Szilvia Solyom

BRCA/FANCONI ANEMIA PATHWAY GENES IN HEREDITARY PREDISPOSITION TO BREAST CANCER
SZILVIA SOLYOM

BRCA/FANCONI ANEMIA PATHWAY GENES IN HEREDITARY PREDISPOSITION TO BREAST CANCER

Academic dissertation to be presented with the assent of the Faculty of Medicine of the University of Oulu for public defence in Auditorium 5 of Oulu University Hospital, on 29 April 2011, at 12 noon

UNIVERSITY OF OULU, OULU 2011
Abstract

Two major genes are involved in hereditary predisposition to breast and ovarian cancer – BRCA1 and BRCA2. However, germline mutations in these tumor suppressors account for a maximum 20% of the familial breast cancer cases. A significant portion of the genes predisposing to this disease is unknown and therefore needs to be discovered. The aim of this study was to identify novel breast cancer susceptibility genes from the interweaving BRCA/Fanconi anemia (FA) pathway. Five candidate genes – MERIT40, ABRAXAS, BRIP1, CHK1, and FANCA – were screened for mutations by utilizing conformation-sensitive gel electrophoresis and sequencing, or with multiplex ligation-dependent probe amplification in blood DNA samples of Finnish familial breast cancer patients.

Investigation of the MERIT40 gene revealed novel nucleotide changes, being the first report on mutation screening of this gene. None of the observed alterations, however, appeared to be disease related, suggesting that germline mutations in MERIT40 are rare or absent in breast cancer patients.

A missense alteration (c.1082G>A, leading to Arg361Gln) was identified in ABRAXAS in 3 out of 125 Northern Finnish breast cancer families (2.4%), but not in any of the 867 healthy controls. The prevalence of the mutation between familial and control cases was statistically significantly different (p=0.002). ABRAXAS c.1082G>A appears to have pathological significance based on its exclusive occurrence in cancer cases, evolutionary conservation, disruption of a putative nuclear localization signal, reduced nuclear localization of the protein, and defective accumulation at DNA damage sites.

The BRIP1 (FANCJ) and CHK1 genes were screened for large genomic rearrangements, but no abnormalities were detected, ruling out a significant contribution to breast cancer susceptibility in the Northern Finnish population.

A novel large heterozygous deletion was identified in the FANCA gene in one out of 100 breast cancer families, removing the promoter and the first 12 exons. The deletion allele was not present in the tested controls, suggesting that it might contribute to breast cancer susceptibility. This is the first report on the association of a large-size germline deletion in a gene acting in the upstream part of the FA signaling pathway with familial breast cancer.

Keywords: ABRAXAS, breast cancer, BRIP1, CHK1, FANCA, FANCJ, Fanconi anemia, genetic predisposition to disease, germline mutation, MERIT40
**Tiivistelmä**


\textit{MERIT40}-geenissä havaittiin useita aikaisemmin raportoimattomia nukleotidimuutoksia, mutta yhdenkään niistä ei havaittu liittyvän rintasyöpäalttiuteen. \textit{MERIT40}-geenimuutosten mahdollista yhteyttä rintasyöpäalttiuteen ei ole tutkittu aikaisemmin.

\textit{ABRAXAS}-geenissä havaittiin missense-mutaatio (c.1082G>A, joka johtaa Arg361Gln aminohappokorvautumiseen) kolmessa pohjoissuomalaisessa rintasyöpäperheessä (3/125, 2.4 %). Muutosta ei havaittu terveissä kontrollieissa (N=867), ja ero mutaation esiintyvyydessä familiaalisten rintasyöpätapausten ja terveiden kontrollien välillä oli tilastollisesti merkitsevä (p=0.002). \textit{ABRAXAS} c.1082G>A-muutos on todennäköisesti patogeeninen, sillä kyseinen aminohappopaaikk on evolutiivisesti konservointunut ja sijaitsee todennäköisellä tumaanohaussignaalialueella. Funktionaliset kokeet osoittivat, että mutaatioon sisältynyt proteiinimuutos helpottaa tumaa ja sen ohjaaminen DNA-vaurioalueen päätyttyä.

\textit{BRIP1} (\textit{FANCJ}) ja \textit{CHK1}-geeneistä etsittiin laajoja genomisia uudelleenjärjestelyjä, mutta niitä ei havaittu. Näin ollen kyseisillä muutoksilla ei ole merkittävää roolia perinnöllisessä rintasyöpäalttiudessa suomalaisessa väestössä.

\textit{FANCA}-geenissä havaittiin laaja heterotsygoottinen deleetto yhdessä tutkistusta 100 rintasyöpäperheestä. Deleetto poistaa geenin promoottorialueen lisää sen 12 ensimmäistä eksonia. Deleettoalueella ei havaittu terveissä kontrolloissa, joten se mahdollisesti liittyy perinnölliseen rintasyöpäalttiuteen. Tutkimus on ensimmäinen, jossa raportoidaan laaja genomin deleetto FA-signaalinsiirtoreitien ylävirran geenissä familiaalisessa rintasyövää.

\textit{Asiasanat:} ABRAXAS, BRIP1, CHK1, FANCA, FANCJ, Fanconin anemia, ituradan muutos, MERIT40, perinnöllinen alttius, rintasyöpä
Acknowledgements

This study was carried out at the Laboratory of Cancer Genetics, Department of Clinical Genetics, Biocenter Oulu, University of Oulu and Oulu University Hospital during January 2008 - January 2011. I wish to thank all those who participated in the project.

My supervisor, Prof. Robert Winqvist is acknowledged for giving me the opportunity to join his research group and to participate in the most inspiring field of biology, cancer research. His guidance and support have been invaluable.

My second supervisor, Dr. Katri Pylkäs is equally acknowledged for her professional guidance and support.

I wish to thank all my present and former colleagues for the supportive atmosphere, their collaboration on research projects and for their unending help in practical matters in Oulu, from translating Finnish texts to apartment hunting: Jenni Nikkilä, M.Sc, Maria Haanpää, MD, Mikko Vuorela, MSc, Hannele Erkko, PhD, Sanna-Maria Karppinen, PhD, furthermore Meeri Otsukka and Helmi Konola for fantastic technical assistance. In addition, members of the Laboratory of Genetics are acknowledged for the friendly atmosphere they created at the Oulu University Hospital.

My co-authors and collaborators Prof. Roger Greenberg, Dr. Jaana Kauppinen, Dr. Hannaleena Kokkonen, Dr. Katrin Rapakko, Prof. Arto Mannermaa and Dr. Jeffrey Patterson-Fortin are thanked for their important contributions to my publications.

The official referees Dr. Marianne Tiainen and Dr. Hannu Norppa are thanked for reviewing the thesis. Ville Varjonen and Juhani Räisänen are acknowledged for their invaluable help with formatting the thesis and Anna Vuolteenaho for language review.

I wish to express my gratitude to my family and friends, without whom I might not have been able to complete my PhD studies. The support of Eva, my mother is thanked above all, because without her I would not have been able to make my dreams come true. I am happy that I have a sister, Dori, whose optimism often cheered me up. I am lucky that I had the opportunity to do my doctoral study in Finland, as I found true friends here. Particularly Liisa – together with Jorma – and Helmi, who stood by me all the time, no matter what.

All the patients and their family members are thanked for participating in this study.
Financial support of my PhD studies by CIMO, the Finnish Cancer Organizations, the Ida Montini Foundation, the Northern Finnish Cancer Foundation, the Sigrid Jusélius Foundation, the Academy of Finland, the University of Oulu, the Tyyni Tani Foundation and the Oulu University Hospital is gratefully acknowledged.

Oulu, January 2011

Szilvia Solyom
Abbreviations

AML  acute myeloid leukemia
AR   androgen receptor
AT   ataxia telangiectasia
ATM  ataxia telangiectasia mutated
ATLD ataxia telangiectasia-like disorder
BARD1 BRCA1-associated ring domain 1
BASC BRCA1-associated genome surveillance complex
BLM  Bloom syndrome gene
BRCA BRCA1 (breast cancer 1) or BRCA2 (breast cancer 2)
BRCT BRCA1 carboxy-terminal repeat
BRIP1 BRCA1 interacting protein 1
c.*… nomenclature for the 3'UTR of a gene according to the Human Genome Variation Society
CDH1 cadherin 1
CHEK2 checkpoint kinase 2
CHK1 checkpoint kinase 1
chr chromosome
Cl  confidence interval
CNV copy number variant
CSGE conformation-sensitive gel electrophoresis
DEB diepoxybutane
ER  estrogen receptor
FA  Fanconi anemia
FC familial cancer cases
GTP guanosine 5'-triphosphate
GWAS genome-wide SNP association study
HNPCC hereditary non-polyposis colorectal cancer
ICL interstrand cross-link
kb kilobase
KBCP Kuopio Breast Cancer Project
LOH loss of heterozygosity
MERIT40 mediator of RAP80 interactions and targeting 40
mir microRNA
mTOR mechanistic target of rapamycin
MLH1 mutL protein homolog 1
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLPA</td>
<td>multiplex ligation-dependent probe amplification</td>
</tr>
<tr>
<td>MMC</td>
<td>mitomycin C</td>
</tr>
<tr>
<td>MMR</td>
<td>mismatch repair</td>
</tr>
<tr>
<td>MPN</td>
<td>Mpr1p, Pad1p N-terminal domain</td>
</tr>
<tr>
<td>MRE11</td>
<td>meiotic recombination 11</td>
</tr>
<tr>
<td>MRN</td>
<td>MRE11-RAD50-NBS1 complex</td>
</tr>
<tr>
<td>MSH2</td>
<td>mutS protein homolog 2</td>
</tr>
<tr>
<td>MSI</td>
<td>microsatellite instability</td>
</tr>
<tr>
<td>NBS</td>
<td>Nijmegen breakage syndrome</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>ND</td>
<td>not determined</td>
</tr>
<tr>
<td>NF1</td>
<td>neurofibromin 1 or neurofibromatosis type 1</td>
</tr>
<tr>
<td>NLS</td>
<td>nuclear localization signal</td>
</tr>
<tr>
<td>OR</td>
<td>odds ratio</td>
</tr>
<tr>
<td>P</td>
<td>probability value</td>
</tr>
<tr>
<td>PALB2</td>
<td>partner and localizer of BRCA2</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PHTS</td>
<td>PTEN hamartoma tumor syndrome</td>
</tr>
<tr>
<td>PTEN</td>
<td>phosphatase and tensin homolog</td>
</tr>
<tr>
<td>RAS</td>
<td>rat sarcoma viral oncogene</td>
</tr>
<tr>
<td>RB1</td>
<td>retinoblastoma 1</td>
</tr>
<tr>
<td>rs</td>
<td>reference SNP</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>STK11</td>
<td>serine/threonine kinase 11</td>
</tr>
<tr>
<td>TP53</td>
<td>tumor protein 53</td>
</tr>
<tr>
<td>TSG</td>
<td>tumor suppressor gene</td>
</tr>
<tr>
<td>USC</td>
<td>unselected cancer cases</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>V(D)J</td>
<td>recombination of Variable, Diversity and Joining gene segments</td>
</tr>
<tr>
<td>VEGF</td>
<td>vascular endothelial growth factor</td>
</tr>
<tr>
<td>WNT</td>
<td>wingless</td>
</tr>
<tr>
<td>WRN</td>
<td>Werner syndrome gene</td>
</tr>
<tr>
<td>XP</td>
<td>xeroderma pigmentosum</td>
</tr>
</tbody>
</table>
List of original publications

This thesis is based on the following articles, which are referred to in the text by their Roman numerals:


Contents

Abstract
Tiivistelmä
Acknowledgements 7
Abbreviations 9
List of original publications 11
Contents 13
1 Introduction 17
2 Review of the literature 19
  2.1 The malignant phenotype and inheritance patterns ........................................ 19
  2.2 Breast cancer ........................................................................................... 21
  2.3 The penetrance of breast cancer predisposition genes .............................. 22
  2.4 Moderate and high penetrance breast cancer susceptibility genes
     mutated in dominantly inherited cancer-prone disorders ...................... 24
     2.4.1 BRCA1 and BRCA2 ......................................................................... 26
     2.4.2 TP53 (Li-Fraumeni syndrome) ..................................................... 28
     2.4.3 STK11 (Peutz-Jeghers syndrome) ................................................. 28
     2.4.4 PTEN (PTEN Hamartoma Tumor Syndrome, autism
     spectrum disorders and developmental delay/mental
     retardation with macrocephaly) ......................................................... 29
     2.4.5 CDH1 (hereditary diffuse gastric cancer) ........................................ 30
     2.4.6 NF1 (neurofibromatosis type 1) ................................................... 30
  2.5 Moderate and high penetrance breast cancer susceptibility genes
     mutated in recessively inherited disorders ........................................... 31
     2.5.1 ATM (ataxia telangiectasia) .......................................................... 33
     2.5.2 NBS1 (Nijmegen breakage syndrome) .......................................... 33
     2.5.3 RAD50 (NBS-like disorder) ......................................................... 34
     2.5.4 BRCA2, PALB2, RAD51C, BRIP1 (Fanconi anemia) ................. 34
  2.6 Other important breast cancer susceptibility genes .............................. 39
     2.6.1 CHEK2 ......................................................................................... 39
     2.6.2 BARD1 .......................................................................................... 40
  2.7 Low penetrance breast cancer predisposition alleles ......................... 40
  2.8 Modifier genes ........................................................................................ 43
  2.9 Debated genes in breast cancer predisposition ..................................... 44
     2.9.1 Mismatch repair genes (MLH1, MSH2, MSH6, PMS1,
     PMS2) ................................................................................................. 44
2.9.2 Androgen receptor gene ................................................................. 45
2.10 The interaction of breast cancer predisposition and Fanconi anemia gene products in holoenzyme complexes .............................. 45
2.10.1 The Fanconi anemia pathway .................................................... 45
2.10.2 Further BRCA1-containing super-complexes and disease ...... 46
2.11 Candidate genes or novel alleles for breast cancer predisposition based on biological activity in the BRCA/FA signaling complex ...... 48
2.11.1 MERIT40 and ABRAXAS .......................................................... 49
2.11.2 BRIP1 and FANCA ................................................................. 50
2.11.3 CHK1 .................................................................................... 51

3 Aims of the study 53

4 Materials and methods 55
4.1 Cases and controls (I-IV) ................................................................. 55
4.2 DNA isolation (I-IV), RNA extraction and cDNA synthesis (II) ..... 56
4.3 Mutation detection methods .............................................................. 57
4.3.1 Conformation-sensitive gel electrophoresis (CSGE) (I, II, III) .................................................................................... 57
4.3.2 Multiplex Ligation-dependent Probe Amplification (MLPA) (III, IV) ........................................................................... 57
4.3.3 Sequencing (I-IV) .................................................................... 57
4.3.4 Breakpoint analysis (IV) ............................................................ 58
4.4 TagSNP analysis (II) ........................................................................ 58
4.5 Cell culture (II, IV) .......................................................................... 58
4.6 Cytogenetic analysis (IV) ................................................................. 59
4.7 Immunofluorescence (II) .................................................................. 59
4.8 Statistical and bioinformatical analysis (I-IV) ................................. 60
4.9 Ethical issues (I-IV) ......................................................................... 60

5 Results 61
5.1 Mutation screening of the MERIT40 gene (I) ................................. 61
5.2 Mutation screening of the ABRAXAS gene (II) ............................. 63
5.3 Screening for large genomic rearrangements of the BRIP1 and CHK1 genes (III) ............................................................................. 67
5.4 Screening for large genomic rearrangements of the FANCA gene (IV) ......................................................................................... 67

6 Discussion 69
6.1 Lack of mutations in the MERIT40 gene (I) ................................. 69
6.2  *ABRAXAS* is a breast cancer susceptibility gene in the Northern Finnish population (II) ................................................................. 70
6.3  Lack of large genomic rearrangements in the *BRIP1* and *CHK1* genes in Finnish breast cancer patients (III) ........................................ 72
6.4  Identification of the first breast cancer family with an extensive deletion in an upstream Fanconi anemia signaling pathway gene, *FANCA* (IV) ......................................................................................... 73

7  Concluding remarks 77
References 79
Original publications 115
1 Introduction

The impact of cancer has become similar to that of heart disease, being the leading causes of death in developed countries. Breast cancer is the most common cancer in women, affecting at least one out of ten, and to a much lesser extent men too. It arises due to genetic and environmental factors. 5–10% of all breast cancer cases are familial and expected to have a genetic background predisposing to the disease (Claus et al. 1996, Narod & Foulkes 2004, Parkin 2004). Most of the known breast cancer susceptibility genes participate in DNA repair, while others are recognized for their role in cellular adhesion, hormonal or metabolic processes. Thus, breast cancer is genetically and biochemically a heterogeneous disease.

