is 99.55% identical, but the variations can greatly affect the risk of various complex disease. More than 10 million of the single nucleotide polymorphisms (SNPs), with a minor allele frequency of 1% or more, are collected in a database, dbSNP [45]. Sets of SNPs in a genomic region are in strong linkage disequilibrium (LD) inherited in blocks. These blocks contain a large number of SNPs with specific patterns, or haplotypes, that can be identified with a reduced number of tag-SNPs. All tag-SNPs have been categorized through the HapMap project. Because of the high correlation between the variants (high LD), a few hundred thousand of SNPs can be used as markers over the whole genome. In GWAS, genotyping platforms that allow evaluation of hundreds to thousands of SNPs simultaneously, are utilised. In general, in a first phase of a GWAS, tag-SNPs are genotyped in smaller number of cases and controls. In the second and third phase variants with the strongest evidence are narrowed down to a small number and replicated through testing for significance in large association studies. The identified low-risk variants are not always the causal ones; any SNPs within the haplotype block could play the role of casual SNP. Fine-mapping, that contains genotyping of all common variants of risk loci, can bring more clarity to the issue. Furthermore, most of the identified variants are

located in non-coding regions, intergenic- and intronic variants are common. They may have an effect on regulation of gene expression or function. So far more than 180 low risk loci associated with breast cancer risk have been identified [46-70]. Since it is more affordable and less complicated than complete resequencing of whole genomes, candidate gene-association studies have also been used to search for common low-penetrance alleles, but with limited progress. One of few convincing examples is a variant in the CASP8 gene which has been associated with a protective effect. *CASP8* is a tumour suppressor involved in regulation of apoptosis [71].

## 1.3 EVIDENCE OF ASSOCIATION FOR SPECIFIC GENES AND VARIANTS

The following section is a summary for genes where some evidence of an association with breast cancer has been established. Most allelic variants associated with a moderate to high risk of breast cancer are protein-truncating. These are nonsense mutations, frame-shift insertions and deletions and splice site affecting variants that in general causes a premature stop codon and as a result, a non-functioning protein product. Some of the alterations associated with breast cancer risk are missense variants which refer to single base pair substitutions causing a change of one amino acid. Compared to truncating variants, the effect of a missense variant is more difficult to assess, the amino-acid change may not have a dramatic effect on the resulting protein. Risk assessment of the pathogenicity of missense variants can be time-consuming and requires analyses of whether the particular variant affects an evolutionary conserved region, the probable impact on protein function as well as genetic analyses in affected families and tumour subtypes. Variants of uncertain pathogenicity are referred to as variants of unknown significance, VUS. Furthermore, pathogenic missense variants are in general associated with a lower risk than truncating variants. For instance, the *CHEK2* truncating variant *c.1100delC* is associated with a higher risk of breast cancer compared to *CHEK2* missense variant *I157T* (c.470A > G) [72]. There are exceptions though, for instance, the *ATM* missense variant p.Val2424Gly (c.7271T > G) has been associated with a higher risk compared to truncating variants [73]. The Breast Cancer Information Core (BIC), is an open access on-line mutation database that provides all available research and technical support of all mutations and polymorphisms in breast cancer susceptibility genes.

### 1.3.1 High-risk genes

#### 1.3.1.1 *BRCA1* and *BRCA2*

The *BRCA1* gene encodes a multi-domain protein that functions in a number of cellular pathways to maintain genomic stability, including cell cycle checkpoint activation as well as transcriptional regulation and apoptosis [74]. *BRCA1* and *BRCA2* are also important for DNA repair, specifically in homologous recombination of double-strand DNA breaks (DSB) [75]. *BRCA1* and *BRCA2* mutations confer a very high life-time risk for breast cancer in the range of 55-85% for *BRCA1* and 35-60% for *BRCA2* whereas the risk of ovarian cancer is up to 40% for *BRCA1* but slightly lower for *BRCA2* carriers with 10-25% [76-78]. In comparison the estimated risk of breast cancer in Swedish women by the age of 75 is 11% and the risk of ovarian cancer is 1.1% [79]. *BRCA1* and *BRCA2* pathogenic mutations are found throughout the gene´s whole coding sequence and are usually truncating. Pathogenic missense mutations have also been described in functional domains. In addition, different pathogenic mutations show a various spectrum of allelic effects, such as the missense variant, R1699Q (c.5096G>A p.Arg1699Gln) in *BRCA1* and one truncating variant, p.Lys3326 (rs11571833) in *BRCA2* that both are associated with a significant increased risk of breast cancer, but much lower than the average risk for carriers in general [46, 80].

Studies of other tumour types associated with *BRCA1* have been inconclusive [81, 82]. Besides breast- and ovarian cancer in women, tumours linked with *BRCA*2 mutations include pancreatic- and early onset prostate cancer. An association with melanoma has also been reported [82-84].

#### 1.3.1.2 TP53

One of the most intensively studied tumour suppressors is tumour protein 53 (TP53) since loss of wild-type *TP53* activity frequently detected in many different tumour types. The protein regulates the cell cycle, interact in DNA repair, apoptosis, cellular senescence, and metabolism. Inherited *TP53* mutations are associated with the rare autosomal dominant disorder, the Li-Fraumeni syndrome (LFS). Female mutation carriers have a nearly 100% lifetime risk of cancer compared to 73% for males, the difference caused by breast cancer [85]. The high cumulative cancer risk was also seen in two later studies, but these may have been affected by selection bias [86]. Breast cancer is the most common tumour with a 49% risk of being affected before 60 years, but most women are diagnosed before age 40 [87]. Unlike other high-risk genes that mostly display a risk associated with truncating mutations, genotype-phenotype analysis in LFS families has revealed that carriers of germline missense

mutations are more frequent and associated with earlier onset compared to other types of mutations [88]. In addition, there is a more than 20% frequency of de novo mutations [85]. Other than breast cancer in women, *TP53* mutation carriers are at increased risk of early-onset and multiple primary cancers including sarcomas, brain- and adrenocortical tumours [89]. Lymphoma, melanoma, lung, pancreas, prostate and ovarian cancers also seem to be increased [89-91]. Childhood-onset tumours exists, most common are brain-tumours, followed by sarcomas.

#### 1.3.1.3 PTEN

Cowdens syndrome is rare condition caused by germline mutations in tumour suppressor gene *PTEN* [92]. The syndrome confers an estimated 25-50% lifetime risk of breast cancer, with low age of onset. Carriers are also at an increased risk of several other malignancies, especially thyroid- and endometrial cancer with a life time risk reported to be 3-10% and 5-10% respectively [93]. The syndrome is otherwise characterized by multiple hamartomas of the gastrointestinal tract and benign tumours. Cowdens syndrome was at first described as a clinical diagnosis in the 1960´s. In 1996 *PTEN* was identified after years of linkages studies and subsequent sequencing in Cowdens syndrome families. Germ-line mutations in *PTEN* have also been associated with other clinical distinct syndromes, in patients with disparate disorders, referred to as PTEN Hamartoma Tumour Syndrome (PHTS). [94]. Studies that have examined cancer risk in PTHS patients have identified greatly increased lifetime risk of breast- (67-85%), endometrial- (21-28%) thyroid and renal cancer as well as a slightly increased risk for colorectal cancer and melanoma [95, 96].

#### 1.3.1.4 STK11

The tumour suppressor gene *STK11* is another gene with a gene product important for cell cycle regulation and mediation of apoptosis. Deleterious mutations causes Peutz-Jeghers syndrome characterized by intestinal hamartomous polyps and mucocutaneous pigmentation [97]. In addition, the lifetime risk of breast cancer by 60 years has been estimated to 32-54%. Other associated tumours with markedly elevated risk are cancers of gastrointestinal origin, pancreatic- and ovarian cancer [98]. Carriers of *STK11* mutations have a cumulative lifetime risk of any cancer of up to 85% [99].

#### 1.3.1.5 CDH1

The *CDH1* gene encodes a protein responsible for cell-to-cell adhesion and functions as a cell invasion suppressor [100]. Impaired E-cadherin activity leads to increased motility and metastatic potential of tumour cells [101]. E-cadherin germline mutations are responsible of hereditary diffuse gastric cancer (HDGC). Carriers are at a high risk of diffuse gastric carcinoma at young age (cumulative risk 83%). There is also an association with lobular breast cancer, which is the second most frequent tumour type, and colorectal cancer [102, 103]. Estimated lifetime risk of developing lobular breast cancer is 40-52% [104, 105]. Recent studies have provided evidence of lobular breast cancer as the first manifestation of HDGC. Deleterious *CDH1* mutations have been identified in women with bilateral lobular breast cancer without a family history of diffuse gastric cancer. Also novel deleterious *CDH1* alterations have been identified raising the question whether lobular breast cancer can be inherited as an independent E-cadherin syndrome [106, 107]. In contrast to other cancer predisposition syndromes, pathogenic splice-site and missense mutations are common; suggesting that also reduced E-cadherin expression can be deleterious [108].

### 1.3.2 Moderate-risk genes

#### 1.3.2.1 PALB2

*PALB2* (the partner and localizer of *BRCA2* gene) was first associated with breast cancer as a risk gene with moderate penetrance in 2007. Loss-of-function *PALB2* mutations yielded risk estimates of two to four-fold times higher in familial breast cancer compared to non-carriers [42, 43]. In a study by Antoniou et al in 2014 the estimated conferred risk by deleterious *PALB2* mutations was overlapping *BRCA2* when breast cancer risk was analysed in 362 patients in 154 families. The risk of breast cancer in carriers by 70 years was estimated to 35%. Breast cancer risk was depending on family history and ranged from 33-58% with risk in the lower range in women without a family history [109]. As a result, *PALB2* is now entering clinical testing, in women at increased risk of breast cancer. Carriers should be offered measures according to guidelines for hereditary breast cancer with surveillance and prophylactic surgery. *PALB2* was originally identified as *BRCA2* interacting protein that enables homologous recombination and double strand break repair and check point functions [110]. Thereafter, it has also been shown that *PALB2* interacts with *BRCA1* as well [111]. A predisposition for ovarian- and pancreatic cancer conferred by *PALB2* mutations has been indicated, though larger studies are needed to address this question [43, 112, 113].

### 1.3.2.2 CHEK2

Checkpoint kinase 2 (*CHEK2*) is in response to DSB or replicative stress activated by *ATM* [114]. Activated *CHEK2* plays an important role in phosphorylating downstream substrates including tumour suppressors *TP53* and *BRCA1* [115]. Originally *CHEK2* was found mutated in Li-Fraumeni patients and one of the truncating mutations c.1100delC, has been reproducibly associated with breast cancer risk in different populations. It seems to be more frequent in affected carriers in parts of Northern and Eastern Europe (1.2-3.5%) compared to Southern Europe and North Amerika [116-118]. In a Swedish population the *CHEK2*1100delC* mutation was found to be most frequent in younger women with familial breast cancer (<45 years) with 5% compared to 2 % of all affected women with a family history and less than 1% in sporadic breast cancer [119]. Recently, a large study reported a cumulative life-time risk of breast cancer in *CHEK2*1100delC* carriers of about 22% [120].

Other than the *CHEK2*1100delC* variant, two more truncating variants have been associated with similar breast cancer risk (del5395 and IVS2 + 1G > A) whereas one missense mutation, I157T (c.470A > G) confers only slightly increased risk [72, 121, 122]. However, the few studies that have analysed germline mutations of the entire coding sequence of *CHEK2* have identified more potentially deleterious variants. Since the allele frequency of the most common pathogenic *CHEK2* variants are highly population-specific, full-gene sequencing has been suggested in regions where these founder mutations are rare [123]. A 3.5 fold excess risk for *CHEK2* 1100delC* carriers to develop contralateral breast cancer has been found as well as a worse prognosis compared to non-carriers, especially for oestrogen receptor-positive breast cancer [124, 125].

The individual risk of breast cancer for carriers of the most evaluated *CHEK2* mutation, c.1100delC, is difficult to assess due to the moderate penetrance. As a consequence, mutations do not always segregate with breast cancer in the affected families. Furthermore, the conferred risk is affected by family history. In a study of *CHEK2*1100delC* carriers, an odds ratio of 2.7 were estimated in unselected breast cancer, 2.6 for early-onset breast cancer and 4.8 for familial breast cancer. The cumulative risk by 70 years was estimated to 37% in familial breast cancer cases [126]. There has been reports of an association between *CHEK2* pathogenic variants and ovarian cancer and several other additional malignancies, but the results have been inconclusive [127, 128]. A meta-analysis provided evidence that *CHEK2*1100delC* carriers may be at an increased risk of colorectal cancer [129]. In another

large GWAS, the result suggested an association with increased risk of lung cancer even though the I157T variant was associated with reduced risk [130].

#### 1.3.2.3 ATM

Homozygous or compound mutations in the ataxia-telangiectasia mutated gene (*ATM*) is the cause of a neurodegenerative disorder, named after the gene, characterized by cerebellar ataxia, immunological deficiency, hypersensitivity to ionizing radiation and increased risk of cancer. The gene encodes a protein kinase involved in DNA repair of double strand breaks through downstream interaction with *TP53*, *BRCA1* and *CHEK2* proteins in the damage response pathway [131]. A meta-analysis found heterozygous carriers of *ATM* mutations at an estimated relative risk of 2.8 [132].

#### 1.3.2.4 NF1

Neurofibromatosis 1 (*NF1*) is an autosomal dominant cancer syndrome, with a very variable phenotype, characterized by skin pigmentation and tumours of the nervous system.
A Finnish study with a population-based cohort found a 59.6% life-time risk of cancer. Excess risk of breast cancer was moderate, but highest in women younger than 40 years [133]. These results were supported in another study of *NF1* associated breast cancer risk that found an increased risk in women less than 50 years, but thereafter incidence patterns similar to the general population [134, 135]. An estimated increased relative risk of 2.6 has been calculated based on data from two cohort studies [132].

#### 1.3.2.5 NBN

The *MRE11A-RAD50-NBN* (*NBS1*) is an evolutionary conserved protein complex important for detection and early processing of double strand breaks. Carriers of heterozygous *NBN* mutations have a significant increased risk of breast cancer with an estimated odds ratio of 3.1 [136]. In a meta-analysis of seven studies, strong evidence was presented for a moderate increased breast cancer-risk associated with a truncating variant, c.657del5 [137]. Another study provided evidence for an association between the variant and early onset breast cancer [138].

### 1.3.3 Other breast cancer gene susceptibility candidates

The fact that *BRCA1*, *BRCA2* and *PALB2* function in the repair of DNA double-strand breaks has suggested further studies on candidates in pathways controlling DNA integrity. Several other susceptibility genes with protein products involved in DNA damage signalling and repair have in different populations been associated with breast cancer risk in single studies, but not confirmed in large systematic associations studies. Most of the identified pathogenic variants are very rare, which makes a putative association with breast cancer and a robust estimate of penetrance difficult to assess. *RAD51,* a key protein in mediating homologous recombination, forms a complex with a family of accessory proteins known as *RAD51* paralogs. Deleterious mutations in *RAD51* paralogs *RAD51C, RAD51D,* have shown clear evidence of an association with ovarian cancer, but there is insufficient evidence for an association with increased risk of breast cancer [139, 140]. Exome-sequencing and a subsequent association study identified *RAD51* paralog, *XXR2C*, as a putative breast cancer gene [141]. *BRIP1*, that encodes a DNA helicase, has also shown clear evidence of predisposing for ovarian cancer but a role in breast cancer risk has not found support [142]. A recent large-scale association study could not provide evidence for an association between any truncating variant in the *BRIP1* gene and breast cancer [143, 144]. The statistical analysis allowed exclusion of a twofold risk of breast cancer, which in general is considered the lower threshold for a moderate susceptibility gene. Recently, rare mutations in two other helicase genes, *RECQL* and *BLM* have been associated with breast cancer [44, 145]. Biallellic mutations in *BRIP1* as well as *PALB2*, *RAD51C* and *BRCA2* (also known as the *FANCJ*-, *FANCD1*, *FANCN* and *FANCO* genes) cause Fanconi Anemia (FA), which has prompted a search of other FA genes predisposing for breast cancer. Exome sequencing has identified heterozygous truncating mutations in FA genes *FANCC* and also in *FANCM* with a particularly strong predisposition for triple negative breast cancer [145-147]. A limited number of pathogenic germline mutations in the *MRE11* and *RAD50* have also been suggested as breast cancer predisposing variants, but their role is yet to be determined [148].

### 1.3.4 Low penetrance variants

During the last decade genome-wide association studies (GWAS) have been successful in identifying a new category of susceptibility variants common single nucleotide polymorphisms (SNPs) with an allele frequency of up to 50%, in the population, that confer a relative risk of breast cancer of 1.5 or less. Collaborations that can achieve large datasets have proven very successful and is necessary when searching for low penetrance variants. The

Breast Cancer Association Consortium (BCAC), a collaboration with more than 50 participating breast cancer case-controls studies, has conducted several combined large-scale genotyping studies. In a BCAC study of more than 120.000 cases and controls, in 2015, 15 more low risk loci were added to the previously more than 90 identified [46-69]. Provided that the identified low risk loci combine multiplicatively, the authors concluded that risk profiling could identify 5% of women at a 2-fold higher risk and 0.7% of women with a >3-fold higher risk of breast cancer than the population average [69]. The most recent genome wide association study, performed by BCAC and the OncoArray Consortium, added 65 more loci associated with breast cancer risk, 19 of them more strongly associated with ER-positive disease [70]. Most of the identified variants are intergenic or intronic with uncertain function, but they may be important for gene expression and function. Only a few casual protein coding variants have been suggested, one being the SNP rs11552449 which is a missense substitution in the *DCLRE1B* gene (*SNM1B*) involved in DNA cross-link repair [58] Further investigation of a tag SNP in an intron of the *EXO1* gene, with a gene-product involved in mismatch repair, found a strong association with a putative causal variant in the coding region of the gene. The associated missense variant, rs4149909, encodes an amino acid substitution, with potential pathogenic effect related to breast cancer [69] Several GWAS have investigated specific effects by breast cancer subtypes, mainly by oestrogen receptor (ER) status. Studies in ER-negative breast cancer have revealed seven ER-negative specific risk loci [59, 149, 150]. Another study analysed established breast cancer loci in triple negative breast cancer and found that variants from 25 of 74 loci also were associated with increased risk of triple negative breast cancer [151]. BCAC and the OncoArray Consortium recently published results from a GWAS in about 40.000 cases of ER-negative breast cancer and *BRCA*1 mutations carriers. Ten new variants associated with ER-negative breast cancer were identified. Furthermore, previously identified variants were replicated and altogether the 125 associated variants were estimated to explain 16% of the familial risk of ER-negative breast cancer [152].

## 1.4 GENETIC CONTRIBUTION IN BREAST CANCER

Overall, BRCA1 and BRCA2 mutations are estimated to be responsible for about 3% and the other less common high penetrance genes account for less than 1 % of all breast cancer [29, 153]. A recent study though suggests that BRCA1 and BRCA2 accounts for only 1,4 % of all breast cancer and that other results with higher proportions could be due to founder populations or to patient selection [154]. In Sweden, carrier frequency of *BRCA1* was less

than 1 % in an unselected breast cancer cohort in Stockholm County [155]. A population based cohort in Southern Sweden found *BRCA1* or *BRCA2* pathogenic mutations in 9.0% of early onset breast cancer [156]. A Nordic study of mono- and dizygotic twins estimated that hereditary factors contribute to 27% of the total risk of breast cancer [1].

Pathogenic mutations in *BRCA1* and *BRCA2* account for 16% of the genetic contribution in familial breast cancer risk [157]. In total, the high risk genes, *BRCA1/2*, *p53*, *STK11* and *PTEN* account for approximately 20% of the familial risk [158]. Moderate penetrance variants account for up to 5% of the inherited familial risk [29, 159]. The more than 180 identified low-risk loci explain 18% of the familial risk [70]. Altogether more than half of the genetic background in familial breast cancer still is unclear (Figure 1).

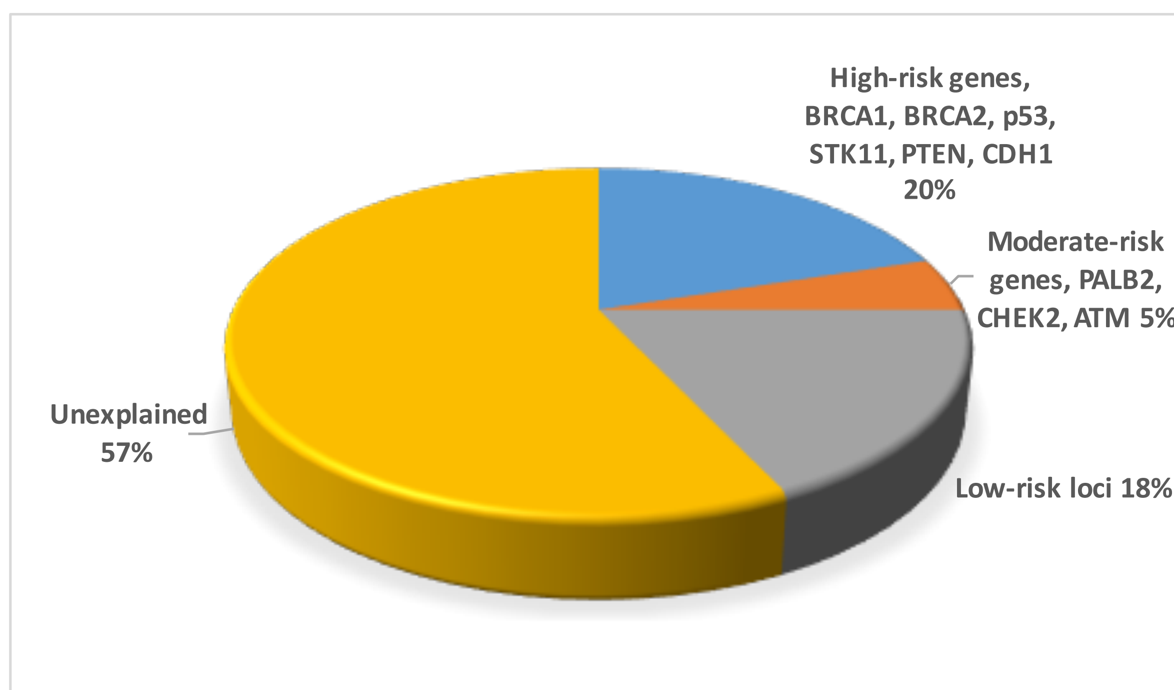


Figure 1. Fraction of the excess familial risk of breast cancer explained by the currently known susceptibility genes/loci.

## 1.5 GENETIC MODIFIERS

Evidence has accumulated that familial breast cancer is a polygenic trait and that the genetic susceptibility is driven by variants at many loci, each conferring a low to moderate risk of the disease [160, 161]. This multiplicative model could explain a part of the residual component also in carriers of high-risk genes such as *BRCA1* and *BRCA2*. Several studies have found support for low-risk variants that modify risk in *BRCA1* and *BRCA2* mutations carriers [162-165]. The markedly higher lifetime risk, in *CHEK2**1100delC* carriers, with a family history

of breast cancer compared to those without, suggests that *CHEK2*1100delC* combines with other genetic variants that together modifies the risk of breast cancer. A recent study investigated whether the multiplicative model was applicable on *CHEK2*1100delC* and common low penetrance susceptibility variants previously associated with breast cancer risk. A polygenic risk score (PRS) was calculated based on the effect of 77 low penetrance variants. The synergistic effect of *CHEK2*1100delC* and the PRS on breast cancer risk was estimated and the result gave strong support for the multiplicative polygenic model. The effect of the PRS was slightly attenuated when adjusting for family history, but remained an independent risk factor. The polygenic risk score was suggested as a tool for stratifying risk in *CHEK2*1100delC* carriers [166]. It is not clear whether the model can be applicable on combinations of moderate to high-penetrance genes. A study of a combined effects of *BRCA*, *ATM* and *CHEK2* mutations, found no support for a multiplicative effect on breast cancer risk. This may reflect their biological interaction in DNA-repair. Impairment of functions up-streams of *BRCA*, such as *ATM* and *CHEK2*, may confer little or no additional risk of breast cancer [167]. Investigating this issue is difficult since carriers of any combinations of high- and moderate penetrance genes are very rare. In our study of Swedish *CHEK2*1100delC* carriers, exome-sequencing was performed in search of variants of moderate to low impact that modify breast cancer risk, but no *CHEK2* specific variants could be identified (paper III).