The multifunctional and highly penetrant BRCA1 and BRCA2 are the two major breast cancer predisposition genes, accounting for about 20% of breast cancer families. A small portion of the cases are due to mutations in some other known high penetrance genes such as TP53 or STK11, but most of the remaining familial breast cancer occurrence is not yet assigned to defectively functioning genes. In fact, most of these cases might be explained by mutations in moderate and low penetrance genes together with environmental factors (Wooster & Weber 2003, Easton et al. 2007, Stratton & Rahman 2008). Identification of these genes is crucial for understanding the genetic etiology of breast cancer and for the design of effective therapies. Furthermore, by knowing the germline defect, diagnostics and genetic counseling can be provided for patients at risk.

It is estimated that most cancers contain up to hundreds or thousands of mutations (Loeb 2001, Bielas et al. 2006, Sjöblom et al. 2006, Stratton et al. 2009) and epigenetic defects (Costello et al. 2000, Toyota & Suyuki 2010). At some point during cancer evolution, the mutator phenotype arises due to chromosomal instability, DNA repair defect, epigenetic changes, aberrant mRNA translation, or constant cellular stress (Loeb 2001, Negrini et al. 2010, Guil & Esteller 2009, Ruggero & Pandolfi 2003, Galhardo et al. 2007, Johnstone & Baylin 2010). This results in an even more rapid mutation rate and mono- or heteroclonal evolution with clones having different abilities to metastasize or to respond to cancer therapies. As a result, in somatic cancers it is difficult – if not impossible – to delineate which locus was the first one to be mutated, even when analyzing preneoplastic lesions. In contrast, hereditary cancer susceptibility provides us with the unique possibility to address this question by tracking the predisposing germline mutation.
Many breast cancer predisposition genes are mutated in the germline of patients having complex hereditary syndromes, such as TP53 in Li-Fraumeni syndrome or BRCA2 in Fanconi anemia (FA). In addition, genes whose protein products interact with either of the BRCA proteins often predispose to breast cancer when mutated (Ghimenti et al. 2002, Seal et al. 2006, Erkko et al. 2007).

The aim of the current study was to identify novel breast cancer susceptibility genes or alleles, based on the interaction of their protein products with BRCA1 or by their participation in the Fanconi anemia signaling pathway. Five such genes were selected for a candidate gene approach-based mutation analysis: MERIT40 and ABRAXAS based on interaction of their protein products with BRCA1; BRIP1 (FANCJ) and FANCA as both BRCA1-interacting and FA factors; furthermore CHK1 based on its participation in both FA and BRCA signaling networks. The coding region of the MERIT40 and ABRAXAS genes was examined for conventional small-sized mutations (detectable by PCR-based pre-screening methods), and the exonic regions of BRIP1, FANCA, and CHK1 were screened for large genomic rearrangements in Northern Finnish familial breast cancer patients and controls. The results were further evaluated by statistical, in silico and functional analysis.
2 Review of the literature

2.1 The malignant phenotype and inheritance patterns

Cancer is the uncontrolled growth of cells. The acquired hallmarks of cancer as defined by Hanahan and Weinberg (2000) are evasion of apoptosis, self-sufficiency in growth signals, sensitivity to anti-growth signals, limitless replicative potential, tissue invasion and metastasis, and sustained angiogenesis. Other similarly important qualities of malignant neoplasms are genetic instability, metabolic deregulation, inflammation triggered by interaction of the tumor and the surrounding stroma, and avoidance of immune destruction (Hanahan, http://www.ncr.org.uk/ncriconference/2010abstracts/abstracts/Plen7.htm). Not all of these qualities might be displayed by each neoplasm at a given time point during its evolution, and additional hallmarks of cancer might emerge in the future. These features make cancer the most challenging disease to treat.

Cancer can be of somatic origin, arising from a cell that has no potential to pass on the genetic material from parent to offspring, giving rise to what is often called sporadic cancer cases. In some other instances the origin of malignant potential can be traced back to the germline, so that each cell of the organism harbors the same defect. If such an organism survives till adulthood and the abnormality is not incompatible with fertility, it can be inherited, giving rise to familial cancer syndromes. So far genetic and rare epigenetic defects (Nicholls 1993, Viljoen & Ramesar 1992, Suter et al. 2004, Hitchins et al. 2010) have been described as the source of hereditary disease predisposition.

Most genes playing a key role in cancer can be grouped into two categories: proto-oncopogenes and tumor suppressors. Proto-oncogenes are involved in the positive regulation of cell growth, while tumor suppressors negatively regulate this process and are able to suppress the malignant phenotype. Intriguingly, proto-oncogenes are known to be converted into activated oncopogenes only in a small fraction of familial cancer syndromes (Vogelstein & Kinzler 2004), as most of the known hereditary cancers are initiated by inactivated tumor suppressor genes (TSGs). TSGs often act recessively at the phenotypic level of the cells, as conventionally both of their alleles need to be inactivated to lose their regulatory function. Knudson’s two-hit hypothesis is the classical model of tumorigenesis initiated by mutated tumor suppressors (Knudson 1971). This model – based on the inheritance pattern of the RB1 gene in retinoblastoma – gives an explanation
as to why germline mutation carriers have an earlier disease onset and often present with bilateral tumors and primary cancers at multiple sites, compared to somatic mutation carriers that usually display unilateral tumors. A germline carrier already has a mutated allele in all of his cells as a first hit, therefore only a second hit in the other allele within any cell is needed for tumorigenesis. In contrast, patients without a germline mutation must acquire yet a first somatic hit, and then an additional hit is required in the other allele of the same gene within the same cell. In this classical concept, loss of heterozygosity (LOH) represents the second hit.

However, exceptions from this model are now known to occur. For instance, haploinsufficient and dominant-negative mutations might require only a single hit in TSGs for cancer predisposition, thus LOH might not occur (Nicolaides et al. 1998, Tucker & Friedman 2002, Fodde & Smits 2002, Heikkinen et al. 2006, Birch et al. 1998, Chenevix-Trench et al. 2002). Haploinsufficiency is the inability of the remaining wild type allele to produce sufficient amount of product to maintain the phenotype encoded by both alleles, whereas in dominant negative cases the mutated allele produces a product that interferes with the functioning of the wild type one. However, even in the classical two-hit cases, when deletion of the remaining wild type allele by LOH does happen, the necessity of the second hit occurring right after the first one, preceding other mutations, has not yet been demonstrated. It might well be that the second hit occurs at another genetic locus (maybe at a proto-oncogene) and LOH in the first inactivated TSG occurs later. In contrast to this scenario, some TSGs have more than two inactivating hits. For example, the APC gene in familial adenomatous polyposis is sometimes inactivated in two steps, followed by loss of the germline mutant allele as the third hit (Miyaki et al. 1994, Spirio et al. 1998, Lamlum et al. 1999). However, whether all three hits are required for tumorigenesis, or they just reflect genomic instability, is not currently known.

Recessively inherited cancer-prone disorders are another exception from Knudson’s two-hit hypothesis due to biallelic germline defects. Patients in this case are already born with two TSG-inactivating hits. They usually have severe and early onset disease with congenital defects. Fanconi anemia (FA), ataxia telangiectasia (presented later in chapter 2.5), and xeroderma pigmentosum (XP) are such examples (Cleaver 2000). The risk for skin cancer in XP depends on environmental exposure, as skin lesions might not be formed until UV light acts as a third hit/carcinogen to fully initiate malignant transformation (Daya-Grosjean & Sarasin 2005).
A further deviation from the Knudson model is that some mutated TSGs increase the risk for different disorders, depending on the number of mutated alleles in the germline. Examples include the \textit{ATM} gene and some of the FA genes: heterozygous carriers are at risk for developing breast cancer, while individuals with biallelic mutations have ataxia telangiectasia or FA, respectively (reviewed later in chapter 2.5). This phenomenon sheds light on the importance of gene dosage effects together with tissue specific regulation. Indeed, conditional tissue-specific disruption of \textit{Pten} leads to different tumors in the affected tissues of mice (reviewed by Yin & Shen 2008). Even a subtle germline \textit{Pten} defect lowering the dose of the protein product results in breast cancer predisposition in these animals, and further \textit{Pten} dose reductions culminate in dramatic tissue-specific cancer phenotypes, leading to the proposition of the continuum model for cancer development (Alimonti \textit{et al.} 2010). This model takes us one step closer in understanding one of the biggest enigmas in cancer research that germline mutations in ubiquitously expressed genes manifest disease only in certain organs.

All in all, ultimate inheritance patterns can still in many instances not be explained without precise knowledge of the mutated genes’ biochemical functions and spatiotemporal gene regulation. In addition, inheritance patterns might be complicated by other factors, such as incomplete penetrance, modifier genes, environmental interactions, and mosaicism (Strachan & Read 1996).

### 2.2 Breast cancer

Breast cancer affects about 10\% of the female Western population, being the most common cancer in women (Stratton & Rahman 2008). The disease can also occur in males, but this is a rare condition. Breast cancer is the uncontrolled growth of the cells of the breast, mostly arising from the epithelium of the mammary gland. During the malignant transformation process, the normal breast epithelium first transforms into benign atypical hyperplasia, then into carcinoma \textit{in situ}. Finally, invasive breast cancer arises, which is mostly ductal carcinoma originating from the inner lining of milk ducts, or lobular carcinoma originating from the lobules that supply the ducts with milk. Rare histological subtypes include medullary, inflammatory, comedo, mucinous, papillary, or tubular carcinomas (Li \textit{et al.} 2005). Mortality occurs due to metastasis to vital organs, such as bone, liver, lungs, and brain (Kang 2006).

The most important risk factors for breast cancer are age and gender, breast cancer being 100 times more common among women than men due to the
growth-promoting effects of estrogen and progesterone. Prolonged oral contraceptive use or estrogen replacement therapy thus further increase breast cancer risk. Other important risk determinants are family history of the disease, geographical location, exposure to ionizing radiation, diet, alcohol usage and physical activity (McPherson et al. 2000).

Hereditary predisposition accounts for 5–10% of all breast cancer cases (Claus et al. 1996, Narod & Foulkes 2004, Parkin 2004). It is recognizable by the familial aggregation of the disease and by the frequent presence of ovarian cancer, manifesting as familial breast-ovarian cancer. Further signs of hereditary breast cancer and de novo cases due to germline mutation are early disease onset, bilateral or multiple primary tumors in the same individual, or the presence of male breast cancer (Thull & Vogel 2004).

2.3 The penetrance of breast cancer predisposition genes

Genes with defective alleles predisposing to breast cancer can be grouped according to disease penetrance, i.e. the proportion of individuals with a mutated gene who exhibit clinical symptoms. In Figure 1, breast cancer predisposition genes are grouped in this way. According to the classification proposed by Stratton and Rahman (2008), high penetrance genes confer a 10- to 20-fold relative breast cancer risk, moderate penetrance genes generally increase the risk 2- to 4-fold, while low penetrance alleles usually increase the risk by less than 1.25-fold. The latter are typically due to single nucleotide polymorphisms (SNPs), which are common in the general population.

The primary strategy for the identification of high and moderate penetrance cancer genes has been genome-wide linkage and positional cloning, candidate gene approach, epidemiological studies, and next generation sequencing. Low penetrance cancer genes have recently been identified by genome-wide SNP association studies (GWAS), using large case and control series (reviewed by Stratton & Rahman 2008). The frequency of mutations in highly and moderately penetrant breast cancer predisposing genes is low compared to the common low penetrance breast cancer susceptibility gene variants, where the population frequency can be as high as 50% (Stratton & Rahman 2008). But proving the cancer predisposing potential of alleles that are not highly penetrant is difficult, as already moderately penetrant breast cancer susceptibility alleles fail to segregate with the disease, and pedigrees are often further compounded by additional, mutation negative breast cancer cases that arise sporadically anyway.
Fig. 1. Penetrance of breast cancer susceptibility genes. Y axis: relative risk of the indicated genes (that of the normal population is set to one). Risk estimates are according to Rahman & Stratton 2008, Erkko et al. 2008, Southey et al. 2010, Easton et al. 2007, Cox et al. 2007, Wirtenberger et al. 2006, Ding et al. 2007, Ding et al. 2009, and Hu et al. 2008. Some genes may have alleles that confer lower or higher penetrance to breast cancer than depicted.

*BRCA1* and *BRCA2* (breast cancer 1 and 2) are the two major breast cancer predisposing genes. A few percentages of hereditary breast cancer cases can be explained by mutations in other known, high and moderate penetrance genes, which are interestingly also often mutated in hereditary cancer-prone disorders (chapters 2.4 - 2.5). This suggests that although germline mutations in some genes particularly predispose to breast cancer, so far hardly any genes are known that predispose solely to this tumor type. This might reflect the ease with which the breast can be malignantly transformed, perhaps due to the genotoxic and mitogenic effects of estrogen. In agreement with this notion, breast cancer is a common disease.

A further prominent feature of most moderate and high penetrance breast cancer susceptibility genes is the over-representation of neurological, immunological and skin pigmentation defects in a syndromic representation. Examples of such disorders are summarized in Table 1 and 2 according to their dominant or recessive inheritance patterns, respectively. All of these disorders are
manifested by inactivated TSGs. These TSGs fulfill a wide range of cellular functions, but exactly how they predispose to breast cancer is not known.