## 1.6 OTHER GENETIC RISK FACTORS

### 1.6.1 Epigenetic regulation

Epigenetics refers to heritable changes in gene expression that occur without alteration in DNA sequence. In fact, epigenetic modifications regulate patterns of gene expression that are crucial for normal development and differentiation of cell-lines. It is now well established that the vast majority of human cancers harbours both genetic mutations and epigenetic alterations that interact in a complex network. Silencing of tumour suppressor genes and activation of oncogenes are caused by a variety pf epigenetic mechanisms such as DNA methylation, alteration in the structure of histone proteins and gene regulation by small noncoding microRNAs [168].

Methylation target is cytosin that pairs with guanosin in CpG dinucleotieds. distributed unevenly throughout the genome. Regions with a high frequency of CpG sites, CpG islands, are common in promoter regions. Approximately 60-70% of human genes have CpG islands in their promotor regions. In normal cells, promoter regions are normally hypomethylated and

tend to be transcriptionally active. However, in tumour cells, CpG islands in promoter regions of tumour suppressor genes are often hypermethylated which causes silencing of those genes [169]. In breast cancer, genes involved in apoptosis (*HOXA5* and *TMS1*), cell-cycle inhibition (*p16* and *RASSF1A*) and DNA repair (*BRCA1*) have been found silenced as a result of hypermethylation [170-172]. When hypermethylated, the *BRCA1* gene has shown the same gene expression pattern as the one in *BRCA1*-mutated breast cancer [173]. Hypermethylation in promoter regions is the most common and the most studied type of epigenetic mechanism. Different types of aberrations can occur in cancer. As an example, hypomethylation in two promoter regions have been associated with HER2 positive breast cancer [174].
Epigenetic changes are not limited to CpG methylation, but also include post-translational histone modification and regulation by micro-RNAs (miRNAs). Modification of histones refers to mechanisms that regulates chromatin structure and thereby influence the level of transcription of a gene. One of these mechanisms, reversible histone acetylation has been targeted in the development of anticancer drugs. Histone acetylase inhitibors (HDACi), have been approved for treatment of cutaneous T-cell lymphomas and are also tested as therapy against breast cancer [175]. MicroRNAs are small RNA molecules that can negatively control their target gene expression posttranscriptionally. Aberrant expression of miRNAs has been linked to breast cancer [176]. Increased expression of miR-21 has been associated with advanced stages of breast cancer and upregulated miR-155 with metastasis and poor prognosis [177, 178].

Genomic imprinting is a form of epigenetic inheritance that causes different gene expression depending on parent of origin. Either the maternal or the paternal allele is epigenetically silenced [179]. The imprints are mainly erased and rewritten during egg and sperm formation, but some imprinted genes bypass reprogramming [180, 181]. Imprinting mechanisms have been identified as methylation of imprinting control regions and histone modification [182]. Parent of origin effects have been studied in complex disease. Seven independent variants in known imprinting regions, one of them associated with breast cancer were examined. Five of them, including the breast cancer associated variant, were associated with parent-of-origin effects [183].

### 1.6.2 Copy Number Variation

Copy number variants (CNVs) are structural rearrangements of the genome of regions larger than 50 base-pairs, that can increase or decrease DNA contents. The variability in DNA sequence contributes to our uniqueness but also influences our susceptibility to disease. As for low risk variants, high-throughput sequencing has made it possible to explore the biological impact of copy number variants. Based on the estimated base-pair coverage, that is estimated to 5-10% of the human genome, CNVs are responsible for a larger part of genomic variability than SNPs. CNVs can encompass whole genes, parts of a gene or regulatory sequences leading to altered gene expression [184]. Genes that are thought to always occur in two copies can be found present in three or more copies or missing. The functional effects of CNVs may play an important role in complex disease including breast cancer [185-187].

## 1.7 BREAST CANCER SYNDROMES

A significant feature for the high penetrance breast cancer genes and some of the moderate risk genes is the conferred risk also for other tumour types in the context of breast cancer syndromes. A putative association with other types of cancer can be difficult to assess especially for rare risk genes with an intermediate penetrance. Regional variation of mutations is another complicating factor. An increased risk of other malignancies within families carrying *CHEK2* mutations includes colon, prostate, kidney, and thyroid cancer, but only the association with colon cancer (*CHEK2*1100delC*) has been supported in a meta-analysis [128, 129]. Female carriers of Lynch syndrome mutations are at increased risk of ovarian- and endometrial cancer, in addition to an increased risk of gastrointestinal and urinary tract tumours [188]. One of the mutations behind the syndrome, *MLH1*, has also been associated with increased risk of breast cancer (18.6% cumulative risk to age 70 years 95% CI, 11.3–25.9), but evidence published so far has not been considered robust enough to make specific recommendations of screening programs for carriers [189].
Other studies have focused on investigating familial breast cancer and any putative associations with other tumour types in families without any previously identified pathogenic mutations. In a study of 803 Swedish families with hereditary breast cancer tumours in the colon, ovary, endometrium, pancreas and liver, as well as leukaemia were overrepresented. A putative breast- and endometrial cancer syndrome was suggested [190]. An association between breast- and endometrial cancer was supported again in our study of tumour spectrum in non-*BRCA* breast cancer (paper I). A recently published study based on data from the Swedish Family-Cancer Database calculated relative risk for other types of tumours in

families with aggregation of breast cancer. The strongest associations with breast cancer were found for prostate- and ovarian cancer ($p$-value $<10^{-11}$). Other associated tumour types were for example stomach and male colorectal cancer ($p < 2.5 \times 10^{-6}$), pancreatic cancer ($p < 5 \times 10^{-4}$), thyroid, endometrial and testicular cancer ($p < 0.0025$) [191].

## 1.8 BREAST CANCER PROGNOSIS

The established prognostic factors of breast cancer are tumour size and histologic grade, lymph node metastasis, distant metastasis, proliferation rate and oestrogen-, progesterone- and HER2 receptor status. Gene expression profiling by microarray is more recent way of prognostic profiling that divides breast cancer tumours into four molecular subtypes, Luminal A (ER positive) and Luminal B (ER positive +/- HER2), HER2-enriched and Basal-like (Triple negative). The prognostic factors are used for risk estimation and as predictors of response to adjuvant breast cancer therapy which also affects outcome [192]. An inherited component with impact on prognosis has been investigated from different perspectives. In a Swedish study breast cancer survival was investigated in first degree relatives. Sisters and daughters of women with poor prognosis had a significantly higher breast cancer mortality compared to the corresponding group with good prognosis [193]. The study was reproduced and the results similar after adjustment for prognostic factors and adjuvant treatment [194]. In another study, of 50 000 breast cancer patients (the Sister study), family history was investigated and the results revealed a predominance of maternally inherited cases. The author suggested a differential risk, with a higher risk of breast cancer with maternal inheritance, due to prenatal maternal effects, maternally inherited mitochondrial variants or an imprinting effect [195]. A parent-of-origin effect was indicated in a retrospective Swedish study. The results showed more than twice as many case of familial breast cancer with maternally inherited breast cancer compared to paternally inherited disease. Furthermore the difference increased over time, suggesting a worse prognosis when breast cancer was inherited paternally [196]. If a parent-of-origin effect on breast cancer prognosis exists, studies of underlying putative imprinted loci are warranted. In our study of parent-of-origin effects on prognosis in a cohort of familial non-*BRCA* breast cancer, no support for a worse prognosis with paternal inheritance was found, even though smaller differences could not be excluded due to sample size (paper II).

## 1.9 CLINICAL ASPECTS

Today, definitive clinical recommendations can be drawn only for carriers of a limited number of high- and moderate penetrance genes. However, the risk associated with many missense variants in the established genes seem lower than for truncating variants whereas the risk from many other missense variants still is unclear. International collaborations in large series of patients from different populations could provide reliable estimates of risk. Furthermore, risk assessment of women at increased risk of breast cancer is in general limited due to insufficient knowledge of how a given variant interacts with other genetic risk factors, life-style factors and family history in terms of an absolute risk. Prospective studies combining genetic risk variants with or without presence of a high- or moderate risk variants could provide evidence for decisive risk estimates.

Previously identified variants account for less than half of the inherited risk behind breast cancer, the rest remains unclear. Linkage analysis, in families at high risk of breast cancer, has not been able to find additional high penetrance genes which suggest that if they exist they probably account for only marginal fraction of the familial risk. Further GWAS in large series of patients from different populations will continue to play an important role, but candidate gene studies is also of great importance, especially since the results only to a small extent seem to overlap [197]. Recent studies show that breast cancer is associated with several other types of tumours suggesting that they also could share genetic susceptibility similarly to the previously identified breast cancer syndromes [191, 198] (paper I). One approach in the genetic search can therefore be to target these families with cases of breast cancer and other associated tumour types since the association implicates that they may share genetic susceptibility (paper I, editorial).

# 2 AIMS

Paper I: To identify new breast cancer syndromes in familial non-*BRCA* breast cancer. Putative new breast cancer phenotypes can be targeted in search of susceptibility variants.

Paper II: To investigate if breast cancer prognosis is influenced by parent-of-origin in non-*BRCA* familial breast cancer. If subgroups with worse prognosis are identified, the genetic background can be explored.

Paper III: To analyse exome-sequencing data in *CHEK2*1100delC* carriers and controls in search of candidate variants that may modify breast cancer risk. Subsequent validation of candidate variants in larger cohorts.

# 3 MATERIAL AND METHODS

## 3.1 MATERIAL

The main part of patients and family data for the studies, except for paper III, were collected from the Department of Clinical genetics and the Cancer Counselling Clinics at Södersjukhuset and Radiumhemmet, Karolinska University Hospital (3.1.1).

### 3.1.1 Patients from Cancer Counselling Clinics, Karolinska University Hospital (paper I-III)

Patients and families with inherited increased risk of cancer are referred to or contact themselves the Cancer Counselling Clinics at the Department of Clinical genetics and Oncology at Södersjukhuset and Radiumhemmet, Karolinska University Hospital for genetic counselling. All families with a history of breast- and/or ovarian cancer who received counselling and subsequently were offered *BRCA1* and *BRCA2* screening that was negative (non-*BRCA* families) between February 2000 and January 2012 were eligible for the studies of this thesis (table 1).

1. Two or more 1st or 2nd degree relatives with a least one case of breast cancer and one case of ovarian cancer or a single individual with both breast and ovarian cancer.
2. Three or more 1st or 2nd degree relatives with breast cancer, at least one with onset before 50 years of age.
3. Two or more 1st or 2nd degree relatives with breast cancer, at least one with onset before 40 years of age.
4. One single indivdual with breast cancer before 35 years of age.
5. Two or more 1st or 2nd degree relatives with ovarian cancer.

Table 1. Swedish criteria for BRCA1/2 screening [198].

All genetic testing was performed at the same laboratory at the Department of Oncology at Lund University hospital. For *BRCA* mutation analysis, denaturing high performance liquid chromatography (DHPLC) was used as screening tool between 2000 and 2005. In addition to DHPLC, from 2006 to 2010, multiple ligation-dependent probe amplification (MLPA) was performed to exclude larger genomic rearrangements. Together the DHPLC and MLPA have

a stated sensitivity of 95%. For cases before 2006 blood samples were reanalysed with MLPA, when the technique was introduced. For samples from the year 2010 and after analysis was performed with Next Generation Sequencing with sensitivity over 95 %. Families in the Stockholm County with index patients that have been screened for any pathogenic mutations have a family number and are registered in the database 4D at the Department of Clinical Genetics.

In the counselling procedure pedigrees were constructed of the different cancer diagnoses in the families, age and age at diagnoses. The counsellor verified diagnoses through medical records, Swedish cancer registry or death certificates, if possible, if it was considered of importance in the individual family. Out of 1364 non-*BRCA* families 1206 pedigrees could be collected (Table 2). In 158 families no pedigree could be found, in a random check, a common reason was that there were few cases of cancer in the family, sometimes only one case of breast- or ovarian cancer at low age.

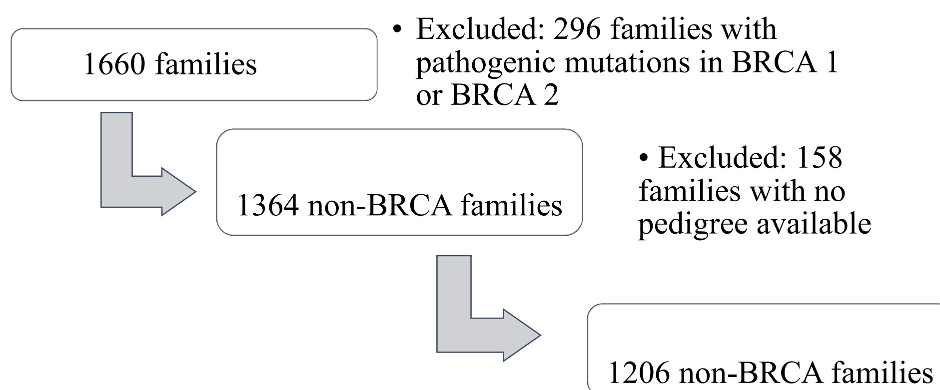


Table 2. Families/pedigrees from Cancer Counselling Clinics, Karolinska University Hospital (paper 1 and II).

All index patients alive in the non-*BRCA* families were asked to participate in a study of genetic risk factors for breast cancer by leaving a blood sample for DNA and by filling in an optional questionnaire regarding hormonal and life style factors (Table 3). Only index patients investigated at the Cancer Counselling Clinics the Department of Clinical Genetics at Karolinska University Hospital and at Södersjukhuset were contacted. The patients were contacted through an invitation letter and were included after a written consent was received. A letter with the questionnaire and sampling tube was then sent by mail. DNA from blood

samples was prepared at the laboratory at the Department of Molecular Medicine and Surgery. The questionnaire consisted of questions regarding hormonal factors (age at onset, age at menarche and menopause, parity, oral contraceptives, hormonal replacement therapy) and life-style questions (alcohol, diet and exercise habits). Three hundred and twelve index patients accepted to participate. Most of the patients were previously affected with breast cancer, but nine of them had been diagnosed with ovarian cancer. Twenty-five of the patients had a history of bilateral breast cancer. The DNA samples were used in paper III, but have also been used in BCAC studies (additional papers). For the BCAC studies, clinical parameters from medical records and questionnaires were collected (hormonal risk factors, prognostic factors, adjuvant treatment, survival, family history).

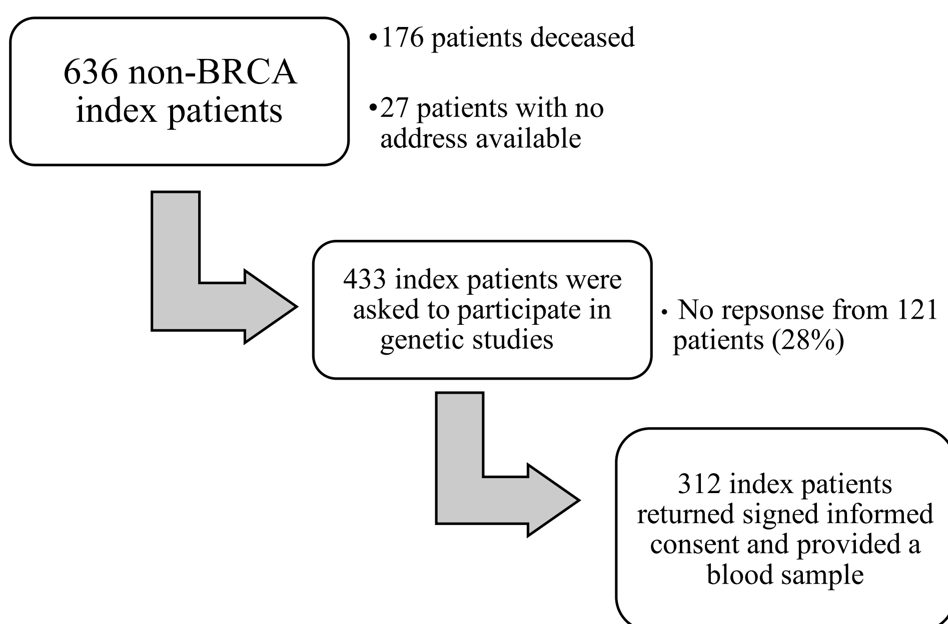


Table 3. Patients from Cancer Counselling Clinics at Department of clinical genetics and Södersjukhuset included in genetic studies (paper III).

### 3.1.2 Patients from other cohorts

#### 3.1.2.1 Patient cohort Södersjukhuset/Huddinge University Hospital (paper II and III)

Patients from a consecutive breast cancer cohort were used in paper II and III. From October 1998 to April 2000 all women with surgically treated primary breast cancer at Sodersjukhuset and Huddinge University Hospital were asked to participate in a study regarding prevalence of breast cancer variants and defining proportions of family history. All participants in the study provided blood samples for DNA. Pathogenic *BRCA*1 mutations were excluded as a part of the study design. Moreover, the women were offered counselling and adequate

additional screening outside the study with the same routines and at same laboratory as for the main study cohort. Patients were classified according to family history of breast cancer. Familial breast cancer patients from the cohort were included in paper II. In paper III, both DNA samples from cases of sporadic breast cancer and familial cancer were used for validation of CHEK2*1100delC variants.

#### 3.1.2.2 SWEA Study (paper III)

In paper III, DNA samples from CHEK2 carriers were collected from the SWEA study. SWEA is a prospective national collaboration, between all Genetic Departments and Cancer Counselling Clinics in Sweden that aim to investigate the genetic background in familial breast cancer. SWEA started in 2012 and includes familial cases, sporadic unselected breast cancer cases and controls. Risk-genes are assessed by complete gene sequencing (*BRCA1* and *BRCA2*, *TP53, PTEN, PALB2, CDH1, ATM, NBN, BARD1, BRIP1, CDKN2A, RAD50, RAD51C/D, MRE11A* and *STK11)* to determine prevalence of established variants and to evaluate candidate variants. A more limited sequencing is also performed in an additional number of breast cancer candidate genes.

#### 3.1.2.3 Swedish Colorectal Cancer Low-Risk Study (paper III)

Samples from healthy spouses of colorectal cancer patients were recruited through the Swedish Colorectal Cancer Low-Risk Study at Karolinska Institutet.

### 3.1.3 Paper I

In paper I all available pedigrees (1206) were used for investigation of tumour spectrum in the non-*BRCA* families from the patient cohort from Karolinska University Hospital (Table 2). Pedigrees from the non-*BRCA* families were examined and included if they contained at least two cases of breast cancer and one cases of other tumour type in 1$^{st}$, 2$^{nd}$ degree relatives or 3$^{rd}$ degree if first cousins on either maternal or paternal branch of the family. Pedigrees from 334 families fulfilled criteria.

### 3.1.4 Paper II

In paper II a parent-of-origin effect on breast cancer prognosis was investigated. The 1206 pedigrees from the mutation negative families were evaluated and included if they contained two cases of breast cancer in 1$^{st}$, 2$^{nd}$ degree relatives or 3$^{rd}$ degree if first cousins (Table 2). In total, index patients from 276 families were included. Families with cases of breast cancer in both family branches were excluded. To increase the study sample, the cohort from Södersjukhuset/Huddinge University Hospital was used (3.1.2.1). Seventy-two of the 487 patients fulfilled inclusion criteria. To avoid overrepresentation of families ascertained on account of the mother being affected, parents of the index case were not counted.

### 3.1.5 Paper III

In paper III we aimed to identify genetic variants that modify breast cancer risk in CHEK2*1100dleC carriers.
Discovery phase: We used 28 samples from *CHEK2*1100delC* carriers from the SWEA study. All *CHEK2*del1100C* carriers were previously affected by breast cancer except for two carriers who had been diagnosed with ovarian cancer. As controls, 28 samples from familial breast cancer cases and 117 familial colorectal cancer cases, recruited through the Cancer Counselling Clinic at the Karolinska University Hospital, Solna, Sweden were used.

Validation phase: We used 72 samples from *CHEK2*1100delC* carriers from the SWEA study. In total 328 cases of sporadic breast cancer, 408 cases of familial breast cancer were and 285 cancer free patients were used in the validation phase. Samples from patients with familial breast cancer were recruited from the patients at the Cancer Counselling Clinics that were included in genetic studies (Table 3) and from the cohort from Södersjukhuset/Huddinge University, which also provided samples from patients with sporadic breast cancer (3.1.2.1). As control cases, 284 healthy spouses of colorectal cancer patients recruited through the Swedish low-risk Colorectal Cancer Study were used (3.1.2.4).

## 3.2 METHODS

### 3.2.1 Paper I

#### 3.2.1.1 Registration of tumour types

All tumour types, other than breast cancer in 1st, 2nd degree relatives or 3rd degree if first cousins (in relation to the index patient), were registered as well as age at onset when data was available. If inclusion criteria were fulfilled in both the maternal and the paternal branches diagnoses from both branches were registered. Each individual cancer diagnosis could only be included once.

#### 3.2.1.2 Reference population

As reference population, the Swedish Cancer Registry was used. To capture changes in incidence two years were chosen, 1970 and 2010 [199, 200].

#### 3.2.1.3 Statistical analysis

Distribution of cancer diagnoses in the data was compared to the distribution of cancer diagnoses in the general Swedish population. Data of cancer diagnoses in the Swedish population was obtained from the National Board of Health and Welfare (Socialstyrelsen). Indirect standardization was here used to adjust the data from the Swedish population to that of the relatives with cancer diagnoses with regards to gender and age. Age was categorized into five-year intervals. For relatives with missing data on gender or age, the method data Missing Completely At Random (Rubin, 1976) was assumed. Cancer cases in the relatives were assumed to be independent of each other. Confidence intervals were calculated for each cancer diagnosis separately, using a binomial distribution. The number of cases was then transformed into proportion of cases, by dividing by the total number of observed cases. Population data were assumed to reflect a true distribution, and were used as reference values. Two reference years were chosen, 1970 and 2010. A cancer diagnosis was regarded as overrepresented in the relatives of the breast cancer patients if the confidence interval was above both population reference values. All statistical analyses were performed in R (R Core Team, 2012). Data entry was performed in EpiData (Lauritsen)

### 3.2.2 Paper II

#### 3.2.2.1 Study arms

Probands previously affected by breast cancer were regarded as the index case, otherwise an index case from the youngest generation in the families was chosen. The index cases were divided into two study arms, paternal or maternal inheritance.

#### 3.2.2.2 Collection of clinical parameters

Tumour characteristics (tumour size, axillary status, ER-status), adjuvant therapy and survival data (recurrence, cause of death) and last date of follow-up were obtained for the index patient from medical records.

#### 3.2.2.3 Statistical analysis

Recurrence-free survival time was calculated from the date of diagnosis to the date of loco-regional recurrence, date of distant recurrence or the date of death, whichever came first. For patients alive and recurrence-free, the time was calculated from the date of diagnosis to the date of last clinical check-up or the date of the most recent contact with health care.
Overall survival was calculated from the date of diagnosis to the date of death or for patients still alive to the date of last clinical check-up or contact with health care. The median follow-up time was estimated using the reversed Kaplan-Meier technique. Survival over time was estimated using the Kaplan-Meier technique with the number of patients still at risk included in the graphs. Proportional hazards regression was used to estimate the univariate and multivariate effect of the different variables on survival time. Results are presented as hazard ratios (HRs) together with 95% confidence intervals. All reported p-values are two-sided and refer to Wald tests.