2.4 Moderate and high penetrance breast cancer susceptibility genes mutated in dominantly inherited cancer-prone disorders

Heterozygous germline mutations of the high penetrance BRCA1, BRCA2, TP53, STK11, PTEN, and the moderate penetrance CDH1 and NF1 breast cancer susceptibility genes are notable for predisposing to multi-system cancer syndromes (Table 1). In fact, germline mutations of TP53 (Rapakko et al. 2001), STK11 (Chen & Lindblom 2000, Guenard et al. 2010), PTEN (FitzGerald et al. 1998), and CDH1 (Masciari et al. 2007, Schrader et al. 2011) are very rarely seen in breast cancer patients without manifestations of other features of the syndromes, and NF1 mutations have not even been reported outside neurofibromatosis type 1.
<table>
<thead>
<tr>
<th>Disease</th>
<th>Mutated gene</th>
<th>Main cancer types / neoplasms</th>
<th>Neurological phenotype</th>
<th>Skin pigmentation problems</th>
<th>Other phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hereditary breast and breast-ovarian cancer</td>
<td>BRCA1, BRCA2</td>
<td>breast, ovarian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li-Fraumeni syndrome</td>
<td>TP53</td>
<td>soft tissue and bone sarcomas, breast, brain, adrenocortical tumor, leukemia, tumors at multiple sites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peutz-Jeghers syndrome</td>
<td>STK11/LKB1</td>
<td>gastrointestinal hamartomatous polyposis, colorectal, gonadal, breast, ovarian, uterine, cervical, lung cancer</td>
<td></td>
<td></td>
<td>hyperpigmentation of lips, buccal mucosa, fingers and toes</td>
</tr>
<tr>
<td>Bannayan-Riley-Ruvalcaba syndrome</td>
<td>PTEN</td>
<td>hamartomatous intestinal polyposis, lymphomas, thyroid, breast, lipoma</td>
<td>macrocephaly, mental retardation</td>
<td>&quot;café-au-lait&quot; on penis</td>
<td>developmental delay</td>
</tr>
<tr>
<td>Cowden syndrome</td>
<td>PTEN</td>
<td>thyroid, breast, endometrial, genitourinary tumors, hamartomatous intestinal polyposis, lesions in multiple organs, lipoma, fibroma</td>
<td>macrocephaly, mental retardation</td>
<td>hamartomas, facial papules, trichilemmomas</td>
<td></td>
</tr>
<tr>
<td>Hereditary diffuse gastric cancer</td>
<td>CDH1 (E-cadherin)</td>
<td>diffuse gastric cancer, breast</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurofibromatosis type 1</td>
<td>NF1 (neurofibromin)</td>
<td>neurofibroma, breast, glioma, peripheral nerve sheath tumor, GIST, juvenile chronic myelogenous leukemia, rhabdomyosarcoma, phaeochromocytoma</td>
<td>macrocephaly, attention deficit hyperactivity disorder</td>
<td>&quot;café-au-lait&quot; spots, freckling, Lisch nodules</td>
<td>bone dysplasia, scoliosis, short stature</td>
</tr>
</tbody>
</table>
2.4.1 BRCA1 and BRCA2

The major breast cancer predisposing genes BRCA1 and BRCA2 confer high penetrance and clear dominant autosomal pattern of inheritance. According to estimates, mutations in these genes account for the majority of the familial breast-ovarian cancer cases (Ford et al. 1998), but for maximum 20% of familial breast cancer cases (Anglian Breast Cancer Study 2000). Inherited breast cancer has been linked to mutated BRCA1 in more cases than to BRCA2 (Ford et al. 1998). The average cumulative risk of female carriers for breast cancer by the age of 70 is 65–87% for BRCA1 (Ford et al. 1994, Antoniou et al. 2003, Brose et al. 2002) and 45–87% for BRCA2 (Antoniou et al. 2003, van der Kolk 2010). The estimates vary largely depending on the studied population, patient selection criteria, and mutation detection method. BRCA1-mutated tumors tend to be high grade, have high mitotic count, exhibit continuous pushing margins and lymphocytic infiltration, and the majority is of the infiltrating ductal or basal-type. They are often triple negative, regarding the lack of expression of the estrogen, progesterone and ERBB2/HER2 receptors, which is associated with a poorer prognosis. BRCA2 mutation-associated tumors appear to behave more similarly to that of non-carriers, and their pathological characteristics are less clear than for BRCA1-associated breast tumors (reviewed by Thompson & Easton 2004).

The BRCA genes are also highly penetrant regarding ovarian cancer (Antoniou et al. 2003) and predispose to a smaller extent to male breast cancer. BRCA1 confers a higher risk of ovarian cancer than BRCA2, and BRCA2 confers a higher risk of male breast cancer than BRCA1 (Szabo & King 1995, Tai et al. 2007). Inactivation of these genes increases the risk for other malignancies as well in varying degrees, such as stomach, pancreas, prostate, uterus, or colon cancer in BRCA1 mutation carriers (Ford et al. 1994, Thompson et al. 2002, Brose et al. 2002) and prostate, pancreas, gall bladder, bile duct, stomach cancer, and melanoma in BRCA2 mutation carriers (The Breast Cancer Linkage Consortium 1999). Mutations in the BRCA genes may predispose to hereditary cancers of the above tumor types outside familial breast cancer too, but such a link is not always obvious. For instance, it is not known whether BRCA2 is mutated in familial gall bladder cancer patients, but the association of the BRCA genes with familial pancreatic cancer has been well demonstrated (Lal et al. 2000, Hahn et al. 2003, Lynch et al. 2005, Couch et al. 2007, Kim et al. 2009).

The BRCA genes were both identified by genetic linkage (Hall et al. 1990, Wooster et al. 1994) and subsequent positional cloning studies (Miki et al. 1994,
Wooster et al. 1995, Tavtigian et al. 1996). BRCA1 is localized on chromosome region 17q21.1, BRCA2 on 13q12.3. Both genes encode large nuclear multifunctional proteins without significant homology either to each other or to any other protein. They are involved in diverse functions, such as various DNA repair pathways, cell cycle checkpoints, chromosome segregation, and chromatin remodeling. In addition, specialized functions have been reported for BRCA1 including E3 ubiquitin ligase activity, regulation of transcription, mitosis, and estrogen receptor (ER) signaling (reviewed by Scully & Livingston 2000, Venkitaraman 2002, O’Donovan & Livingston 2010). A specific hallmark of loss of function of either of the BRCA genes is loss of damage-inducible nuclear RAD51 foci. This provides a useful in vitro method to identify cell lines with loss of BRCA function (Turner et al. 2004).

Mutations in these two genes appear to be evenly distributed across the coding sequences, with no obvious hot-spots. Even though the mutations generally confer high breast cancer risk, BRCA alleles with milder effects might also exist as hypomorphic variants, and some alleles have currently unknown effects. What is more, contrary to the notion that most truncating mutations are highly pathogenic, a polymorphic stop codon in BRCA2 has been described (Mazoyer et al. 1996), which, however was found to be potentially deleterious in other cancer types, such as familial pancreatic cancer (Martin et al. 2005). Interestingly, although BRCA1 and BRCA2 show LOH in 21% and 33% of sporadic breast tumors, respectively, the frequency of somatic mutations in these genes is low in breast cancer, except for promoter hypermethylation, an alternative mechanism to inactivate BRCA1 (as shown and reviewed by Janatova et al. 2005).

The strongest predisposition to hormone-dependent cancers is remarkable in BRCA carriers, but its explanation has mostly remained elusive. Possible explanations are the association with tissue- and developmental stage-specific factors, such as ERα either with BRCA1 (Fan et al. 1999, Zheng et al. 2001, Kawai et al. 2002, Eakin et al. 2007) or with the coding region of its protein partner, BARD1 (Creekmore et al. 2007). As BRCA1 forms multi-protein complexes, in theory, mutations in any of the genes in this complex could influence hormonal pathways. Indeed, BRCA2 and p53 are also known to be regulated by estrogen (Hilakivi-Clarke 2000, Spillman & Bowcock 1996, Hurd et al. 1997), and these proteins are included in BRCA1 complexes (described later in chapter 2.10). Perhaps estrogen exposure in women who carry germline mutations in a breast cancer predisposition gene increases breast cancer risk,
because the ER signaling pathway would be able to cause DNA damage and genetic instability without being opposed by BRCA pathway-induced DNA repair. As an alternative explanation, reduced BRCA activity may be linked to a dedifferentiated state of mammary gland morphology, making malignant transformation easier than from terminally differentiated tissues (Hilakivi-Clarke 2000). The fact that hereditary breast and breast-ovarian cancer arise only in adults must be the result of hormone-dependent maturation of these organs, and thus the interplay of hormone signaling with breast cancer susceptibility gene products. This picture is further complicated by the discovery that biallelic BRCA2 germline mutation carriers have Fanconi anemia (FA-D1 complementation group, Howlett et al. 2002) (Table 2), a childhood disease with again a different spectrum of tumors than in monoallelic mutation carriers (chapter 2.5.4). Female hormone-mediated pathogenesis does not seem to be obvious in FA, as FA patients do not classically have breast or ovarian cancer. However, as the malignancies of FA patients usually arise in early childhood, and most of them die at a young age, it could provide a natural explanation for the lack of these post-puberty related malignancies.

2.4.2 TP53 (Li-Fraumeni syndrome)

Patients with Li-Fraumeni syndrome are predisposed to bone and soft tissue sarcomas, breast cancer, brain tumor, adrenal cortical carcinoma, leukemia, and an increased risk for multiple primary tumors. The syndrome is primarily caused by germline mutations in TP53 (Malkin et al. 1990, Srivastava et al. 1990) (chr. 17p13.1), a gene generally called the guardian of the genome. In addition, TP53 is among the most frequently mutated genes in human cancer (Sengupta & Harris 2005), including breast tumors (Sjöblom et al. 2006). The multifunctional p53 protein impacts on several pathways, such as those involving DNA repair, cell division, transcription, apoptosis, senescence, angiogenesis, metabolism, and tumor-stroma interactions (Sengupta & Harris 2005, Vousden & Ryan 2009, Teodoro et al. 2007, Bar et al. 2010). The risk of breast cancer of TP53 germline mutation carriers is 84% already by the age of 45 (Chompret et al. 2000).

2.4.3 STK11 (Peutz-Jeghers syndrome)

STK11/LKB1 (chr. 19p13.3, serine/threonine kinase 11) is mutated in Peutz-Jeghers syndrome (Hemminki et al. 1998) that is characterized by hamartomatous
polyposis, cancers in the gastrointestinal tract and multiple organs, as well as skin hyperpigmentation. The cumulative risk for breast cancer was estimated to be 45% by the age of 70 (Hearle et al. 2006). STK11 plays a role in cellular energy metabolism, cell polarization, cell growth, ER signaling, p53-dependent apoptosis, VEGF (vascular endothelial growth factor) regulation, WNT (wingless) and mTOR (mechanistic target of rapamycin) signal transduction (Nath-Sain et al. 2000, Fan et al. 2009).

2.4.4 PTEN (PTEN Hamartoma Tumor Syndrome, autism spectrum disorders and developmental delay/mental retardation with macrocephaly)

PTEN might be the second most commonly mutated gene in human cancers after TP53 (Simpson & Parsons 2001). The PTEN Hamartoma Tumor Syndrome (PHTS) comprises Cowden syndrome (Liaw et al. 1997), Bannayan-Riley-Ruvalcaba syndrome (Marsh et al. 1997), Proteus and Proteus-like syndrome (Zhou et al. 2000, Zhou et al. 2001). PTEN/MMAC1 (phosphatase and tensin homolog) is located on chr. 10q23.3. It encodes a phosphatase that negatively regulates the AKT signaling pathway and is involved in regulating cell size, chromosomal integrity, DNA repair, cell cycle, and apoptosis (Orloff & Eng 2008).

A unifying feature of PHTS in humans is hamartomatous polyposis and germline mutation in PTEN. The risk for breast cancer in Cowden syndrome is estimated to be 25–50% (Eng 2003), but is unclear in Bannayan-Riley-Ruvalcaba syndrome. Cowden syndrome is further characterized by mucocutaneous lesions, macrocephaly, and an increased risk for benign and malignant diseases of the thyroid and endometrium. Male Cowden syndrome patients with breast cancer have also been reported (Fackenthal et al. 2001). Other features of Bannayan-Riley-Ruvalcaba syndrome are macrocephaly, lipomatosis, hemangiommas, pigmented maculae of the penis, and developmental delay. Furthermore, Lhermitte-Duclos disease or dysplastic gangliocytoma of the cerebellum is another PHTS allelic disorder (Iida et al. 1998, Zhou et al. 2003). It may be associated with Cowden syndrome, and is characterized by global thickening of the cerebellar folia (presumably due to hamartomatous overgrowth of ganglion cells), ataxia, symptoms of increased intracranial pressure, and seizures.

Proteus syndrome consists of disproportionate, asymmetric overgrowth of body parts, cerebriform connective tissue nevi, epidermal nevi, vascular
malformations, and dysregulated adipose tissue. Germline PTEN mutations account for a subset of Proteus, Proteus-like and SOLAMEN (Segmental Overgrowth, Lipomatosis, Arteriovenous Malformation and Epidermal Nevus) syndrome cases (Zhou et al. 2000, Zhou et al. 2001, Smith et al. 2002, Loffeld et al. 2006, Caux et al. 2007). Asymmetric overgrowth of body parts and tissues has been explained by germline or sporadic genetic mosaicism (Zhou et al. 2000, Loffeld et al. 2006, Caux et al. 2007). Breast cancer has been reported in Proteus syndrome(-like) patients in at least two instances (Iqbal et al. 2006, Caux et al. 2007), while in others it might be absent due to premature death.

Interestingly, PTEN germline mutations have also been identified in autism spectrum disorders and in patients with developmental delay/mental retardation with macrocephaly (Goffin et al. 2001, Butler et al. 2005, Varga et al. 2009), thus being new allelic disorders to PHTS. It is not yet known whether these patients are at risk for breast cancer.

### 2.4.5 CDH1 (hereditary diffuse gastric cancer)

Hereditary diffuse gastric cancer is caused by germline mutations in CDH1 (cadherin 1) (Guilford et al. 1998). Mutation carriers have a lifetime risk of 39–52% for lobular breast cancer (Pharoah et al. 2001, Kaurah et al. 2007). CDH1 mutation-driven lobular breast carcinoma is one of the best examples of genotype-phenotype correlations in hereditary breast cancer. The gene is located on chr. 16q22.1 and encodes E-cadherin. It has a role in cell adhesion and hence in the maintenance of cell architecture and differentiation of epithelial cells. The cytoplasmic domain of E-cadherin directs the β-catenin-mediated interaction with the actin cytoskeleton and its downregulation directly leads to invasive potential (reviewed by Berx & Van Roy 2001).