### 3.2.3 Paper III

#### 3.2.3.1 Sample preparation

In the discovery phase whole-exome sequencing was performed in the samples from *CHEK2*del1100C* carriers, colorectal cancer patients and patients with familial breast cancer. Genomic DNA was extracted from peripheral blood samples and subjected to whole exome sequencing at the National Genomics Infrastructure in Uppsala, Sweden. Exome-enriched sequencing libraries were prepared using the Agilent SureSelectXT Human All Exon V5 XT2 + UTR kit (Agilent, Santa Clara, California, US). Cluster generation and 125

cycle paired-end sequencing was performed using the Illumina HiSeq 2500 system and v4 sequencing chemistry (Illumina, San Diego, California, US). Next generation sequencing was performed at SciFiLab, University of Uppsala.

#### 3.2.3.2 Selection of candidate alleles

The cases with hereditary colon cancer and cases with hereditary breast cancer served as controls of variants identified by sequencing in the *CHEK2*del1100C* cases. All non-synonymous mutations in *CHEK2*del1100C* cases were evaluated to select candidate alleles for further evaluation in a second validation phase. The mutations were assessed by the following criteria that all should be fulfilled:
Allele frequency: Ratios of the allele frequencies of the mutations were calculated. A ratio of two or more between *CHEK2*del1100C* cases and colon cancer cases and/or a ratio of 1.5 or more between *CHEK2*del1100C* cases and familial breast cancer cases was required (the group familial breast cancer 56 cases included also the 28 *CHEK2*del1100C* cases which is why a lower ratio was required).
Gene function: Genes/mutations that were selected should display a function of a putative cancer driver gene when evaluated by on-line genome browser databases (OMIM, GeneCards) and scientific publications available on PubMed.
Reference databases: A more than 30% higher allele frequency in *CHEK2*del1100C* carriers compared to regional reference databases was required (ExAC non Finnish population, 1000genome2014oct European, 249 Swedes/Swedish individuals, SweGen Variant Frequency Browser).

Variant tools: The possible impact of an amino acid substitution was assessed by nine bioinformatics/variant tools (SIFT, Polyphen2 HDIV/HVAR, LRT, Mutation Taster, FATHMM, RadialSVM, LR and MutationsAssessor). Four of the nine tools, or more, should indicate that the variant was at least possibly damaging.
Sequencing accuracy: Only mutations with a sequencing accuracy of 65% or more were included, a criteria that was applied on all study groups.

All mutations fulfilling criteria were subject for validation. In a second round, mutations almost fulfilling criteria were evaluated again and four more mutations were added, (rs811925, rs34983477, rs34523498, rs4926600) that did not fulfil the variant tool criteria (3/9). In addition, a missense mutation in the breast cancer gene *PALB2*, rs152451, was also selected although it has previously been suggested as a polymorphism (table 1).

#### 3.2.3.3 Validation of candidate alleles

Eleven SNPs (rs2297809, rs17860405, rs8176786, rs34523498, rs117739035, rs34983477, rs152451, rs811925, rs7962217, rs34492126 and rs2287749) were genotyped using TaqMan SNP genotyping assay (Thermo Fisher Scientific, Waltham, Massachusetts, USA). rs35932273 was genotyped by Sanger sequencing following PCR. Genomic DNA used in the validation screen was extracted from peripheral blood by phenol method or a Qiagen DNA extraction kit for blood (Qiagen, Duesseldorf, Germany).

#### 3.2.3.4 Statistical analysis

Odds ratios, 95% confidence intervals and p-values were calculated to test the association with allele frequency using the DeFinetti programme provided as an online source (https://ihg.gsf.de/cgi-bin/hw/hwa1.pl).

# 4 RESULTS AND DISCUSSION

## 4.1 PAPER I

Breast cancer inherited with other tumour types represents breast cancer syndromes, such as the breast- and ovarian cancer syndrome in families with *BRCA1* or *BRCA2* mutations. In this study tumour spectrum in non-*BRCA* families was investigated in search of associations between breast cancer and other types of cancer in terms of proportions of tumour types compared to the Swedish general population in the reference years 1970 and 2010. We found increased proportions of several tumour types in the 334 non-*BRCA* families including endometrial cancer, an association that had been reported before [190]. In addition, liver-, prostate- ovarian-, pancreatic- and liver cancer were overrepresented (Table 1). A few of the cancer types were present in a smaller proportion than expected. All of the underrepresented tumour types except for lymphoma (2.97%, CI 1.84-4.24, proportion 5.75% 1970 and 6.28% 2010), were rare types of cancer with expected proportion of less than 1%.

| Cancer site | Observed number | Proportion [%] | LL 95% | UL 95% | Proportion [%] in Sweden 1970** | Proportion [%] in Sweden 2010** | Reference outside CI |
|---|---|---|---|---|---|---|---|
| **Prostate** | 103 | 14.57 | 12.02 | 17.26 | 6.98 | 11.95 | **CI above reference** |
| **Ovary/Fallopian tube** | 73 | 10.33 | 8.2 | 12.59 | 4.51 | 1.93 | **CI above reference** |
| **Uterus** | 45 | 6.36 | 4.67 | 8.2 | 3.07 | 2.64 | **CI above reference** |
| **Pancreas** | 33 | 4.67 | 3.11 | 6.22 | 2.79 | 1.54 | **CI above reference** |
| **Liver** | 15 | 2.12 | 1.13 | 3.25 | 0.28 | 0.13 | **CI above reference** |

Table 1. Overrepresented tumour types in the non-*BRCA* families.

A strength with the study was using a well-defined cohort where pathogenic *BRCA* mutations had been excluded. The risk of misclassification of gynaecological cancers due to recall bias is considered low since the majority of these diagnoses where verified through medical records or death certificates. One limitation with using this cohort is the risk of selection bias affecting the result, especially for ovarian cancer since patients are referred to counselling because of the presence of ovarian cancer in the family. The overrepresentation of ovarian cancer is certainly at least partly due to selection bias. The method, calculation of proportions, requires awareness of that changes in individual tumour proportions affect each other.

Endometrial cancer was present in 45 cases in 43 of the non-*BRCA f*amilies. Eleven of these families had only cases of breast- and endometrial cancer. There were 44 cases of 17 other tumour types in the remaining 32 families were various (Table 2).

| Cancer site | Number of cases |
|---|---|
| Stomach | 5 |
| Lung | 4 |
| Ovary/Fallopian tube | 4 |
| Uterine cervix | 4 |
| Kidney | 4 |
| Brain, nervous system | 3 |
| Colon | 3 |
| Prostate | 3 |
| Pancreatic | 2 |
| Rectum and anus | 2 |
| Lymphoma | 2 |
| Malignt melanoma | 2 |
| Skin | 2 |
| Testis | 1 |
| Larynx | 1 |
| Urinary tract | 1 |
| Multiple myeloma | 1 |
| Total number | 44 |

Table 2. Number of cases with other tumour types in the 32 out of 43 families with breast-, endometrial cancer and additional cancer diagnoses.

Colorectal-, ovarian-, urinary tract and gastric cancer are tumour types that are present in Hereditary non-polyposis colorectal cancer (HNPCC), caused by inherited defects in DNA mismatch repair genes. Endometrial cancer has been associated with HNPCC and an association with breast cancer has been indicated, but evidence is equivocal [201]. A family history indicating HNPCC would have been captured and excluded genetically in the counselling procedure, why it is unlikely that there are undetected Lynch syndrome families in the cohort. Results from all genetic testing for cancer predisposition, not only breast cancer syndromes, were investigated before inclusion in the study.

We now have strong support for an association between breast- and endometrial cancer, suggesting that the tumour types may, besides hormonal risk factors, also share genetic risk factors. A shared genetic susceptibility could involve oestrogen metabolism or cancer driver genes which warrant candidate gene association studies. Another approach is to investigate if breast- and endometrial cancer share low-risk loci identified in genome-wide association studies. In a recent study a breast- and endometrial cancer syndrome was again supported as well as our finding of an association between breast cancer and prostate- ovarian- and pancreatic cancer, suggesting shared susceptibility [191].

## 4.2 PAPER II

We aimed to investigate the effect of parent of origin on breast cancer prognosis. No significant difference in overall or recurrence-free survival between maternal and paternal inheritance of breast cancer was observed with HRs of 0.99 (95% CI=0.54 to 1.80) and 1.22 (95% CI=0.78 to 1.92) respectively. In a previous study a clear difference in prognosis was observed with a worse prognosis when breast cancer was inherited paternally. Furthermore, the difference increased with time. The present study was dimensioned with this taken into account. However, rather few families fulfilled inclusion criteria; only 23% of them could be included. A common reason for exclusion was inheritance of breast cancer on both lineages or presence of only one case of breast cancer or only index patient and her mother (parents were not counted as cases). The cohort was otherwise representative of the general breast cancer population in Sweden except for low mean age at onset, which is expected in familial breast cancer.

An interesting finding was the predominance of maternally inherited breast cancer. The asymmetry was also observed in the index patients collected from the cohort of unselected sporadic breast cancer (3.1.2.1), cases with maternal inheritance were twice as many as paternally. This may indicate that patients with paternal inheritance of breast cancer are less aware of their family history. However, the accuracy of self-reported history of breast cancer in first and second relative with breast cancer is in general high, which indicates that an underlying genetic mechanism may partly explain the asymmetry [202, 203].

We hypothesised that genomic imprinting, or other non-standard genetic mechanisms, may affect breast cancer prognosis in a parent-of-origin dependent manner. Imprinting has been shown to contribute to phenotypic variation even though imprinting has been confirmed in

less than 1% of all genes [180]. In general, genome-wide association studies do not differentiate between paternal and maternal alleles why little is known about the number and effect of imprinted genes. In that perspective it would be of interest to investigate, in a larger cohort, if a smaller parent-or-origin effect on breast cancer exists.

### 4.3 PAPER III

In paper III we investigated cases of CHEK2*1100delC carriers in search of genetic variants that may modify breast cancer risk in this patient group. After selection of candidate alleles, 11 selected variants (rs2297809, rs17860405, rs8176786, rs34523498, rs117739035, rs34983477, rs152451, rs811925, rs7962217, rs34492126 and rs2287749) were tested in a validation. No CHEK2 specific modifier could be identified, none of the variants showed significant difference in allele frequency in CHEK2*1100delC carriers compared to controls.

The investigated samples from *CHEK2*1100delC* carriers were all identified cases over the last five years in Sweden, which makes it all well-defined cohort. Criteria were constructed to assess the non-synonymous mutations in the carriers contra the controls of cases of familial colorectal- and familial breast cancer in order to identify candidate variants. Since none of the selected variants in CHEK2*1100delC carriers were confirmed as candidate alleles in the validation we need to re-evaluate the criteria in favour of another approach. Searching for more rare truncating variants is one option, another to investigate several individuals in selected families. The limited number of CHEK2 cases has implications on statistical power. In the future, international collaborations would be valuable, at least for validation of putative candidate alleles.

*CHEK2*1100delC* carriers have a 2-3-fold risk of breast cancer but the risk is considerably higher in carriers with a family history [120, 126, 132]. The individual risk for carriers is thereby difficult to predict. This is a clinical dilemma, since carriers in the lower risk spectrum may be near the risk of breast cancer of the general population, whereas many other carriers are at high risk and should be offered preventive programmes. Seventy-seven common low-risk breast cancer variants have demonstrated a modifying effect in *CHEK2*1100delC* carriers explaining a part of the excess familial risk [166]. Further studies are warranted to identify more genetic variants that modify risk in these families.

Two missense mutations, one in the DLG1 gene and one in the PRDM1 gene showed higher allele frequencies in both the familial cohort and the CHEK2*del1100C cohort compared to controls (rs34492126 OR: 1.19 and 1.15 and rs811925 OR: 1.19 and 1.09), but the differences were non-significant. Since these genes are interesting as putative cancer driver genes, testing the variants in larger cohorts would be of interest to investigate if they may be low-risk variants predisposing to breast cancer.

# 5 CONCLUSIONS AND CLINICAL IMPLICATIONS

Paper I

We found strong support for an association between breast cancer and endometrial cancer, which is a first step towards detecting new susceptibility variants. Future studies are warranted to identify shared genetic susceptibility that may explain some of the residual genetic component behind breast cancer. More identified risk genes/variants can improve the individual risk prediction in the families at risk.

Paper II

We found no parent-of-origin effect on breast cancer prognosis. A study in a larger cohort could confirm or rule out a smaller effect. Little is known about genetic mechanisms that confers an inherited inter-lineage effect. An incidental finding of our result, the predominance of maternally inherited cases, indicates that parent-of-origin may not have an effect on breast cancer prognosis, but rather the risk of being affected. Future studies could target both of these issues.

Paper III

We could not identify any variants that specifically modify breast cancer risk in *CHEK2*1100delC* carriers. Because of the imprecise risk prediction available for *CHEK2*1100delC* carriers, screening is so far not clinical routine in familial breast cancer. Continuous studies of genetic modifiers are thereby of importance. For counselling purpose, a future model that combine estimated relative risks and convert them into calculated absolute risks would be beneficial.

# 6 ACKNOWLEDGEMENTS

**Annika Lindblom,** my brilliant main supervisor, for sharing your endless knowledge in the cancer genetic research field and for your ability to inspire me and others. For always encouraging me and believing in me. For always making time for questions. For rewarding discussions and good laughs.

**Sara Margolin,** my equally brilliant supervisor. You are deeply committed to both breast cancer research and clinics, and generously share your profound knowledge. For helping me keep my calm and for great clinical tutoring, as my mentor, during my first years at SÖS. For good advice in other perspectives of life and for good laughs.

**Anna von Wachenfeldt Väppling,** for great inspiration and for getting me interested in the field from the beginning. Besides working hard at Södra Station, we have lots of fun together with **Linda Thorén,** my dear colleague and friend since we both started at the Oncology unit at SÖS, and **Lotta Bodell,** also a dear dedicated colleague. Together they form the rest of the team working with counselling for hereditary breast cancer families at the Department of Oncology at SÖS.

**Brita Arver**, for interesting discussions and valuable comments on manuscripts.

My colleagues at Annika´s lab, **Jessada Thutkawkorapin** and **Xiang Jiao,** and former colleague **Tao Liu,** for tutoring me in oncogenetics, for invaluable work in the lab and for good company.

**Hemming Johansson** for valuable help with the statistics.

**All collaborators** in the SWEA-study.

**Erika Rindsjö**, for excellent help with administrative issues.

**Marianne Iiristo**, head of the breast-section at the Department of Oncology at Södersjukhuset and Karolinska University Hospital for many years. For giving me the opportunity to do research, always struggling to make room for science. For many rewarding discussions in clinical issues.

**Ingveldur Björnsdottir,** for being so supportive and for always solving insoluble problems.

**Ulla Blom Goldman, Gerard Winblad, Tommy Fornander, Jenny Lundin, Ellinor Elinder, Ylva Sandeder, Elin Barnekow** and **Anna Larsson Wråke,** my dear colleagues and friends at the breast-section at the Department of Oncology at Södersjukhuset. For always being encouraging and supportive! And for all the fun moments and interesting discussions.

**Present and former colleagues** at Annika´s lab and the Department of Clinical Genetics, for valuable support and for creating a stimulating atmosphere, especially **Johanna Rantala, Kristina Lagerstedt-Robinson, Erik Björck** and **Emma Tham**.

The dedicated colleagues working with oncogenetic counselling at the Department of Oncology, Karolinska/RAH, **Svetlana Bajalica Lagercrantz, Ulla Platten, Anne Kinhult Ståhlbom, Annelie Liljegren** and **Daria Glaessgen,** for good cooperation and interesting discussions.

**All present and former nurses** at the breast-section at the Department of Oncology, Karolinska at SÖS, nowadays the Department of Oncology, SÖS, for invaluable work in the daily care of our out-patients and for good cooperation.

**All other present and former colleagues** at the Department of Oncology, Karolinska at SÖS, nowadays the Department of Oncology, SÖS.

And, of course, **all the patients** who participated.

My mother, **Kristina**, this would not have been possible without you giving a helping hand and keeping things running, and my father, **Jan-Allan,** because you were always there for me, helping out in big things and in small.

**Mårten,** my beloved husband, super fun to be with, smart and always supportive. I love you!

And finally, my children, **Agnes, Jonah** and **Carla,** love you endlessly.

# 7 REFERENCES


1. Lichtenstein, P., et al., *Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland.* N Engl J Med, 2000. **343**(2): p. 78-85.

2. Pharoah, P.D., et al., *Family history and the risk of breast cancer: a systematic review and meta-analysis.* Int J Cancer, 1997. **71**(5): p. 800-9.

3. Smithers, D.W., *Family Histories of 459 Patients with Cancer of the Breast.* British Journal of Cancer, 1948. **2**(2): p. 163-167.

4. Meijers-Heijboer, H., et al., *Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations.* Nat Genet, 2002. **31**(1): p. 55-9.

5. Vahteristo, P., et al., *A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer.* Am J Hum Genet, 2002. **71**(2): p. 432-8.

6. Claus, E.B., N. Risch, and W.D. Thompson, *Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction.* Cancer, 1994. **73**(3): p. 643-651.

7. Hulka, B.S. and A.T. Stark, *Breast cancer: cause and prevention.* Lancet, 1995. **346**(8979): p. 883-7.

8. *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, 2002. **360**(9328): p. 187-95.

9. van den Brandt, P.A., et al., *Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk.* Am J Epidemiol, 2000. **152**(6): p. 514-27.

10. *Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies.* Lancet, 1996. **347**(9017): p. 1713-27.

11. *Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer.* Lancet, 1997. **350**(9084): p. 1047-59.

12. Thomas, H.V., G.K. Reeves, and T.J. Key, *Endogenous estrogen and postmenopausal breast cancer: a quantitative review.* Cancer Causes Control, 1997. **8**(6): p. 922-8.

13. Hankinson, S.E., et al., *Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women.* J Natl Cancer Inst, 1998. **90**(17): p. 1292-9.

14. McCormack, V.A. and I. dos Santos Silva, *Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(6): p. 1159-69.

15. Byrne, C., et al., *Effects of mammographic density and benign breast disease on breast cancer risk (United States).* Cancer Causes Control, 2001. **12**(2): p. 103-10.

16. Page, D.L., et al., *Atypical lobular hyperplasia as a unilateral predictor of breast cancer risk: a retrospective cohort study.* Lancet, 2003. **361**(9352): p. 125-9.

17. Hancock, S.L., M.A. Tucker, and R.T. Hoppe, *Breast Cancer After Treatment of Hodgkin's Disease.* JNCI: Journal of the National Cancer Institute, 1993. **85**(1): p. 25-31.
18. Schatzkin, A. and M.P. Longnecker, *Alcohol and breast cancer. Where are we now and where do we go from here?* Cancer, 1994. **74**(3 Suppl): p. 1101-10.
19. Smith-Warner, S.A., et al., *Alcohol and breast cancer in women: a pooled analysis of cohort studies.* Jama, 1998. **279**(7): p. 535-40.
20. Adami, H.O., et al., *Familiality in breast cancer: a case-control study in a Sweden population.* Br J Cancer, 1980. **42**(1): p. 71-7.
21. Tulinius, H., et al., *Epidemiology of breast cancer in families in Iceland.* J Med Genet, 1992. **29**(3): p. 158-64.
22. Claus, E.B., N.J. Risch, and W.D. Thompson, *Age at onset as an indicator of familial risk of breast cancer.* Am J Epidemiol, 1990. **131**(6): p. 961-72.
23. Anderson, D.E., *Breast cancer in families.* Cancer, 1977. **40**(4 Suppl): p. 1855-60.
24. Claus, E.B., N. Risch, and W.D. Thompson, *Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction.* Cancer, 1994. **73**(3): p. 643-51.
25. Go, R.C., et al., *Genetic epidemiology of breast cancer and associated cancers in high-risk families. I. Segregation analysis.* J Natl Cancer Inst, 1983. **71**(3): p. 455-61.
26. Newman, B., et al., *Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families.* Proc Natl Acad Sci U S A, 1988. **85**(9): p. 3044-8.
27. Williams, W.R. and D.E. Anderson, *Genetic epidemiology of breast cancer: segregation analysis of 200 Danish pedigrees.* Genet Epidemiol, 1984. **1**(1): p. 7-20.
28. Peto, J. and T.M. Mack, *High constant incidence in twins and other relatives of women with breast cancer.* Nat Genet, 2000. **26**(4): p. 411-414.
29. Ghoussaini, M. and P.D. Pharoah, *Polygenic susceptibility to breast cancer: current state-of-the-art.* Future Oncol, 2009. **5**(5): p. 689-701.
30. Hall, J.M., et al., *Linkage of early-onset familial breast cancer to chromosome 17q21.* Science, 1990. **250**(4988): p. 1684-9.
31. Narod, S.A., et al., *Familial breast-ovarian cancer locus on chromosome 17q12-q23.* Lancet, 1991. **338**(8759): p. 82-3.
32. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
33. Stratton, M.R., et al., *Familial male breast cancer is not linked to the BRCA1 locus on chromosome 17q.* Nat Genet, 1994. **7**(1): p. 103-107.
34. Wooster, R., et al., *Identification of the breast cancer susceptibility gene BRCA2.* Nature, 1995. **378**(6559): p. 789-92.
35. Wooster, R., et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.* Science, 1994. **265**(5181): p. 2088-90.

36. Li, F.P. and J.F. Fraumeni, Jr., *Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome?* Ann Intern Med, 1969. **71**(4): p. 747-52.

37. Malkin, D., et al., *Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms.* Science, 1990. **250**(4985): p. 1233-8.

38. Srivastava, S., et al., *Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome.* Nature, 1990. **348**(6303): p. 747-9.

39. Easton, D.F., *Cancer risks in A-T heterozygotes.* Int J Radiat Biol, 1994. **66**(6 Suppl): p. S177-82.

40. Renwick, A., et al., *ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles.* Nat Genet, 2006. **38**(8): p. 873-5.

41. Savitsky, K., et al., *A single ataxia telangiectasia gene with a product similar to PI-3 kinase.* Science, 1995. **268**(5218): p. 1749-53.

42. Rahman, N., et al., *PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene.* Nat Genet, 2007. **39**(2): p. 165-7.

43. Casadei, S., et al., *Contribution of inherited mutations in the BRCA2-interacting protein PALB2 to familial breast cancer.* Cancer Res, 2011. **71**(6): p. 2222-9.

44. Cybulski, C., et al., *Germline RECQL mutations are associated with breast cancer susceptibility.* Nat Genet, 2015. **47**(6): p. 643-6.

45. Altshuler, D., M.J. Daly, and E.S. Lander, *Genetic mapping in human disease.* Science, 2008. **322**(5903): p. 881-8.

46. Michailidou, K., et al., *Large-scale genotyping identifies 41 new loci associated with breast cancer risk.* Nat Genet, 2013. **45**(4): p. 353-61, 361e1-2.

47. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. **447**(7148): p. 1087-93.

48. Hunter, D.J., et al., *A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.* Nat Genet, 2007. **39**(7): p. 870-4.

49. Stacey, S.N., et al., *Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.* Nat Genet, 2007. **39**(7): p. 865-9.

50. Stacey, S.N., et al., *Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer.* Nat Genet, 2008. **40**(6): p. 703-6.

51. Ahmed, S., et al., *Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2.* Nat Genet, 2009. **41**(5): p. 585-90.

52. Zheng, W., et al., *Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1.* Nat Genet, 2009. **41**(3): p. 324-8.

53. Thomas, G., et al., *A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1).* Nat Genet, 2009. **41**(5): p. 579-84.