### 2.4.6 NF1 (neurofibromatosis type 1)

Neurofibromatosis type 1 is characterized by café-au-lait spots, freckling, neurofibroma, iris Lisch nodules, optic nerve and other central nervous system gliomas, malignant peripheral nerve sheath tumors and scoliosis (abnormal curvature of the spine). It is caused by germline mutations in NF1 (neurofibromin 1, chr. 17q11.2) (Wallace et al. 1990, Cawthon et al. 1990, Viskochil et al. 1990). Women with neurofibromatosis type 1 aged less than 50 years have a five-fold increased risk of breast cancer (Sharif et al. 2007). NF1 codes for neurofibromin,
which is a protein accelerating GTP (guanosine 5′-triphosphate) hydrolysis on RAS (rat sarcoma viral oncogene homolog), thereby attenuating signaling from the RAS pathway (Basu et al. 1992).

2.5 Moderate and high penetrance breast cancer susceptibility genes mutated in recessively inherited disorders

Heterozygous mutation carriers of the high penetrance BRCA2, PALB2, RAD51C, and moderate penetrance BRIP1, ATM, NBS1, and RAD50 genes are all at increased risk for breast cancer. However, biallelic mutation carriers of the same genes succumb to entirely different, autosomal recessively inherited syndromes (Table 2). Biallelic mutations can be homozygous or compound heterozygous, and breast cancer is not typical in these patients, only in heterozygous family members.
Table 2. High and moderate penetrance breast cancer susceptibility genes conferring recessively inherited disorders.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Mutated gene</th>
<th>Main cancer types / neoplasms</th>
<th>Neurological phenotype</th>
<th>Skin pigmentation problems</th>
<th>Immune deficiency</th>
<th>Other phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ataxia telangiectasia (AT)</td>
<td>ATM</td>
<td>lymphoma, leukemia</td>
<td>cerebellar degeneration, apraxia</td>
<td>dermatitis, scleroderma-like changes, “café-au-lait” spots, eczema</td>
<td>+</td>
<td>ataxia, ocular telangiectasia, sterility, Cl, RS</td>
</tr>
<tr>
<td>Nijmegen breakage syndrome (NBS)</td>
<td>NBS1 (nibrin)</td>
<td>lymphoma, leukemia, medulloblastoma</td>
<td>microcephaly, structural brain abnormalities, mental retardation</td>
<td>“café-au-lait”-like spots, vitiligo spots, freckles</td>
<td>+</td>
<td>ataxia, growth retardation, bird-like face, Cl, RS</td>
</tr>
<tr>
<td>NBS-like disorder</td>
<td>RAD50</td>
<td>- (?)</td>
<td>microcephaly, mental retardation</td>
<td>hyper- and hypopigmentation</td>
<td>?</td>
<td>bird-like face, short stature, Cl, RS</td>
</tr>
<tr>
<td>Fanconi anemia (FA)</td>
<td>BRCA2 (FANCD1), PALB2 (FANCN), BRIP1 (FANCJ), RAD51C (FANCO)</td>
<td>AML, squamous cell carcinoma, childhood cancers</td>
<td>microcephaly, malformations, hearing loss</td>
<td>“café-au-lait” spots, hyper- or hypopigmentation</td>
<td>?</td>
<td>short stature, skeletal abnormalities, malformations of internal organs, bone marrow failure, developmental delay, Cl, RS</td>
</tr>
</tbody>
</table>

Abbreviations: Cl – chromosomal instability, RS – radiosensitivity
2.5.1 ATM (ataxia telangiectasia)

According to studies of individuals with ataxia telangiectasia (AT), it was suggested that female relatives heterozygous for ATM mutations have a two- to sevenfold increased breast cancer risk (Swift et al. 1976, Swift et al. 1987, reviewed by Thompson et al. 2005). Based on these observations, ATM (AT mutated, chr. 11q22.3) was established as a breast cancer susceptibility gene (Thompson et al. 2005, Renwick et al. 2006). The relative risk of breast cancer associated with ATM mutations has been estimated to be 2.4 (Renwick et al. 2006) or as high as 12- or 15-fold (Pylkäs et al. 2007, Chenevix-Trench et al. 2002).

Biallelic ATM mutations result in AT (Savitsky et al. 1995), characterized by cerebellar degeneration, telangiectasia (dilated blood vessels), neurodegeneration, greying of hair, immunodeficiency, chromosomal instability, predisposition to lymphoid cancers, and extreme radiation sensitivity. The protein encoded by ATM belongs to the PI3/PI4-kinase family and plays a prominent role in cell cycle checkpoint, apoptosis, V(D)J (Variable, Diversity and Joining gene segments) recombination, DNA repair and telomere maintenance, regulating a wide variety of downstream proteins, including p53, BRCA1, CHEK2, and NBS1 (Perkins et al. 2002, reviewed by Pandita 2002, Derheimer & Kastan 2010).

2.5.2 NBS1 (Nijmegen breakage syndrome)

Nibrin, the gene product of NBS1 (Nijmegen breakage syndrome 1, chr. 8q21), is involved in DNA double-strand break repair, V(D)J recombination, meiotic recombination, DNA damage-induced checkpoint activation, and telomere maintenance (Zhu et al. 2000, Helming et al. 2009, reviewed by Lamarche et al. 2010). Nijmegen breakage syndrome (NBS) is caused by biallelic NBS1 mutations, resulting in severe immunodeficiency, chromosomal and radiosensitivity, predisposition to malignancy, a bird-like facial appearance, microcephaly, short stature, and mental retardation (Varon et al. 1998).

Heterozygous NBS1 germline carriers have about a 3-fold increased risk for breast cancer (Gorski et al. 2003, Steffen et al. 2006a, Bogdanova et al. 2008), as well as a variety of other malignancies, such as prostate, ovarian, melanoma, leukemia, lymphoma, gastrointestinal, colorectal, and lung cancer (Cybulski et al. 2004a, Plisiecka-Halasa et al. 2002, Debiak et al. 2003, Varon et al. 2001, Ebi et al. 2007, Steffen et al. 2004, Steffen et al. 2006b).
2.5.3 RAD50 (NBS-like disorder)

The protein product of RAD50 (chr. 5q31) is important for non-homologous end-joining and V(D)J recombination, meiotic recombination, cell cycle checkpoint activation, and telomere maintenance (Helmink et al. 2009, reviewed by Lamarche et al. 2010). Heterozygous carriers are at risk for developing breast cancer (Heikkinen et al. 2003, Heikkinen et al. 2006, Tommiska et al. 2006). The RAD50 mutant allele 687delT has been reported to result in a 4.3-fold increased breast cancer risk (Heikkinen et al. 2006). In addition, RAD50 was also found to be mutated in hereditary pancreatic cancer by one study (Wang et al. 2008).

So far, only one patient has been described with biallelic RAD50 mutations in NBS-like disorder, thus a general list of clinical symptoms is not yet available. However, this patient presented with microcephaly, mental retardation, bird-like face, short stature, chromosomal instability, and radiosensitivity. When the follow-up was made, the patient had neither immunological defects nor cancer (Waltes et al. 2009).

2.5.4 BRCA2, PALB2, RAD51C, BRIP1 (Fanconi anemia)

Fanconi anemia (FA) is characterized by bone marrow failure, aplastic anemia, congenital defects, genomic instability, radiosensitivity and increased cancer incidence (Moldovan & D’Andrea 2009). Congenital defects include short stature, radial aplasia, microphthalmia, malformation of the kidneys, and skin hyperpigmentation. FA cells display increased chromosomal aberrations, hypersensitivity to DNA interstrand cross-linking (ICL) agents and subsequent accumulation of the cells at the G2/M cell cycle checkpoint. FA patients have an elevated risk for acute myelogenous leukemia (AML), head and neck squamous cell carcinoma, as well as hepatic, esophageal, and gynecologic malignancies (Alter 2003, Alter et al. 2003, Kutler et al. 2003, Rosenberg et al. 2003).

FA is a genetically heterogeneous disease, caused by biallelic mutations in at least 14 distinct genes: FANCA, FANCB, FANCC, FANCD1, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCJ, FANCL, FANCN, and FANCO (Figure 2). FA-B has X chromosome-linked inheritance. The most common complementation group is FA-A, which is attributed to the hypermutable nature of the FANCA gene (Wijker et al. 1999). All 14 FA complementation group proteins and the associated factors function in ICL DNA repair signaling, thus being sensitive to mitomycin C (MMC) or diepoxylutane (DEB), which form the
basis of diagnostic laboratory tests (German et al. 1987, Auerbach 1993, Mathew 2006). Thus, the biochemical homogeneity of the pathway is attributed to the participation of the proteins in DNA ICL repair and to their recruitment into multienzyme complexes (described in detail in chapter 2.10.1).

The FA pathway and cancer susceptibility in monoallelic mutation carriers

Hereditary breast cancer and FA are strongly intertwined genetically (Figure 2) as to date four FA genes acting downstream of the ubiquitination of FANCD2 and FANCI, namely BRCA2/FANCD1, PALB2/FANCN (partner and localizer of BRCA2, chr. 16p12.2), BRIP1/FANCJ/BACH1 (BRCA1 interacting protein 1, chr. 17q22.2) and RAD51C/FANCO (a homologous recombination factor, chr. 17q25.1) are known to be mutated in both diseases (Howlett et al. 2002, Xia et al. 2007, Reid et al. 2007, Erkko et al. 2007, Rahman et al. 2007, Cantor et al. 2004, Levitus et al. 2005, Levran et al. 2005a, Litman et al. 2005, Seal et al. 2006, Vaz et al. 2010, Meindl et al. 2010). Monoallelic RAD51C mutations have so far been identified in families with both breast and ovarian cancer (Meindl et al. 2010), and biallelic RAD51C mutations have been found in one family with a FA-like disorder, most probably representing the FA-O complementation group (Vaz et al. 2010). Biallelic BRCA1 mutations have not been identified, maybe because of embryonic lethality.
Fig. 2. Supercomplex formation of proteins involved in Fanconi anemia (FA) and breast cancer at DNA double-strand breaks. Proteins that are the products of genes with strong evidence of breast cancer susceptibility are striped. Proteins that are mutated exclusively in FA are in grey, and proteins associated with hereditary susceptibility to both syndromes are in black. FANCD2 and FANCI monoubiquitination represents the central part of the FA signaling pathway. Proteins above are upstream and proteins below in black are the downstream FA signaling components (further detailed in chapter 2.10). Upstream and central FA proteins are also involved in DNA repair by interacting with DNA or by localizing at DNA damage sites. Not all interactions are depicted, and the formation of some of the protein sub-complexes can be mutually exclusive.

In contrast to the requirement of biallelic inactivation for the manifestation of FA, monoallelic defects in the same genes are enough for breast cancer predisposition. Despite the moderate to high penetrance of the four FA/breast cancer predisposition genes, epidemiologic studies on cancer occurrence in heterozygous carrier relatives of FA patients, not stratified according to their complementation
group, have failed to show an overall elevated cancer occurrence. The only exception was a higher rate of breast cancer among carrier grandmothers (especially those with \textit{FANCC} mutations) and possibly prostate cancer (Mathew \textit{et al.} 2006, Berwick \textit{et al.} 2007, Tischkowitz \textit{et al.} 2008). However, breast cancer risk conferred by the shared BRCA/FA pathway genes is clearly elevated, and is about 2-fold for \textit{BRIP1} (Seal \textit{et al.} 2006) and 6- to 30-fold for \textit{PALB2} (Erkko \textit{et al.} 2008, Southey \textit{et al.} 2010). \textit{PALB2} alleles conferring a moderate 2.3-fold increased breast cancer risk have been also reported. These mutations were associated with incomplete penetrance for breast cancer in the investigated pedigrees (Rahman \textit{et al.} 2007). The breast cancer risk for \textit{RAD51C} is not yet known, but is presumed to be high based on complete segregation of the mutations with the disease and their absence in healthy controls (Meindl \textit{et al.} 2010). The risk conferred by \textit{BRCA2} mutations is described in chapter 2.4.1. Besides, mutated \textit{BRCA2} has been implicated to cause a fraction of Li-Fraumeni-like syndrome or familial breast cancer/sarcoma cases (Evans \textit{et al.} 2008, Manoukian \textit{et al.} 2007). Apart from predisposing to breast cancer, \textit{PALB2} has been found to be mutated in hereditary pancreatic cancer as well (Jones \textit{et al.} 2009, Tischkowitz \textit{et al.} 2009), whereas \textit{BRIP1} was shown to be involved in hereditary predisposition to prostate cancer (Kote-Jarai \textit{et al.} 2009).

Direct mutation analysis studies of FA genes acting upstream and central in the pathway have failed to show an elevated breast cancer risk for heterozygotes, but a missense variant was found in \textit{FANCA} and \textit{FANCE} in familial breast cancer (Garcia \textit{et al.} 2009, Seal \textit{et al.} 2003). An elevated risk for other cancer types is implied for some upstream FA gene alleles, as \textit{FANCC} and \textit{FANCG} mutations were found in inherited pancreatic cancer (Couch \textit{et al.} 2005, van der Heijden \textit{et al.} 2003). Furthermore, mutation screens of \textit{FANCA} in sporadic AML cases revealed a low frequency of both missense mutations (Condie \textit{et al.} 2002) and heterozygous exonic deletions (Tischkowitz \textit{et al.} 2004). However, remission samples or other tissues were not available to determine whether these were germline or somatic mutations deriving from leukemic cells in the blast phase (Mathew 2006).

\textit{Distinct Fanconi anemia phenotypes}

Some cancer patients might be undiagnosed FA cases without the classical symptoms, assigned to FA only upon very serious response to radiation- or chemotherapy (Gyger \textit{et al.} 1989, Mathew 2006). Such an example is provided by
an AML case with homozygous FANCD2 mutations (Borriello et al. 2007). One FANCG mutation together with an unclassified variant was also identified in a sporadic AML (Meyer et al. 2006), but it was not determined whether the two alterations resided on separate chromosomes, and thus if the patient was in fact an undiagnosed FA case.

In particular, BRCA2 and PALB2 biallelic mutations predispose to a distinct set of malignancies compared to the other FA genes, involving very early onset medulloblastoma, glioblastoma, Wilms tumor, and AML. The cumulative probability of malignancies was found to be as high as 97% by the age of 5.2 years in biallelic BRCA2 mutation carriers (Alter et al. 2007). Currently, these two complementation groups represent the only evident genotype-phenotype correlation in FA, as mutations in all the other FA genes – irrespective of the mutation site – cause similar defects at the cellular and clinical level (Neveling et al. 2009). This genotype-phenotype correlation might be explained by unusually high spontaneous and ICL-induced chromosomal instability, possibly reflecting a more severe defect in homologous recombination (Hirsch et al. 2004, Wagner et al. 2004, Reid et al. 2007). As an alternative point of view, it was proposed that BRCA2 mutations in FA-D1 children with brain tumors, primarily medulloblastomas, might constitute a new syndromic association (Offit et al. 2003), a phenotype that is now known to be shared with the FA-N complementation group (Xia et al. 2007, Reid et al. 2007). What is more, Wilms tumor and childhood brain tumors are also reported in biallelic BRCA2 mutation carriers who do not display evident FA features (Reid et al. 2005, DeWire et al. 2009). This potentially represents an even more distinctive variant of FA, alternatively a separate disease. In agreement with a potentially new disease entity, biallelic BRCA2 mutations have been associated with familial Wilms tumor (Reid et al. 2005).