54. Turnbull, C., et al., *Genome-wide association study identifies five new breast cancer susceptibility loci.* Nat Genet, 2010. **42**(6): p. 504-7.

55. Antoniou, A.C., et al., *A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population.* Nat Genet, 2010. **42**(10): p. 885-92.

56. Fletcher, O., et al., *Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study.* J Natl Cancer Inst, 2011. **103**(5): p. 425-35.

57. Haiman, C.A., et al., *A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer.* Nat Genet, 2011. **43**(12): p. 1210-4.

58. Ghoussaini, M., et al., *Genome-wide association analysis identifies three new breast cancer susceptibility loci.* Nat Genet, 2012. **44**(3): p. 312-8.

59. Siddiq, A., et al., *A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11.* Hum Mol Genet, 2012. **21**(24): p. 5373-84.

60. Gaudet, M.M., et al., *Identification of a BRCA2-specific modifier locus at 6p24 related to breast cancer risk.* PLoS Genet, 2013. **9**(3): p. e1003173.

61. Couch, F.J., 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): p. e1003212.

62. French, J.D., 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): p. 489-503.

63. Garcia-Closas, M., et al., *Genome-wide association studies identify four ER negative-specific breast cancer risk loci.* Nat Genet, 2013. **45**(4): p. 392-8, 398e1-2.

64. Bojesen, S.E., et al., *Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer.* Nat Genet, 2013. **45**(4): p. 371-84, 384e1-2.

65. Chen, F., et al., *A genome-wide association study of breast cancer in women of African ancestry.* Hum Genet, 2013. **132**(1): p. 39-48.

66. Long, J., et al., *Genome-wide association study in east Asians identifies novel susceptibility loci for breast cancer.* PLoS Genet, 2012. **8**(2): p. e1002532.

67. Milne, R.L., et al., *Common non-synonymous SNPs associated with breast cancer susceptibility: findings from the Breast Cancer Association Consortium.* Hum Mol Genet, 2014. **23**(22): p. 6096-111.

68. Cai, Q., et al., *Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1.* Nat Genet, 2014. **46**(8): p. 886-90.

69. Michailidou, K., et al., *Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer.* Nat Genet, 2015. **47**(4): p. 373-380.

70. Michailidou, K., et al., *Association analysis identifies 65 new breast cancer risk loci. .* Nature;. **23 Oct 2017; DOI: 10.1038/nature24284**.

71. Barnhart, B.C., et al., *The death effector domain protein family.* Oncogene, 2003. **22**(53): p. 8634-44.

72. Kilpivaara, O., et al., *CHEK2 variant I157T may be associated with increased breast cancer risk.* Int J Cancer, 2004. **111**(4): p. 543-7.

73. Goldgar, D.E., et al., *Rare variants in the ATM gene and risk of breast cancer.* Breast Cancer Res, 2011. **13**(4): p. R73.

74. Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection.* Nat Rev Cancer, 2012. **12**(1): p. 68-78.

75. Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection.* Nat Rev Cancer, 2011. **12**(1): p. 68-78.

76. King, M.C., J.H. Marks, and J.B. Mandell, *Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2.* Science, 2003. **302**(5645): p. 643-6.

77. Fackenthal, J.D. and O.I. Olopade, *Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations.* Nat Rev Cancer, 2007. **7**(12): p. 937-48.

78. Antoniou, 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**(5): p. 1117-30.

79. Socialstyrelsen, *Cancer Incidence in Sweden.* 2014.

80. Spurdle, A.B., et al., *BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk.* J Med Genet, 2012. **49**(8): p. 525-32.

81. Ford, D., et al., *Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium.* Lancet, 1994. **343**(8899): p. 692-5.

82. Moran, A., et al., *Risk of cancer other than breast or ovarian in individuals with BRCA1 and BRCA2 mutations.* Fam Cancer, 2012. **11**(2): p. 235-42.

83. *Cancer risks in BRCA2 mutation carriers.* J Natl Cancer Inst, 1999. **91**(15): p. 1310-6.

84. van Asperen, C.J., et al., *Cancer risks in BRCA2 families: estimates for sites other than breast and ovary.* J Med Genet, 2005. **42**(9): p. 711-9.

85. Chompret, A., et al., *P53 germline mutations in childhood cancers and cancer risk for carrier individuals.* Br J Cancer, 2000. **82**(12): p. 1932-7.

86. Mai, P.L., et al., *Risks of first and subsequent cancers among TP53 mutation carriers in the National Cancer Institute Li-Fraumeni syndrome cohort.* Cancer, 2016. **122**(23): p. 3673-3681.

87. Hwang, S.J., et al., *Germline p53 mutations in a cohort with childhood sarcoma: sex differences in cancer risk.* Am J Hum Genet, 2003. **72**(4): p. 975-83.

88. Bougeard, G., et al., *Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families.* J Med Genet, 2008. **45**(8): p. 535-8.

89. Petitjean, A., et al., *Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database.* Hum Mutat, 2007. **28**(6): p. 622-9.

90. Birch, J.M., et al., *Cancer phenotype correlates with constitutional TP53 genotype in families with the Li-Fraumeni syndrome.* Oncogene, 1998. **17**(9): p. 1061-8.

91. Olivier, M., et al., *Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and TP53 genotype.* Cancer Res, 2003. **63**(20): p. 6643-50.

92. Nelen, M.R., et al., *Localization of the gene for Cowden disease to chromosome 10q22-23.* Nat Genet, 1996. **13**(1): p. 114-6.

93. Farooq, A., et al., *Cowden syndrome.* Cancer Treat Rev, 2010. **36**(8): p. 577-83.

94. Eng, C., *PTEN: one gene, many syndromes.* Hum Mutat, 2003. **22**(3): p. 183-98.

95. Tan, M.H., et al., *Lifetime cancer risks in individuals with germline PTEN mutations.* Clin Cancer Res, 2012. **18**(2): p. 400-7.

96. Nieuwenhuis, M.H., et al., *Cancer risk and genotype-phenotype correlations in PTEN hamartoma tumor syndrome.* Fam Cancer, 2014. **13**(1): p. 57-63.

97. Tomlinson, I.P. and R.S. Houlston, *Peutz-Jeghers syndrome.* J Med Genet, 1997. **34**(12): p. 1007-11.

98. Syngal, S., et al., *ACG clinical guideline: Genetic testing and management of hereditary gastrointestinal cancer syndromes.* Am J Gastroenterol, 2015. **110**(2): p. 223-62; quiz 263.

99. Hearle, N., et al., *Frequency and spectrum of cancers in the Peutz-Jeghers syndrome.* Clin Cancer Res, 2006. **12**(10): p. 3209-15.

100. Takeichi, M., *Cadherin cell adhesion receptors as a morphogenetic regulator.* Science, 1991. **251**(5000): p. 1451-5.

101. Christofori, G. and H. Semb, *The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene.* Trends Biochem Sci, 1999. **24**(2): p. 73-6.

102. Pharoah, P.D., P. Guilford, and C. Caldas, *Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families.* Gastroenterology, 2001. **121**(6): p. 1348-53.

103. Keller, G., 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): p. 337-42.

104. Kangelaris, K.N. and S.B. Gruber, *Clinical implications of founder and recurrent cdh1 mutations in hereditary diffuse gastric cancer.* JAMA, 2007. **297**(21): p. 2410-2411.

105. Corso, G., et al., *CDH1 germline mutations and hereditary lobular breast cancer.* Familial Cancer, 2016. **15**(2): p. 215-219.

106. Benusiglio, P.R., et al., *<em>CDH1</em> germline mutations and the hereditary diffuse gastric and lobular breast cancer syndrome: a multicentre study.* Journal of Medical Genetics, 2013. **50**(7): p. 486-489.

107. Petridis, C., et al., *Germline CDH1 mutations in bilateral lobular carcinoma in situ.* Br J Cancer, 2014. **110**(4): p. 1053-7.

108. Suriano, G., et al., *E-cadherin germline missense mutations and cell phenotype: evidence for the independence of cell invasion on the motile capabilities of the cells.* Hum Mol Genet, 2003. **12**(22): p. 3007-16.

109. Antoniou , A.C., et al., *Breast-Cancer Risk in Families with Mutations in PALB2.* New England Journal of Medicine, 2014. **371**(6): p. 497-506.

110. Xia, B., et al., *Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2.* Mol Cell, 2006. **22**(6): p. 719-29.

111. Zhang, F., et al., *PALB2 links BRCA1 and BRCA2 in the DNA-damage response.* Curr Biol, 2009. **19**(6): p. 524-9.

112. Ramus, S.J., et al., *Germline Mutations in the BRIP1, BARD1, PALB2, and NBN Genes in Women With Ovarian Cancer.* JNCI: Journal of the National Cancer Institute, 2015. **107**(11): p. djv214-djv214.

113. Jones, S., et al., *Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene.* Science, 2009. **324**(5924): p. 217.

114. Matsuoka, S., M. Huang, and S.J. Elledge, *Linkage of ATM to cell cycle regulation by the Chk2 protein kinase.* Science, 1998. **282**(5395): p. 1893-7.

115. Nevanlinna, H. and J. Bartek, *The CHEK2 gene and inherited breast cancer susceptibility.* Oncogene, 2006. **25**(43): p. 5912-9.

116. Walsh, T., et al., *Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer.* Jama, 2006. **295**(12): p. 1379-88.

117. Iniesta, M.D., et al., *Absence of CHEK2*1100delC mutation in families with hereditary breast cancer in North America.* Cancer Genet Cytogenet, 2010. **202**(2): p. 136-40.

118. Apostolou, P. and F. Fostira, *Hereditary breast cancer: the era of new susceptibility genes.* BioMed research international, 2013. **2013**.

119. Margolin, S., et al., *CHEK2 1100delC is prevalent in Swedish early onset familial breast cancer.* BMC Cancer, 2007. **7**: p. 163.

120. Schmidt, M.K., et al., *Age- and Tumor Subtype-Specific Breast Cancer Risk Estimates for CHEK2*1100delC Carriers.* J Clin Oncol, 2016. **34**(23): p. 2750-60.

121. Bogdanova, N., et al., *Association of two mutations in the CHEK2 gene with breast cancer.* Int J Cancer, 2005. **116**(2): p. 263-6.

122. Cybulski, C., et al., *A deletion in CHEK2 of 5,395 bp predisposes to breast cancer in Poland.* Breast Cancer Res Treat, 2007. **102**(1): p. 119-22.

123. Desrichard, A., et al., *CHEK2 contribution to hereditary breast cancer in non-BRCA families.* Breast Cancer Research : BCR, 2011. **13**(6): p. R119-R119.

124. Kriege, M., et al., *Survival and contralateral breast cancer in CHEK2 1100delC breast cancer patients: impact of adjuvant chemotherapy.* British Journal of Cancer, 2014. **111**(5): p. 1004-1013.

125. Weischer, M., et al., *CHEK2*1100delC heterozygosity in women with breast cancer associated with early death, breast cancer-specific death, and increased risk of a second breast cancer.* J Clin Oncol, 2012. **30**(35): p. 4308-16.

126. Weischer, M., et al., *CHEK2*1100delC genotyping for clinical assessment of breast cancer risk: meta-analyses of 26,000 patient cases and 27,000 controls.* J Clin Oncol, 2008. **26**(4): p. 542-8.

127. Walsh, T., et al., *Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing.* Proc Natl Acad Sci U S A, 2011. **108**(44): p. 18032-7.

128. Cybulski, C., et al., *CHEK2 is a multiorgan cancer susceptibility gene.* Am J Hum Genet, 2004. **75**(6): p. 1131-5.

129. Xiang, H.P., et al., *Meta-analysis of CHEK2 1100delC variant and colorectal cancer susceptibility.* Eur J Cancer, 2011. **47**(17): p. 2546-51.

130. Wang, Y., et al., *Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer.* Nat Genet, 2014. **46**(7): p. 736-41.

131. Ahmed, M. and N. Rahman, *ATM and breast cancer susceptibility.* Oncogene, 0000. **25**(43): p. 5906-5911.

132. Easton, D.F., et al., *Gene-Panel Sequencing and the Prediction of Breast-Cancer Risk.* New England Journal of Medicine, 2015. **372**(23): p. 2243-2257.

133. Uusitalo, E., et al., *Incidence and mortality of neurofibromatosis: a total population study in Finland.* J Invest Dermatol, 2015. **135**(3): p. 904-906.

134. Seminog, O.O. and M.J. Goldacre, *Age-specific risk of breast cancer in women with neurofibromatosis type 1.* Br J Cancer, 2015. **112**(9): p. 1546-1548.

135. Madanikia, S.A., et al., *Increased risk of breast cancer in women with NF1.* Am J Med Genet A, 2012. **158a**(12): p. 3056-60.

136. Bogdanova, N., et al., *Nijmegen Breakage Syndrome mutations and risk of breast cancer.* Int J Cancer, 2008. **122**(4): p. 802-6.

137. Zhang, G., et al., *Significant association between Nijmegen breakage syndrome 1 657del5 polymorphism and breast cancer risk.* Tumour Biol, 2013. **34**(5): p. 2753-7.

138. Steffen, J., 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): p. 472-5.

139. Coulet, F., et al., *Germline RAD51C mutations in ovarian cancer susceptibility.* Clin Genet, 2013. **83**(4): p. 332-6.

140. Le Calvez-Kelm, F., et al., *RAD51 and breast cancer susceptibility: no evidence for rare variant association in the Breast Cancer Family Registry study.* PLoS One, 2012. **7**(12): p. e52374.

141. Park, D.J., et al., *Rare mutations in XRCC2 increase the risk of breast cancer.* Am J Hum Genet, 2012. **90**(4): p. 734-9.

142. Rafnar, T., et al., *Mutations in BRIP1 confer high risk of ovarian cancer.* Nat Genet, 2011. **43**(11): p. 1104-7.

143. Seal, S., et al., *Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles.* Nat Genet, 2006. **38**(11): p. 1239-41.

144. Easton, D.F., et al., *No evidence that protein truncating variants in BRIP1 are associated with breast cancer risk: implications for gene panel testing.* J Med Genet, 2016. **53**(5): p. 298-309.

145. Thompson, E.R., et al., *Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles.* PLoS Genet, 2012. **8**(9): p. e1002894.

146. Gracia-Aznarez, F.J., et al., *Whole exome sequencing suggests much of non-BRCA1/BRCA2 familial breast cancer is due to moderate and low penetrance susceptibility alleles.* PLoS One, 2013. **8**(2): p. e55681.

147. Kiiski, J.I., et al., *Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(42): p. 15172-15177.

148. Damiola, F., 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, 2014. **16**(3): p. R58.

149. Figueroa, J.D., et al., *Associations of common variants at 1p11.2 and 14q24.1 (RAD51L1) with breast cancer risk and heterogeneity by tumor subtype: findings from the Breast Cancer Association Consortium.* Hum Mol Genet, 2011. **20**(23): p. 4693-706.

150. Broeks, A., et al., *Low penetrance breast cancer susceptibility loci are associated with specific breast tumor subtypes: findings from the Breast Cancer Association Consortium.* Hum Mol Genet, 2011. **20**(16): p. 3289-303.

151. Purrington, K.S., et al., *Genome-wide association study identifies 25 known breast cancer susceptibility loci as risk factors for triple-negative breast cancer.* Carcinogenesis, 2014. **35**(5): p. 1012-9.

152. Milne, R., . et al. , *Identification of ten variants associated with risk of estrogen receptor negative breast cancer.* . Nature Genetics; 23 Oct 2017; DOI: 10.1038/ng.3785.

153. *Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group.* Br J Cancer, 2000. **83**(10): p. 1301-8.

154. Palomaki, G.E., *Is it time for BRCA1/2 mutation screening in the general adult population?: impact of population characteristics.* Genet Med, 2015. **17**(1): p. 24-6.

155. Margolin, S., et al., *BRCA1 mutations in a population-based study of breast cancer in Stockholm County.* Genet Test, 2004. **8**(2): p. 127-32.

156. Loman, N., et al., *Family history of breast and ovarian cancers and BRCA1 and BRCA2 mutations in a population-based series of early-onset breast cancer.* J Natl Cancer Inst, 2001. **93**(16): p. 1215-23.

157. Peto, J., et al., *Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer.* J Natl Cancer Inst, 1999. **91**(11): p. 943-9.

158. Antoniou, A.C., et al., *Predicting the likelihood of carrying a BRCA1 or BRCA2 mutation: validation of BOADICEA, BRCAPRO, IBIS, Myriad and the Manchester scoring system using data from UK genetics clinics.* J Med Genet, 2008. **45**(7): p. 425-31.

159. Stratton, M.R. and N. Rahman, *The emerging landscape of breast cancer susceptibility.* Nat Genet, 2008. **40**(1): p. 17-22.

160. Antoniou, A.C., et al., *A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes.* Br J Cancer, 2002. **86**(1): p. 76-83.

161. Antoniou, A.C. and D.F. Easton, *Models of genetic susceptibility to breast cancer.* Oncogene, 2006. **25**(43): p. 5898-905.

162. Antoniou, A.C., et al., *Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers.* Am J Hum Genet, 2008. **82**(4): p. 937-48.

163. Antoniou, A.C., et al., *Common variants in LSP1, 2q35 and 8q24 and breast cancer risk for BRCA1 and BRCA2 mutation carriers.* Human Molecular Genetics, 2009. **18**(22): p. 4442-4456.

164. Antoniou, A.C., et al., *Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers.* Breast Cancer Res, 2012. **14**.

165. Antoniou, A.C., et al., *Common alleles at 6q25.1 and 1p11.2 are associated with breast cancer risk for BRCA1 and BRCA2 mutation carriers.* Human Molecular Genetics, 2011. **20**(16): p. 3304-3321.

166. Muranen, T.A., et al., *Genetic modifiers of CHEK2*1100delC associated breast cancer risk.* Genetics in medicine : official journal of the American College of Medical Genetics, 2016: p. 10.1038/gim.2016.147.

167. Turnbull, C., et al., *Gene-gene interactions in breast cancer susceptibility.* Hum Mol Genet, 2012. **21**(4): p. 958-62.

168. Allis, C.D. and T. Jenuwein, *The molecular hallmarks of epigenetic control.* Nat Rev Genet, 2016. **17**(8): p. 487-500.

169. Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer.* Nat Rev Genet, 2002. **3**(6): p. 415-28.

170. Lustberg, M.B. and B. Ramaswamy, *Epigenetic Therapy in Breast Cancer.* Current Breast Cancer Reports, 2011. **3**(1): p. 34-43.

171. Esteller, M., *CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future.* Oncogene, 2002. **21**(35): p. 5427-40.

172. Radpour, R., et al., *Integrated epigenetics of human breast cancer: synoptic investigation of targeted genes, microRNAs and proteins upon demethylation treatment.* PLoS One, 2011. **6**(11): p. e27355.

173. Hedenfalk , I., et al., *Gene-Expression Profiles in Hereditary Breast Cancer.* New England Journal of Medicine, 2001. **344**(8): p. 539-548.

174. Park, S.Y., et al., *Alu and LINE-1 hypomethylation is associated with HER2 enriched subtype of breast cancer.* PLoS One, 2014. **9**(6): p. e100429.

175. Damaskos, C., et al., *Histone Deacetylase Inhibitors: An Attractive Therapeutic Strategy Against Breast Cancer.* Anticancer Res, 2017. **37**(1): p. 35-46.

176. Christodoulatos, G.S. and M. Dalamaga, *Micro-RNAs as clinical biomarkers and therapeutic targets in breast cancer: Quo vadis?* World Journal of Clinical Oncology, 2014. **5**(2): p. 71-81.

177. Iorio, M.V., et al., *MicroRNA gene expression deregulation in human breast cancer.* Cancer Res, 2005. **65**(16): p. 7065-70.

178. Mattiske, S., et al., *The oncogenic role of miR-155 in breast cancer.* Cancer Epidemiol Biomarkers Prev, 2012. **21**(8): p. 1236-43.

179. Reik, W. and J. Walter, *Genomic imprinting: parental influence on the genome.* Nat Rev Genet, 2001. **2**(1): p. 21-32.

180. Morison, I.M., J.P. Ramsay, and H.G. Spencer, *A census of mammalian imprinting.* Trends Genet, 2005. **21**(8): p. 457-65.

181. Chong, S. and E. Whitelaw, *Epigenetic germline inheritance.* Curr Opin Genet Dev, 2004. **14**(6): p. 692-6.

182. Ferguson-Smith, A.C., *Genomic imprinting: the emergence of an epigenetic paradigm.* Nat Rev Genet, 2011. **12**(8): p. 565-575.

183. Kong, A., et al., *Parental origin of sequence variants associated with complex diseases.* Nature, 2009. **462**(7275): p. 868-74.

184. Zarrei, M., et al., *A copy number variation map of the human genome.* Nat Rev Genet, 2015. **16**(3): p. 172-83.

185. Fanciulli, M., E. Petretto, and T.J. Aitman, *Gene copy number variation and common human disease.* Clin Genet, 2010. **77**(3): p. 201-13.

186. Krepischi, A.C., P.L. Pearson, and C. Rosenberg, *Germline copy number variations and cancer predisposition.* Future Oncol, 2012. **8**(4): p. 441-50.

187. Palma, M.D., et al., *The relative contribution of point mutations and genomic rearrangements in BRCA1 and BRCA2 in high-risk breast cancer families.* Cancer Res, 2008. **68**(17): p. 7006-14.

188. Peltomaki, P., *Update on Lynch syndrome genomics.* Fam Cancer, 2016. **15**(3): p. 385-93.

189. Harkness, E.F., et al., *Lynch syndrome caused by MLH1 mutations is associated with an increased risk of breast cancer: a cohort study.* J Med Genet, 2015. **52**(8): p. 553-6.

190. von Wachenfeldt, A., et al., *A hypothesis-generating search for new genetic breast cancer syndromes--a national study in 803 Swedish families.* Hered Cancer Clin Pract, 2007. **5**(1): p. 17-24.

191. Zheng, G., et al., *Familial associations of female breast cancer with other cancers.* Int J Cancer, 2017.

192. Cao, S.-S. and C.-T. Lu, *Recent perspectives of breast cancer prognosis and predictive factors.* Oncology Letters, 2016. **12**(5): p. 3674-3678.

193. Hartman, M., et al., *Is breast cancer prognosis inherited?* Breast Cancer Res, 2007. **9**(3): p. R39.

194. Verkooijen, H.M., et al., *Breast cancer prognosis is inherited independently of patient, tumor and treatment characteristics.* Int J Cancer, 2012. **130**(9): p. 2103-10.

195. Weinberg, C.R., et al., *Asymmetry in family history implicates nonstandard genetic mechanisms: application to the genetics of breast cancer.* PLoS Genet, 2014. **10**(3): p. e1004174.

196. Lindblom, A., et al., *Hereditary breast cancer in Sweden: a predominance of maternally inherited cases.* Breast Cancer Res Treat, 1992. **24**(2): p. 159-65.

197. Chang, C.Q., et al., *A systematic review of cancer GWAS and candidate gene meta-analyses reveals limited overlap but similar effect sizes.* Eur J Hum Genet, 2014. **22**(3): p. 402-8.

198. *Swedish Breast Cancer Group, Guidelines 2014.*

199. *Cancer incidence in Sweden, Socialstyrelsen 2010.*

200. *Cancer incidence in Sweden, Socialstyrelsen 1970.*

201. Win, A.K., N.M. Lindor, and M.A. Jenkins, *Risk of breast cancer in Lynch syndrome: a systematic review.* Breast Cancer Res, 2013. **15**(2): p. R27.

202. Eerola, H., et al., *Familial breast cancer in southern Finland: how prevalent are breast cancer families and can we trust the family history reported by patients?* Eur J Cancer, 2000. **36**(9): p. 1143-8.

203. Parent, M.E., et al., *The reliability of recollections of family history: implications for the medical provider.* J Cancer Educ, 1997. **12**(2): p. 114-20.