Breast cancer is interestingly not a cancer type associated with the phenotype of FA, though at least four FA patients have been reported to manifest it, as reviewed by Alter et al (2003) and three patients were also pointed out by Kutler et al (2003). These patients’ median age at diagnosis was calculated to be 37 (Alter et al. 2003). Another FA-A patient was also diagnosed with breast cancer at the age of 37 (Huck et al. 2006). The young age of breast cancer onset indicates genetic predisposition in these patients. The lack of breast cancer among most FA patients might be due to early lethality or hypogonadism with decreased estrogen levels in FA women (D’Andrea, 2010). However, FA patients with milder symptoms due to hypomorphic mutations (like the patient reported by Huck et al.)
or mosaicism have exceptional longevity (Neveling et al. 2009), thus in theory they could have an increased breast cancer risk. Mosaicism is frequent in FA due to compensating mutations or reversion of the germline mutation to wild type, and was shown to be correlated with improved symptom status and survival (Lo Ten Foe et al. 1997, Waisfisz et al. 1999, Gregory et al. 2001, Gross et al. 2002, Mankad et al. 2006). It may be that because of their mild symptoms, some FA patients are not diagnosed with the disease, thus their breast cancer incidence remains elusive.

2.6 Other important breast cancer susceptibility genes

**CHEK2** and **BARD1** are considered to be moderate penetrance breast cancer susceptibility genes without firm syndromic connections.

2.6.1 **CHEK2**

**CHEK2** (chr. 22q12.1) is the upstream serine/threonine kinase of the highly penetrant breast cancer predisposition factors BRCA1 and p53 (Lee et al. 2000, Chebab et al. 2000). It is activated in response to ionizing radiation through phosphorylation by ATM (Lee & Paull 2005), contributing to critical regulation of cell cycle checkpoints and cell survival (Lee et al. 2000, Chebab et al. 2000). **CHEK2** is considered to be a moderate penetrance breast cancer susceptibility gene. In particular, the 1100delC mutation has been implicated to confer a 2-fold risk for breast cancer in women and a 10-fold increase in men (Meijers-Heijboer et al. 2002, Vahteristo et al. 2002). A pathogenic large genomic rearrangement has been also described as the cause of breast cancer (Walsh et al. 2006).

**CHEK2** mutations also predispose to prostate cancer (Dong et al. 2003a, Seppälä et al. 2003, Cybulski et al. 2004b, Cybulski et al. 2006, Wu et al. 2006). Mutations in this gene have been reported to confer an increased risk for thyroid, colon, kidney, bladder, and HNPCC-related colorectal cancer too (Cybulski et al. 2004c, Zlowocka et al. 2008, Wasielewski et al. 2008). In addition, **CHEK2** had been implicated in Li-Fraumeni and Li-Fraumeni-like syndrome (Bell et al. 1999, Lee et al. 2001, Vahteristo et al. 2001). However, its contribution to these familial cancer syndromes has been questioned due to the relatively high frequency of the reported mutations in control cases and lack of the typical cancer spectrum characteristic to Li-Fraumeni(-like) syndrome in most **CHEK2** mutation positive families (Sodha et al. 2002, Evans et al. 2008, Ruijs et al. 2009).


2.6.2 BARD1

BARD1 (BRCA1-associated ring domain 1, chr. 2q34-q35) constitutes a ubiquitin ligase together with BRCA1, regulates BRCA1’s subcellular localization, participates in homology-directed repair of chromosome breaks, mRNA polyadenylation, and apoptosis (reviewed by Irminger-Finger & Jefford 2006). Germline mutations in BARD1 have been found in breast cancer families (Ghimenti et al. 2002, Ishitobi et al. 2003, Karppinen et al. 2004, Stacey et al. 2006, De Brakeleer et al. 2010). In particular, the common Cys557Ser variant was associated with a 2- to 4-fold increased breast cancer risk (Karppinen et al. 2004, Stacey et al. 2006), but was not enriched in male breast, ovarian, prostate, and colorectal cancer (Karppinen et al. 2006). BARD1 is the only known non-low penetrance breast cancer susceptibility gene to date that does not appear to be associated with other hereditary cancer types. However, as the reported BARD1 mutation positive families also harbor other cancer cases than breast or ovarian, it would not be entirely unexpected if its role in other hereditary cancer forms was also established (with the possible exception of the Cys557Ser variant as reported by Karppinen and coworkers, 2006). In agreement with this assumption, common variations in BARD1 have been reported to influence susceptibility to high-risk neuroblastoma (Capasso et al. 2009).

2.7 Low penetrance breast cancer predisposition alleles

Some genes whose protein product interact with such important breast cancer predisposition factors as BRCA1, BRCA2 or p53 and play a crucial role in genome integrity maintenance, have not been shown to contribute to breast cancer susceptibility in a moderate or highly penetrant manner. Instead, they might act as low penetrance breast cancer predisposition genes. Notable examples of such genes include RAD51, BLM (Ding et al. 2009), and WRN (Wirtenberger et al. 2006, Ding et al. 2007). Their association with the BRCA1 holoenzyme complex is described in detail in chapter 2.10.2. BLM is mutated in Bloom syndrome (Ellis et al. 1995) and WRN in Werner syndrome (Yu et al. 1996), which are autosomal recessively inherited accelerated aging disorders. The results of the above SNP studies are in agreement with the role of genetic variations in genome integrity genes in age-related breast carcinogenesis.

CASP8 (caspase-8, involved in apoptosis) was identified as a breast cancer susceptibility gene by the genotyping of known SNPs in cases and controls.
(MacPherson et al. 2004, Cox et al. 2007). The CASP8 D320H allele conferred a protective effect against breast cancer, indicating the importance of inherited variation in the apoptosis pathway in breast cancer susceptibility (MacPherson et al. 2004).

The most recent technology used for the ascertainment of common, already annotated variants’ association with low cancer risk is oligonucleotide microarray-based GWAS. Its advantage is that no prior knowledge about the gene, its location or function is needed. However, due to the small risk these alleles confer to breast cancer susceptibility, some of the recent GWAS results have not been confirmed by others. Firmly confirmed variants involve those in FGFR2 (fibroblast growth factor receptor 2, a tyrosine kinase receptor playing a role in growth signaling), TNRC9/TOX3 (a putative high mobility group gene with a tri-nucleotide repeat, implicated in breast cancer metastasis to bone), MAP3K1 (mitogen-activated protein kinase kinase kinase 1, involved in growth signaling), LSP1 (lymphocyte-specific protein 1, an F-actin binding protein), H19 (imprinted non-coding RNA), and chr. 8q24 (all described by Easton et al. 2007). The relative risks of breast cancer associated with carrying a single copy of each risk allele ranged from 1.07 to 1.26 (Easton et al. 2007, Stratton & Rahman 2008).

Unexpectedly, none of these SNPs resided in protein coding regions: they were intronic or only in linkage disequilibrium with the indicated genes. This leaves open the possibility that the SNPs are in reality linked with other genes or perhaps cryptic genetic elements. What is more, the variant in 8q24 is located in a gene desert. It is interesting that SNPs in 8q24 were also associated for instance with multiple independent prostate (Amundadottir et al. 2006, Gudmundsson et al. 2007, Haiman et al. 2007a, Yeager et al. 2007) and colon cancer loci (Haiman et al. 2007b, Tomlinson et al. 2007, Zanke et al. 2007). Clustering of these predisposition alleles is unlikely to be a coincidence and may therefore indicate a shared mechanism for cancer predisposition (Stratton & Rahman 2008). This might be the regulation of MYC, which is the nearest gene to these loci – although being tens to hundreds of kilobases away. Tissue-specific transcriptional enhancer activities through long-range interaction with MYC (Ahmadiyeh et al. 2010, Wasserman et al. 2010) have been implicated as causative mechanisms for breast cancer susceptibility.

The genome-wide association study of Easton et al. (2007) sheds light on the possible involvement of oncogenes in breast cancer predisposition. Intriguingly, one of the two reported oncogenes, MYC, directly induces the expression of the maternal H19 RNA allele (which were both reported in the same study of Easton.
et al. 2007) and the reciprocally regulated \textit{IGF2} gene, thus contributing to breast carcinogenesis (Barsyte-Lovejoy \textit{et al.} 2006). \textit{MYC} is a widely implicated oncogene in breast tumors (reviewed by Liao & Dickson 2000). What is more, epigenetic deregulation of \textit{H19} and \textit{IGF2} is implicated in Beckwith-Wiedemann syndrome, an overgrowth disorder associated with embryonal tumors, and in its genetically and clinically opposite counterpart, Silver-Russell syndrome, a growth retardation disorder (reviewed by Eggermann \textit{et al.} 2008). The second implicated oncogene, \textit{FGFR2} – which behaves as a context dependent oncogene – is mutated in primary breast cancer (Stephens \textit{et al.} 2005) and in congenital skeletal disorders (Crouzon, Jackson-Weiss, Apert, Pfeiffer, Beare-Stevenson and Saethre-Chotzen syndrome) (Katoh 2009).

The loci characterized by Easton \textit{et al.} (2007) are estimated to account for 3.6\% of the familial risk of breast cancer in European populations. They represent the justification of the common variant – common disease hypothesis. Numerous additional novel low-risk breast cancer variants are being reported, thus their real impact on breast cancer predisposition will probably be clarified in the near future. In particular, many of the novel loci were specifically associated with a distinct hormone receptor status. For example, a SNP near \textit{TNRC9} was found to confer susceptibility to ER-positive breast cancer (Stacey \textit{et al.} 2007), similarly to SNPs in \textit{FGFR2} and \textit{MRPS30/PDCD9} (mitochondrial ribosomal protein S30 / programmed cell death protein 9) (Stacey \textit{et al.} 2008). Notably, one of the disease-associated \textit{FGFR2} SNPs generated a putative ER binding site (Easton \textit{et al.} 2007). In contrast, SNPs in \textit{MERIT40} were associated with an increased risk of triple-negative breast cancer (Antoniou \textit{et al.} 2010) and ovarian cancer (Bolton \textit{et al.} 2010). \textit{MERIT40} is a BRCA1 interacting protein (Shao \textit{et al.} 2009a, Feng \textit{et al.} 2009, Wang \textit{et al.} 2009) and BRCA1 mutation positive breast tumors are usually triple-negative (Thompson & Easton 2004). Thus, the association of \textit{MERIT40} variants with triple-negative breast cancer and their modifier effect also on \textit{BRCA1} mutation carriers is in agreement with the important functional connection of the two proteins.

From the SNP association studies the important conclusions have emerged that simple dominant and recessive models could be rejected for the studied alleles, and there was a higher breast cancer risk for homozygous than for heterozygous carriers in a dose-dependent manner (Easton \textit{et al.} 2007, Stratton & Rahman 2008). Also, the involvement of many of these effected genes in cancer predisposition syndromes is reminiscent of that of moderate and high penetrance breast cancer predisposition genes.
It is notable that apart from the H19 non-protein coding RNA, also the role of other non-coding RNAs is being established in breast cancer susceptibility. SNPs in pre-microRNAs (pre-mir) 196a2 and 499 have been associated with a 17% and 19% increased risk of breast cancer in heterozygotes, respectively (Hu et al. 2009). Targets of mir-196a2 include homeobox genes, MYC, LSP1, and TNRC9, while those of mir-499 include NBS1, among others. A SNP in mir-146a has been found to influence breast and ovarian cancer onset, and the microRNA has been shown to modulate BRCA1 and BRCA2 expression (Shen et al. 2008). In addition, a SNP in mir-125a, whose targets include ERBB2 and ERBB3, has been demonstrated to be a germline mutation in six unrelated breast cancer cases (Duan et al. 2007, Li et al. 2009). Furthermore, a risk variant in the miRNA-125b binding site in the 3’UTR of BMPRIB (bone morphogenic receptor type 1B encoding a serine/threonine kinase) is associated with breast cancer (Saetrom et al. 2009).

It is expected that polymorphic DNA copy number variants (CNVs) have a similar impact on breast cancer predisposition as SNPs. Their role in hereditary predisposition to neuroblastoma is already established (Diskin et al. 2009). Currently, the only known polymorphic CNV in breast cancer significantly modifying cancer risk is an exonic MTUS1 (mitochondrial tumor suppressor gene 1) deletion, decreasing the risk by 42% (Frank et al. 2007).

2.8 Modifier genes

Beast cancer susceptibility allele products may interact with each other multiplicatively and modify the risk conferred by another allele (Antoniou et al. 2002). For example, one of the RAD51 SNPs was confirmed to be a significant modifier of breast cancer risk in BRCA2 mutation carriers (Levy-Lahad et al. 2001, Wang et al. 2001, Kadouri et al. 2004). In fact, one study has reported a potential disease-causing RAD51 missense mutation in two patients with bilateral breast cancer (Kato et al. 2000). Thus it seems that some RAD51 variants have a high impact on breast cancer susceptibility, while some other alleles have smaller effects (Ding et al. 2009).
2.9 Debated genes in breast cancer predisposition

2.9.1 Mismatch repair genes (MLH1, MSH2, MSH6, PMS1, PMS2)

Hereditary non-polyposis colorectal cancer (HNPCC or Lynch syndrome) is an autosomal dominantly inherited disorder that is associated with germline mutations in mismatch repair (MMR) genes. The most affected genes in HNPCC are MLH1 and MSH2 (Bronner et al. 1994, Fishel et al. 1993). A germline deficiency in any of them leads to a strong mutator phenotype with accumulation of replication errors, typically at short repetitive DNA sequences. Thus, the molecular hallmark of HNPCC is microsatellite instability (MSI). Individuals carrying germline mutations in MMR genes have a high lifetime risk of developing colorectal cancer as well as extracolonic tumors in the endometrium, renal pelvis, ureter, stomach, small intestine, bile ducts, brain, and ovary (Lynch et al. 2006). It is strongly debated whether breast cancer is an inherent component of HNPCC and if mutated MMR genes predispose to breast cancer, as studies on the statistical occurrence and MSI status of this tumor in HNPCC have constantly yielded contradictory results (studies in favor of involvement of breast cancer in HNPCC or of the predisposition of MMR genes to breast cancer: Risinger et al. 1996, Walsh et al. 1998, Boyd et al. 1999, Stone et al. 2001, Westenend et al. 2005, Shanley et al. 2009, Walsh et al. 2010, Bianchi et al. 2010, Jensen et al. 2010; contra: Anbazhagan et al. 1999, Jonsson et al. 1995, Vasen et al. 2001, Muller et al. 2002). Interestingly, a germline MLH1 epimutation (promoter hypermethylation) has been identified in a HNPCC patient with breast cancer (Suter et al. 2004).

Intriguingly, biallelically mutated MMR genes also predispose to neurofibromatosis type 1 (reviewed by Banidpalliam et al. 2005). These cases are now described to have constitutional mismatch repair-deficiency syndrome, characterized by neurofibromatosis type 1, early onset hematological malignancy, gastrointestinal neoplasia, central nervous system tumors, and café-au-lait spots (reviewed by Wimmer & Eitzler 2008). Importantly, such clinical features have not been detected in the parents of the reported individuals (Banidpalliam et al. 2005), suggesting that this disease is caused by MMR deficiency leading to a mutation in the NF1 gene (Wang et al. 1999, Alotaibi et al. 2008). Indeed, NF1 has been found to be a mutational target of MMR deficiency in cancer cell lines (Wang et al. 2003) and in the germline of patients, causing neurofibromatosis type 1 (Alotaibi et al. 2008). The connection of HNPCC and constitutional
mismatch repair-deficiency with neurofibromatosis type 1 is remarkable from the angle of breast cancer susceptibility, due to the established and debated involvement of *NF1* and the MMR genes in this process, respectively. At least one breast cancer case has been described in patients with biallelically mutated MMR genes (Hackman et al. 1997).

### 2.9.2 Androgen receptor gene

Androgen receptor (*AR*) gene mutations have been implicated in male breast cancer (Wooster et al. 1992, Lobaccaro et al. 1993a, Lobaccaro et al. 1993b), but have not been confirmed by other studies (Haraldsson et al. 1998, Syrjäkoski et al. 2003). It has been proposed that the mutated *AR* has acquired the ability to bind to estrogen response elements and therefore to activate estrogen-regulated genes. Alternatively, the role of *AR* mutations in male breast cancer may relate to elevated estrogen/androgen activity ratio (Lobaccaro et al. 1993).

### 2.10 The interaction of breast cancer predisposition and Fanconi anemia gene products in holoenzyme complexes

#### 2.10.1 The Fanconi anemia pathway

A central event in the FA pathway is the monoubiquitination of the FANCD2 and FANCI proteins upon DNA damage, which is mediated by the upstream FA core complex (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, and FANCM) acting as a nuclear E3 ubiquitin ligase (Figure 2). The monoubiquitinated FANCD2/FANCI heterodimer interacts with FA proteins acting downstream relative to this post-translational modification (Kee & D’Andrea 2010). Altogether four downstream FA proteins – FANCD1/BRCA2, FANCN/PALB2, FANCJ/BRIP1, and FANCO/RAD51C – are known to participate in homologous recombination together with BRCA1 (Garcia-Higuera et al. 2001) and RAD51 (Scully et al. 1997a), beyond their role in ICL repair. There is a direct connection of some of these genes’ protein products, as BRIP1 binds to BRCA1 (Cantor et al. 2001) and PALB2 bridges the interplay between BRCA1 and BRCA2 (Sy et al. 2009, Zhang et al. 2009a, Zhang et al. 2009b).

In addition to these core FA proteins, there are several FA pathway-associated proteins whose functions are critical to the pathway; however, mutations have not
been found in the corresponding genes in FA patients. These include Fanconi-associated protein 24 (FAAP24), FAAP100, FANCM-associated histone fold protein 1 (MHF1) and MHF2, which are all required for efficient FANCD2 monoubiquitination. Fan1 (Fanconi anemia-associated nuclease 1) is a novel downstream FA protein that will certainly be scrutinized for involvement in FA and hereditary breast cancer predisposition. The activated FA pathway is inactivated for example by the USP1/UAF1 deubiquitinating enzyme complex (reviewed by Kee & D’Andrea 2010).

2.10.2 Further BRCA1-containing super-complexes and disease

Both BRCA1 and BRCA2 interact with p53 and RAD51 (Sharan et al. 1997, Mizuta et al. 1997, Chai et al. 1999, Scully et al. 1997a, Marmorstein et al. 1998). It is not known whether p53 and RAD51 are able to link the BRCA proteins like PALB2 does. BRCA1 also associates with RNA polymerase II (Scully et al. 1997b), the SWI/SNF chromatin remodeling complex (Bochar et al. 2000) and CtiP (CtBP-interacting protein) (Yu et al. 1998), thus affecting the transcription of numerous genes. Newly identified BRCA1 binding partners are RAP80, ABRAXAS, MERIT40, BRCC36, and BRCC45 (reviewed in chapter 2.11.1). DNA damage response through this BRCA1 holoenzyme signaling is achieved in part through a (de)ubiquitination cascade that has emerged as a central event in DNA damage signaling (Dong et al. 2003b, Wang et al. 2007a, Greenberg 2008, Shao et al. 2009a, Shao et al. 2009b). Interestingly, a potentially pathogenic amino acid deletion (del81E) was recently reported in RAP80 in two breast cancer patients (Nikkilä et al. 2009). This mutation abrogated ubiquitin binding and DNA damage response signaling for RAP80 and other members of the BRCA1 complex.

Several other BRCA1-containing super-complexes have been reported. One of them is BASC, the BRCA1-associated genome surveillance complex (Wang et al. 2000) that includes BARD1, ATM, BLM, the MRN (MRE11-RAD50-NBS1) complex, MSH2, MSH6, MLH1, RFC (replication factor C), and HDAC1 (histone deacetylase 1). WRN is probably also a BASC component (Wang et al. 2000). WRN and BLM – similarly to BRIP1 – are helicases that interact with each other, with p53, and NBS1 (von Kobbe et al. 2002, Spillare et al. 1999, Blander et al. 1999, Garkavstev et al. 2001, Cheng et al. 2004). For the RAD51-mediated homologous recombination, RAD51’s association with BLM and WRN is also important (Wu et al. 2001, Otterlei et al. 2006).
In the MRN complex, RAD50, NBS1, and MRE11 (meiotic recombination 11) mutually stabilize each other (Stewart et al. 1999, Delia et al. 2004, Bartkova et al. 2008), which could explain the very similar phenotypes conferred by inactivation of any one of the three cooperating proteins. MRE11 also interacts with ATM, thus it is notable that patients with biallelically mutated MRE11 have AT-like disorder (ATLD) (Stewart et al. 1999). ATLD is a similar, but less severe disease than AT, as patients do not have telangiectasia, display a later onset of ataxia, slower progression of neurodegeneration, and are not predisposed to cancer (Stewart et al. 1999, Delia et al. 2004, Fernet et al. 2005). MRE11 encodes an endonuclease that is involved in homologous recombination, non-homologous end-joining, V(D)J recombination, and telomere length maintenance (Zhu et al. 2000, Helmink et al. 2009, reviewed by Lamarche et al. 2010). Two reports have been published on the involvement of MRE11 in breast cancer predisposition. A truncating mutation and another, most probably also pathogenic variant (Arg202Gly) was found in familial breast cancer cases (Bartkova et al. 2008). A further potentially disease-related variant (Arg305Trp) was also identified, though in an ovarian cancer patient (Heikkinen et al. 2003). Due to the limited number of studies carried out so far, the role of MRE11 in hereditary breast cancer predisposition has not yet been firmly established, but based on the functional similarities of all three MRN complex proteins and the phenotypic outcome of the mutations in these genes, MRE11 might be a moderate penetrance breast cancer susceptibility gene as well.

Besides MRE11, ATM also associates with BRCA1, p53, and STK11 upon DNA damage (Fernandes et al. 2005). In addition, ATM-dependent phosphorylation of FANCD2 converges the FA and AT signaling pathways (Taniguchi et al. 2002), which might explain the phenotypic overlap between FA and AT.

Association of the BASC complex with the MMR proteins might explain the debated involvement of MMR genes in breast cancer from an alternative direction, perhaps not via MSI, but through deregulated BRCA1 complex function or ICL repair activity. In agreement with this notion, the FANCJ-MLH1 interaction has been found to be required for correction of ICL sensitivity in FA-J cells (Peng et al. 2007). Thus, mutations in either FANCJ or MLH1 might further explain the presence of microsatellite stable cancers in HNPCC (Xie et al. 2010).

It is notable that the BRCA/FA network’s members are all bona fide tumor suppressors, which seem to be engaged in almost all major DNA repair pathways: the FA pathway in ICL and oxidative damage repair (reviewed by Neveling et al.
the BRCA complex together with RAD51 in homologous recombination; the MRN complex in non-homologous end-joining; the MSH2 and MLH1 proteins in MMR; and signaling kinases in signal transmission by phosphorylation of target proteins. Nucleotide excision repair factors and translesion synthesis have been also reported to be necessary for proper FA signaling (McCabe et al. 2008, Shen et al. 2009, Bhagwat et al. 2009, Niedzwiedz et al. 2004, Song et al. 2010).

2.11 Candidate genes or novel alleles for breast cancer predisposition based on biological activity in the BRCA/FA signaling complex

Given the numerous breast cancer predisposition genes that are part of a BRCA multiprotein complex and are involved in genome integrity maintenance (Figure 2), all related factors are good candidates for breast cancer susceptibility, until disproven. Prior knowledge of the involvement of a candidate gene in a recessively inherited genomic instability disorder might further raise the possibility of its role in cancer predisposition.

The candidate gene approach involves the systematic screening of entire genes for disease-causing variants in cancer cases compared to controls. It might not only be employed because of a priori knowledge on the gene’s biological activity, but this might be the only effective strategy to detect rare high or moderate penetrance cancer susceptibility loci. Disease-causing mutations in moderate penetrance breast cancer susceptibility genes do not generally result in large pedigrees with multiple breast cancer cases and their segregation with the disease is incomplete – thus, further such susceptibility genes are not easily mapped by genetic linkage analysis. Moreover, uncommon disease-causing alleles of both moderate and high penetrance cancer susceptibility genes are unlikely to be detected by association studies. The usage of familial rather than population-based breast cancer cases is expected to increase the mutation detection rate, as even lower-penetrance breast cancer susceptibility alleles are usually enriched in familial compared to non-familial series. Usage of population isolates with founder mutations can also empower gene identification studies (reviewed by Stratton & Rahman 2008). The Finns, originating from a small number of settlers, constitute an isolated population with genetic founder effects. Especially the Northern Finnish population has remained largely isolated. Therefore, it is particularly suitable for the identification of novel cancer
predisposition alleles by using this study approach (Heikkinen et al. 2006, Erkko et al. 2007, Pylkäs et al. 2007).

2.11.1 MERIT40 and ABRAXAS

**MERIT40** (mediator of RAP80 interactions and targeting 40 kD/HSPC42/NBA1/c19orf62, chr. 19p13.11) encodes a BRCA1 interacting protein (Shao et al. 2009a, Feng et al. 2009, Wang et al. 2009). Other BRCA1 interacting proteins in this complex are ABRAXAS (Wang et al. 2007a, Kim et al. 2007a, Liu et al. 2007), RAP80 (Kim et al. 2007b, Sobhian et al. 2007, Wang et al. 2007a), BRCC45 (Dong et al. 2003b, Rual et al. 2005, Ewing et al. 2007), and BRCC36 (Dong et al. 2003b, Chen et al. 2006). Interaction with MERIT40 is essential to maintain the integrity of the complex and for DNA double-strand break targeting via lysine-63-ubiquitin deubiquitination activity. Consequently, MERIT40 is required for G2 checkpoint execution and viability responses to ionizing radiation (Shao et al. 2009a, Feng et al. 2009, Wang et al. 2009). The protein possesses a von Willebrand factor A (VWA) domain at its N-terminus that is homologous to the VWA domain of the S5A proteasome subunit (Wang et al. 2009). To date, MERIT40 appears primarily to play a potential scaffolding role for the BRCA1 complex, although other activities cannot be ruled out, such as proteasome-linked protein processing activities through the VWA domain. The central region (amino acids 93-320) of MERIT40 is evolutionarily conserved down to invertebrates and plants. Deletion of the last 30 residues prevents MERIT40 nuclear accumulation, formation of ionizing radiation-induced foci, as well as interaction with BRCA1, RAP80, ABRAXAS, and BRCC36 (Shao et al. 2009a).

**ABRAXAS** (CCDC98/ABRA1/FAM175A, coiled coil domain containing 98, chr. 4q21.23) is proposed to be a central organizer of the MERIT40-RAP80-BRCC36-BRCC45-containing BRCA1 holoenzyme complex (Wang et al. 2007a, Kim et al. 2007a, Liu et al. 2007, Wang & Elledge 2007, Feng et al. 2009, Wang et al. 2009), binding via its phosphorylated SPTF (serine-proline-threonine-phenylalanine) motif to one of the BRCT domains of BRCA1 (Wang et al. 2007a, Kim et al. 2007a, Liu et al. 2007). Through these interactions ABRAXAS targets BRCA1 to sites of DNA damage, participates in the G2/M checkpoint and in homologous recombination. Deletion of the last 159 residues renders ABRAXAS cytoplasmic and prevents it from BRCA1 binding (Kim et al. 2007a). Regarding post-translational modifications, both ABRAXAS and BRCC36 possess an MPN
(Mpr1p, Pad1p N-terminal) domain (but that of ABRAXAS is catalytically inactive), and such a domain pair is observed in the proteasome lid complex and the COP9 signalosome. It is speculated that the ABRAXAS-BRCC36 MPN domain pair may play a role in the RAP80-mediated DNA damage-induced (de)ubiquitination of the BRCA1 complex (Wang et al. 2009). No ABRAXAS mutations have been reported prior to this study (Novak et al. 2009), except for a missense variant (Met299Ile) with a suspected deleterious effect (Osorio et al. 2009).

2.11.2 BRIP1 and FANCA

BRIP1/FANCJ/BACH1 is a BRCA1 Associated C-terminal Helicase and DNA-dependent ATPase, necessary for BRCA1’s tumor suppressive functions (Cantor et al. 2001, Cantor et al. 2004). BRIP1 is mutated in FA-J (Levitus et al. 2005, Levran et al. 2005, Litman et al. 2005) and in a fraction of familial breast cancer cases (Cantor et al. 2004, Sigurdson et al. 2004, Seal et al. 2006, De Nicolo et al. 2008). The C-terminal part of BRIP1 interacts with the BRCT repeats of BRCA1 (Yu et al. 2003), similarly to MERIT40 and ABRAXAS. Again, in a similar manner, BRIP1 supports the localization of BRCA1 to DNA damage foci (Cantor et al. 2001) and the G2/M checkpoint function (Rodriguez et al. 2003, Yu et al. 2003). Both BRIP1 and BRCA1 are critical for the repair of DNA double-strand breaks by homologous recombination, resistance against MMC-induced DNA damage, and for chromosomal integrity (Litman et al. 2005). BRIP1 is an established breast cancer susceptibility gene, but interestingly, no conventional BRIP1 mutations have been identified in Finnish breast cancer patients (Karppinen et al. 2003, Vahteristo et al. 2006).

FANCA (chr. 16q24.3) encodes an upstream FA pathway protein that is part of the FA nuclear core complex (reviewed by Moldovan et al. 2009). It also interacts with BRCA1 (Folias et al. 2002), with the chromatin remodeling BRG1 protein (Otsuki et al. 2001), as well as with CENP-E (centromere-associated protein E) (Du et al. 2009). FANCA is mutated in two-thirds of all FA cases. It has been described as being hypermutable due to the abundance of Alu elements, homopolymeric tracts and direct repeats (Fanconi anaemia/Breast cancer consortium 1996, Levran et al. 1997, Levran et al. 1998, Centra et al. 1998, Wijker et al. 1999, Levran et al. 2005b). Recombination between such elements has been ascribed for the high prevalence of FANCA deletions, accounting for one-third of all FANCA mutations (Morgan et al. 1999, Human Gene Mutation
Database, http://www.hgmd.cf.ac.uk/). Although, none of the upstream FA genes are acknowledged to be involved in breast cancer susceptibility, some data indicated such a possibility for FANCA. For instance, the prevalence of large genomic rearrangement in the interacting BRCA1 gene, also with high density of Alu sequences, and their contribution to breast cancer susceptibility has been demonstrated in several populations (Mazoyer et al. 2005). It is also notable that the 16q24.3 genomic region, where FANCA resides is a common target for LOH in breast tumors (Cleton-Jansen et al. 2001). In addition, an intronic FANCA SNP has been implicated to be associated with an 8% increase in breast cancer risk (Haiman et al. 2008), one potential FANCA missense mutation has been identified in UK breast cancer families (Seal et al. 2003), and large FANCA deletions have been reported in sporadic acute myeloid leukemia (Tischkowitz et al. 2004).

2.11.3 CHK1

CHK1 is an evolutionarily conserved serine/threonine kinase, the major downstream target of ATR phosphorylation, a critical maintainer of cell cycle checkpoints and genomic stability (Liu et al. 2000, Takai et al. 2000, Bartek & Lukas 2003). It is required for homologous recombination in concert with RAD51 (Sorensen et al. 2005), regulates chromosomal instability at fragile sites similarly to ATR, BRCA1, and FANCD2 (Durkin et al. 2006), and is involved in the regulation of mitotic catastrophe (Huang et al. 2005). There is emerging evidence that CHK1 is involved in BRCA1 (Yarden et al. 2002, Ting & Lee 2004), BRCA2 (Bahassi et al. 2008) and FA signaling pathways (Wang et al. 2007b, Collis et al. 2008, Zhi et al. 2009, Guervilly et al. 2008). Furthermore, CHK1 is directly involved in the signaling events of other high penetrance breast cancer susceptibility genes, such as TP53 (Shieh et al. 2000) and PTEN (Puc et al. 2005). CHK1 has been found to be mutated in sporadic cancers, such as colorectal (Bertoni et al. 1999, Kim et al. 2007c), stomach (Menoyo et al. 2001), endometrial cancer (Bertoni et al. 1999), melanoma (Papp et al. 2007), and mesothelioma (Kumar et al. 2005), and a potential germline mutation was found in lung cancer (Haruki et al. 2000). Only a few studies have been set out to search for possible germline CHK1 mutations in breast cancer families, but so far no pathogenic alterations have been found (Vahteristo et al. 2001, Marsh et al. 2007). Though LOH is frequently observed in chromosomal location 11q24, where CHK1 resides, no mutations or inactivating epigenetic changes could be unraveled in primary breast tumors (Allinen et al. 2002a).
3 Aims of the study

The vast majority of breast cancer predisposition genes remain to be discovered. In search for susceptibility genes that could explain an additional portion of familial breast cancer clustering in Finland, we set out to evaluate the presence of germline mutations in MERIT40, ABRAXAS, BRIP1, CHK1, and FANCA. The specific aims of the studies were:

1. To screen MERIT40 for conventional mutations, and evaluate its potential involvement in breast cancer susceptibility.
2. To screen the ABRAXAS gene for conventional mutations, and investigate the functional effect of a potential Finnish founder mutation.
3. To establish whether large genomic rearrangements exist in BRIP1 in Finnish breast cancer patients, given that small-size mutations were previously found to be absent from these patients. Furthermore, to assess for the first time, if CHK1 might be inactivated by large genomic rearrangements in the germline of cancer patients.
4. To determine whether large FANCA deletions might have an impact on breast cancer predisposition.
4 Materials and methods

4.1 Cases and controls (I-IV)

Familial breast cancer cases (I-IV)

Mutation screening was performed on 125 breast and breast-ovarian cancer families originating from Northern Finland. For statistical purposes, one index patient from each family was chosen according to the youngest age of breast cancer onset. Inclusion criteria and the number of families are summarized in Table 3. In studies III and IV, utilizing the multiplex ligation-dependent probe amplification (MLPA) method, the number of analyzed families was 111 and 100, respectively. Altogether 15 of the 125 studied index cases had previously been tested positive for known high penetrance breast cancer associated germline mutations in BRCA1 or BRCA2 (11 cases, Huusko et al. 1998), TP53 (one case, Huusko et al. 1999), and PALB2 (three cases, Erkko et al. 2007). In study II, an additional cohort of 139 familial breast cancer patients from the Tampere region in Southern Finland were used specifically for testing the presence of ABRAXAS c.1082G>A. In studies I and II, DNA from some additional affected family members was also screened for mutations from the Northern Finnish familial breast cancer cohort.

Table 3. Inclusion criteria of Northern Finnish breast cancer families in the mutation screening studies.

<table>
<thead>
<tr>
<th>Risk category</th>
<th>Study no. (and no. of families)</th>
<th>Inclusion criteria I</th>
<th>Inclusion criteria II</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>I (72)</td>
<td>3 or more cases of breast cancer, potentially in combination with single ovarian cancers in first- or second-degree relatives</td>
<td>2 cases of breast cancer in first- or second-degree relatives, of which at least one with early disease onset (&lt;35 years), bilateral breast cancer, or multiple primary tumors including breast or ovarian cancer in the same individual</td>
</tr>
<tr>
<td></td>
<td>II (73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>III (64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IV (57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>I (53)</td>
<td>2 cases of breast cancer in first- or second-degree relatives</td>
<td>1 case of breast cancer under the age of 35, or 1 breast cancer and other types of cancer in the same family (2 cases)</td>
</tr>
<tr>
<td></td>
<td>II (52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>III (47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IV (43)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Unselected breast cancer cases (II, IV)

For the ABRAXAS c.1082G>A genotyping and tagSNP analysis (study II) DNA from an unselected cohort of breast cancer patients (N=991) was analyzed. These breast cancer patients were not selected for or against family history of the disease. This sample set consisted of 544 Northern Finnish cases (same as for study IV) and 447 patients with invasive breast cancer from the Kuopio Breast Cancer Project (KBCP), originating from the province of Northern Savo in Eastern Finland and diagnosed at the Kuopio University Hospital between 1990 and 1995.

For genotyping the observed FANCA deletion allele (study IV), DNA from a cohort of breast cancer patients was collected without selection for family history of breast cancer. This sample set consisted of 540 Northern Finnish cases operated at the Oulu University Hospital during the years 2000–2007.

Healthy controls (I, II, IV)

The control samples (N=124–192 for study I and IV) derived from anonymous female cancer-free Finnish Red Cross blood donors (age ≥45 years) originating from Northern Finland.

Altogether 868 Finnish control cases were used for genotyping and tagSNP analysis for study II: the Northern Finnish control samples (where the number of studied individuals varied between 88 and 506, depending on the specific analysis) derived from anonymous female cancer-free Finnish Red Cross blood donors (age ≥45 years), and the age and area-of-residence matched KBCP cohort consisted of DNA from 362 female control subjects selected from the National Population Register during the same time period as the unselected breast cancer patients.

4.2 DNA isolation (I-IV), RNA extraction and cDNA synthesis (II)

DNA from blood lymphocytes was extracted using the standard phenol-chloroform method or the Puregene D-50K purification kit (Gentra). Tumor DNA isolated from soft tissue sarcoma was available from patient B92 in study III. DNA was extracted from paraffin-embedded tissue material with the QIAamp DNA FFPE Tissue kit and QIAcube (Qiagen).

To evaluate the effect of ABRAXAS c.1082G>A at the mRNA level, RNA was isolated using the QIAamp RNA Blood Mini kit (Qiagen) and was reverse
transcribed with iScript cDNA Synthesis Kit (Bio-Rad) according to the manufacturer’s instructions. The effect of the mutation at the mRNA level was then studied by cDNA sequencing.

4.3 Mutation detection methods

4.3.1 Conformation-sensitive gel electrophoresis (CSGE) (I, II, III)

Mutation screening of candidate genes was done by using CSGE on PCR products of exons and exon-intron boundaries. The assay is based on denaturating gel separation of amplicons with conformational changes due to mismatches in the double-stranded DNA.

4.3.2 Multiplex Ligation-dependent Probe Amplification (MLPA) (III, IV)

The SALSA MLPA kit P240-A1 BRIP1-CHEK1 and P031/P032 FANCA (MRC Holland) were used for the analysis of genomic deletions/amplifications according to the manufacturer’s instructions. After PCR amplification with IRD800 labeled primers the samples were analyzed with Li-Cor IR2 4200-S DNA Analysis system (Li-Cor Inc.) and Gene Profiler 4.05 analysis program (Scanalytics, Inc.).

Dosage data were interpreted using the P240-A1, P031 or P032 FANCA spreadsheets (National Genetics Reference Laboratory) in Microsoft Excel. Threshold values were set so that dosage quotients below 0.65 were considered to signify deletions and those above 1.35 amplifications. The standard deviation of the control ligation products did not exceed the 0.1 quality value for any of the investigated samples.

4.3.3 Sequencing (I-IV)

Sequencing was done for the following amplicons: those exhibiting deviating CSGE patterns, those that could not be analyzed by CSGE due to GC-richness or large size, and samples that were directly examined to unravel deletion breakpoint or to confirm mutation status of family members. Sequencing was performed with the Li-Cor IR2 4200-S DNA Analysis system (Li-Cor Inc.) using the SequiTherm
EXEL™II DNA Sequencing Kit-LC (Epicentre Technologies) or with ABI3730 (ABI Perkin Elmer) using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

4.3.4 Breakpoint analysis (IV)

The breakpoints of the FANCA deletion were confirmed with Affymetrix Genome-Wide Human SNP Array 6.0 platform and subsequent nested PCR amplification using the GeneAmp High Fidelity PCR System (Applied Biosystems). Allele-specific PCR was designed in a multiplex format with primers amplifying a control fragment to monitor PCR success and with FANCA primers specific to the site of the breakpoint. A successful PCR from DNA positive to the deletion yielded two amplicons, while deletion-negative samples yielded one PCR amplicon.

4.4 TagSNP analysis (II)

Genotyping of ABRAXAS c.1082G>A and tagging single nucleotide polymorphisms (tagSNPs) rs12499395, rs12649417 and rs13125836 was done using MassARRAY® mass spectrometer (Sequenom Inc., San Diego, CA, USA) and iPLEX® Gold (Sequenom Inc.) on 384-well plate format. TagSNPs were selected using the HapMap Genome Browser release 2. TagSNPs for region chr4:84593218-84633217 were picked out for the CEU population using the Tagger multimarker algorithm. MassARRAY was used for spectra acquisitions from the SpectroCHIP (Sequenom Inc.). Data analysis and genotype calling were done by using TyperAnalyzer Software version 4.0.3.18 (Sequenom Inc.). Each 384-well plate contained a minimum of eight non-template controls. For the c.1082G>A mutation DNA from three heterozygous mutation carriers was used as positive controls on each plate. For quality control, duplicate analysis was done for 6.5% of the samples from Oulu and for 6.7% of the samples from Kuopio.

4.5 Cell culture (II, IV)

Epstein-Barr virus transformed B-lymphoblastoid cell lines were established from mutation carriers and controls. Cells were grown in RPMI 1640 medium (Invitrogen) containing 20% FBS (Invitrogen), 1% L-glutamine (Invitrogen) and gentamycin (10 μg/ml, Roussel) at 37°C in a 5% CO2 atmosphere.
4.6 Cytogenetic analysis (IV)

The cytogenetic analysis of mitomycin C (MMC)-treated and untreated peripheral blood T-lymphocytes of two healthy *FANCA* deletion carriers and of non-carriers (N=4 for MMC-treatment and N=9 for untreated) was carried out in short-term three-day cultures. The four control samples for MMC-treatment derived from the *FANCA* deletion family and the nine controls used in standard chromosomal analysis were from unrelated healthy females. For the genotoxic treatment, MMC solution was added to the cultures at final concentration of 10 ng/ml. A minimum of 50 Giemsa-banded metaphases per sample were evaluated for chromosomal abnormalities under a light microscope and photographed with an automatic chromosome analyzer (CytoVision version 3.92, Applied Imaging). Samples were monitored for chromatid/chromosome breaks and deletions, simple chromosomal rearrangements (inversions, ring chromosomes, translocations, ≤3 break rearrangements), and complex chromosomal rearrangements (translocations, ≥4 break rearrangements and marker chromosomes) as published by Heikkinen *et al.* (2006).

4.7 Immunofluorescence (II)

U2OS or HeLa cells were cultured on glass coverslips and were used for transient transfections with LipoD293 (Signagen), while siRNA was transfected with Lipofectamine RNAiMax (Invitrogen) according to manufacturer’s protocols. Cells were treated with 10 Gy ionizing radiation and were allowed to recover for 4–6 hours in a 37°C incubator. Cells were washed with PBS, pre-extracted with 0.5% triton solution for 5 minutes at 4°C and then fixed in 3% paraformaldehyde/2% sucrose containing solution for 10 minutes at room temperature. Cells were subsequently permeabilized with 0.5% triton solution for 5 minutes at 4°C and then incubated with mouse monoclonal antibody HA.11 (Covance) at 1:000 dilutions for 20 minutes at 37°C to detect ABRAXAS or with rabbit polyclonal antibody against 53BP1 (Novus) at 1:250 dilution. Cells were then washed with PBST and incubated with secondary antibody for 20 minutes at 37°C. After four washes in PBST, coverslips were mounted onto glass slides using Vectashield mounting media containing DAPI (Vector Labs) and visualized using a Nikon Eclipse 80i fluorescent microscope. A minimum of 200 cells were counted per condition for quantification.
4.8 Statistical and bioinformatical analysis (I-IV)

Carrier frequencies between patients and healthy controls were compared by using Pearson Chi-Square or Fisher’s exact test (two-sided, SPSS version 17.0 for Windows). For ABRAXAS tagSNP data, the overall association as well as the Hardy-Weinberg equilibrium, allele-specific P, OR and CI were computed using Cochran-Armitage trend test. All alterations were checked with NNSplice software for potential effects on splicing (http://www.fruitfly.org/seq_tools/splice.html) and mutation candidates were also evaluated by using PolyPhen software (http://genetics.bwh.harvard.edu/pph). The vicinity of the FANCA deletion breakpoint was checked by the RepeatMasker program (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) for transposable elements.

4.9 Ethical issues (I-IV)

All of the biological specimens and clinical information of the familial breast cancer cases investigated were collected at the Oulu University Hospital, with the informed consent of the patients. Family members were only contacted with permission from the proband. Each patient-derived sample was labeled with a research code and handled anonymously. Results were only for research purposes, and no information on the outcome of the mutation analyses was given to patients or family members.

Approval to perform the study was obtained from the Ethical Board of the Northern Ostrobothnia Health Care District and the Finnish Ministry of Social Affairs and Health. The KBCP has been approved by the joint ethics committee of the University of Kuopio and the Kuopio University Hospital. The current research use of the Tampere breast cancer material has been approved by the Ethical Board of the Pirkanmaa Health Care District.
5 Results

5.1 Mutation screening of the MERIT40 gene (I)

Screening of the index patients of breast cancer families for germline alterations in MERIT40 revealed two intronic and five exonic variants. Four of the changes are not reported in the NCBI (National Center for Biotechnology Information) SNP database (http://www.ncbi.nlm.nih.gov/SNP/). Details and occurrence of the observed nucleotide variants are shown in Table 4. Only one alteration resulted in an amino acid change (Lys274Arg). However, according to the result of PolyPhen software analysis, it is not predicted to affect protein function. c.*87G>A, located in the 3’ UTR (untranslated region) of the gene was observed in one family but not in the controls, whereas all the other variants were found at similar or only at slightly lower frequencies in the control population.

The nucleotide alterations were also assessed for possible effects on consensus splice sites. Of the observed changes, c.*87G>A might be implicated in introducing a new splice acceptor site (score 0.68, versus 0.53 for the wild-type sequence), the effect of which is unclear.
Table 4. Observed sequence variations in the \textit{MERIT40} gene.

<table>
<thead>
<tr>
<th>Location</th>
<th>Nucleotide change</th>
<th>Effect on protein</th>
<th>rs number</th>
<th>Frequency of heterozygotes, % (n/N)</th>
<th>\textit{P} (OR; 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>familial cases</td>
<td>controls</td>
<td>familial cases</td>
<td>controls</td>
<td>familial cases</td>
<td>controls</td>
</tr>
<tr>
<td>exon 3</td>
<td>c.342C&gt;T</td>
<td>Asn114Asn</td>
<td>-</td>
<td>4.0 (5/125)</td>
<td>3.7 (7/192)</td>
</tr>
<tr>
<td>intron 3</td>
<td>c.344+41A&gt;T</td>
<td>-</td>
<td>rs10420922</td>
<td>49.6 (62/125)</td>
<td>44.3 (85/192)</td>
</tr>
<tr>
<td>exon 4</td>
<td>c.393C&gt;T</td>
<td>Phe131Phe</td>
<td>-</td>
<td>0.8 (1/125)</td>
<td>1.6 (3/192)</td>
</tr>
<tr>
<td>intron 8</td>
<td>c.787-6C&gt;T</td>
<td>-</td>
<td>rs10406920</td>
<td>39.2 (49/125)</td>
<td>37.5 (72/192)</td>
</tr>
<tr>
<td>exon 9</td>
<td>c.621A&gt;G</td>
<td>Lys207Arg</td>
<td>-</td>
<td>1.6 (2/125)</td>
<td>1 (2/192)</td>
</tr>
<tr>
<td>exon 9</td>
<td>c.637G&gt;A</td>
<td>Lys279Lys</td>
<td><strong>rs8170</strong></td>
<td>39.2 (49/125)</td>
<td>37.5 (72/192)</td>
</tr>
<tr>
<td>exon 9</td>
<td>c.&quot;87G&gt;A</td>
<td>-</td>
<td>-</td>
<td>0.8 (1/125)</td>
<td>- (0/192)</td>
</tr>
</tbody>
</table>

In bold: rs8170, the SNP associated with slightly increased breast (Antoniou \textit{et al.} 2010) and ovarian cancer risk (Bolton \textit{et al.} 2010)

Abbreviations: rs – reference SNP, \textit{P} – probability value, OR – odds ratio, CI – confidence interval
5.2 Mutation screening of the ABRAXAS gene (II)

The screening of 125 index patients of Northern Finnish breast cancer families for germline alterations in ABRAXAS revealed four intronic and six exonic variants (Table 5). Altogether five of the changes have not been reported either in the NCBI SNP database (http://www.ncbi.nlm.nih.gov/SNP/) or by previous studies (Osorio et al. 2009, Novak et al. 2009). Of the identified alterations c.1082G>A results in Arg361Gln, substituting the strongest positively charged amino acid with a non-charged one, and also changing the last residue of a putative nuclear localization signal (NLS) (Kim et al. 2007a). The site of the altered residue is located between the BRCA1-interacting SPTF motif and the RAP80-BRCC36 binding coiled coil domain, but the functional significance of this protein area is not characterized. PolyPhen software analysis classified Arg361Gln to be a possibly damaging alteration. In addition, ABRAXAS mRNA and protein sequence alignments both revealed absolute evolutionary conservation among vertebrates at the site of the wild type sequence, indicating functional significance.

We observed ABRAXAS c.1082G>A in 3 out of 125 Northern Finnish breast cancer families (2.4%) by direct sequencing of exon 9 (Table 5, Figure 3). The alteration was absent from 868 healthy control individuals, as assessed by iPLEX genotyping and sequencing. The exclusive occurrence of this variant among familial cancer cases prompted us to evaluate its prevalence in breast cancer patients unselected for a family history of breast cancer. The mutant allele was identified in one out of 991 unselected breast cancer cases. The prevalence of ABRAXAS c.1082G>A in familial vs. control cases, and also in familial vs. unselected breast cancer cases was found to be statistically significantly different (P=0.002 and P=0.005, respectively), and thus suggests that this variant is not only disease-associated, but also that its distribution is specifically correlated with familial breast cancer. In agreement with this notion is that the only mutation positive unselected breast cancer patient also proved to have a familial cancer background (Figure 3). This unselected breast cancer patient also originated from Northern Finland, similarly to the other three mutation positive breast cancer families. In order to determine if the mutation is characteristic to the whole Finnish population, a further familial breast cancer cohort from the Tampere region of Southern Finland was genotyped using iPLEX. However, the mutation was not present in the 139 breast cancer patients analyzed, indicative of being a Northern Finnish founder mutation.
Table 5. Observed variation in the *ABRAXAS* gene.

<table>
<thead>
<tr>
<th>Location</th>
<th>Nucleotide change</th>
<th>Effect on protein</th>
<th>rs number or reference</th>
<th>Frequency of heterozygotes, % (n/N)</th>
<th>P-value (OR; 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>familial cases</td>
<td>Unselected cases</td>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>intron 2</td>
<td>c.179-34_-38 delAATTA</td>
<td>-</td>
<td>0.8 (1/125)</td>
<td>ND</td>
<td>2.0 (2/100)</td>
</tr>
<tr>
<td>intron 3</td>
<td>c.216-44T&gt;C</td>
<td>Novak et al.</td>
<td>2.4 (3/125)</td>
<td>ND</td>
<td>5.3 (21/400)</td>
</tr>
<tr>
<td>intron 4</td>
<td>c.282+6T&gt;A</td>
<td>-</td>
<td>0.8 (1/125)</td>
<td>ND</td>
<td>- (-400)</td>
</tr>
<tr>
<td>intron 7</td>
<td>c.682-14A&gt;G</td>
<td>-</td>
<td>0.8 (1/125)</td>
<td>ND</td>
<td>0.5 (2/400)</td>
</tr>
<tr>
<td>exon 9</td>
<td>c.1042G&gt;A Ala348Thr</td>
<td>Osorio et al., Novak et al.</td>
<td>52.8 (66/125)</td>
<td>ND</td>
<td>51.1 (92/180)</td>
</tr>
<tr>
<td>exon 9</td>
<td>c.1082G&gt;A Arg361Gln</td>
<td>-</td>
<td>2.4 (3/125)</td>
<td>0.1 (1/991)</td>
<td>- (-867)</td>
</tr>
<tr>
<td>exon 9</td>
<td>rs12242536</td>
<td>Novak et al.</td>
<td>12.8 (16/125)</td>
<td>16.5 (163/990)</td>
<td>13.9 (121/868)</td>
</tr>
<tr>
<td>exon 9</td>
<td>c.*249delG</td>
<td>-</td>
<td>53.6 (67/125)</td>
<td>ND</td>
<td>47.7 (42/88)</td>
</tr>
<tr>
<td>exon 9</td>
<td>c.*347C&gt;T</td>
<td>-</td>
<td>0.8 (1/125)</td>
<td>ND</td>
<td>- (-88)</td>
</tr>
<tr>
<td>exon 9</td>
<td>rs6825184</td>
<td>Novak et al.</td>
<td>7.2 (9/125)</td>
<td>ND</td>
<td>5.7 (5/88)</td>
</tr>
</tbody>
</table>

Abbreviations: rs – reference SNP, P – probability value, OR – odds ratio, CI – confidence interval, ND – not determined, FC – familial cancer cases, CT – control cases, USC – unselected cancer cases
In addition to the mutation screening, a linkage disequilibrium-based analysis of the *ABRAXAS* gene was carried out, comparing the allele distribution of unselected breast cancer cases with those of age and geographically matched unaffected controls. The following tagSNPs within the gene were used: rs12499395, rs12649417 and rs13125836. There was no evidence of significant association of any of the analyzed tagSNPs with breast cancer risk.

As *ABRAXAS* c.1082G>A was the only alteration indicative of pathological significance, we set out to characterize its functional effect. First, we verified the presence of the mutant allele at the mRNA level to rule out aberrant splicing or expression. The mutation and its surrounding area was sequenced in reverse transcribed mRNA from cultured lymphoblastoid cells of two carriers, but no abnormalities were observed.

To study the subcellular localization of the wild type and Arg361Gln mutant *ABRAXAS*, we transfected the respective HA-tagged siRNA-resistant DNA constructs to U2OS and HeLa cells, in which endogenous *ABRAXAS* expression was silenced using siRNA construct CGUUUAGAGAGGGCCUUCACAA. Cells were exposed to 10 Gy ionizing radiation, and after 4 hours of recovery, they were fixed for immunofluorescence and incubated with anti-HA antibody. While wild type *ABRAXAS* showed nuclear localization, the mutant was found both in the nucleus and cytoplasm. An antibody staining for 53BP1, a p53 and BRCA1 interacting protein, was used as a marker for functional DNA damage checkpoint (Wang *et al.* 2002). In both wild type and mutant *ABRAXAS* transfected cells, 53BP1 was localized to nuclear foci indicating DNA damage sites. This result suggests that there is a functional DNA damage checkpoint in the cells. In triton pre-extracted cells (where soluble elements of the cells are removed), wild type ABRAXAS colocalized with 53BP1 to nuclear foci, but the Arg361Gln mutant did not, presumably as it was soluble in the nucleus due to loose association with the chromatin.
Fig. 3. Pedigrees of breast cancer families harboring the ABRAXAS c.1082G>A variant. Black circles represent breast cancer, other cancer types are marked with grey. Arrows point to index patients. The age of diagnosis is indicated below the patients or the current age for healthy carriers. Individuals are marked with a plus sign if mutation positive, or minus sign if negative. A slashed symbol indicates a deceased individual. Only limited information was available on family 98-063.

5.3 Screening for large genomic rearrangements of the BRIP1 and CHK1 genes (III)

111 index patients were screened for large-size germline genomic rearrangements in the BRIP1 and CHK1 genes by MLPA. No large deletions or amplifications were uncovered.

We have complemented this analysis by screening for the presence of a deep intronic cryptic splice site mutation (IVS11-498A>T) in BRIP1 (Levitus et al. 2005, Levran et al. 2005) that had been identified in a FA-J patient, but has not been previously assessed in breast cancer patients. Though this mutation was not detected, we identified another change, IVS11del-435_-439TTCT in one breast cancer patient. This variant is not reported in public databases - however, it has no predictable effect on splicing.

5.4 Screening for large genomic rearrangements of the FANCA gene (IV)

100 patients were screened for large-size germline genomic rearrangements in the FANCA gene by MLPA in Finnish breast cancer families. A novel large deletion affecting the first 12 exons of the gene was uncovered in the index patient of one family. The deletion breakpoints were delineated by Affymetrix Genome-Wide Human SNP 6.0 Array analysis and subsequent nested PCR experiments. The breakpoints lay in the genomic area between the 5’ end of FANCA and the adjacent SPIRE2 gene, and in FANCA intron 12. This resulted in the deletion of 28.8 kb of DNA (c.-4403_1084-126del). The breakpoints occurred at GCTC microhomology sites, and the one in the 5’ region of the FANCA gene was part of an Alu sequence as indicated by the RepeatMasker program. To check whether the deletion allele might be a Finnish founder mutation, allele-specific PCR was designed to screen for its presence in an additional breast cancer cohort unselected for family history of breast cancer (N=540) and in healthy controls.
(N=124). However, no additional carriers were identified. 30 of these controls were also screened by MPLA and found to be negative for any other exonic FANCA deletions or amplifications.

Pedigree analysis of the index patient showed incomplete segregation of the deletion allele with breast cancer, as two healthy carriers were identified at the age of 86 and 31 years (paternal aunt and daughter, respectively). Furthermore, the other breast cancer case (B92, another paternal aunt) in this family appeared not to be a FANCA ex1-ex12 deletion carrier. However, only fragmented tumor DNA was available for mutation testing from this deceased patient. Thus, loss of the mutated FANCA allele or a mutated primer binding site preventing PCR amplification of the deletion allele cannot be ruled out, leaving a slight possibility for the carrier status of this patient. The segregation analysis revealed, however, that the father of the index case with pancreatic cancer is an obligate carrier, and the paternal grandfather of the index case with prostate cancer has 50% chance of being a mutation carrier.

The index patient also had a germline BRCA1 variant with unknown pathogenicity (C3832T encoding Pro1238Leu, Huusko et al. 1998). This variant was also present in the germline of a healthy family member on the maternal side of the index case, and was predicted to be benign by PolyPhen software analysis.

No reported FA symptoms were exhibited by the deletion carriers. Given the hypersensitivity of FA cells to DNA crosslinking agents, two of these individuals were evaluated for a potentially increased frequency of spontaneous and MMC-induced chromosomal aberrations in peripheral blood T-lymphocytes. However, no significant elevation was observed when compared to healthy non-carriers.
6 Discussion

6.1 Lack of mutations in the MERIT40 gene (I)

MERIT40 is involved in the integrity of a BRCA1-containing complex also including RAP80, ABRAXAS, BRCC36, and BRCC45 (Shao et al. 2009a, Feng et al. 2009, Wang et al. 2009). MERIT40 is required for the recruitment of these interaction partners to DNA double-strand breaks upon ionizing radiation and for a (de)ubiquitiniation signaling cascade.

In our present study, the whole coding region and exon-intron boundaries of MERIT40 were screened for germline mutations. Although several new alterations were found, they are unlikely to be pathogenic. Two nucleotide changes, however, may be functionally important. c.*87G>A was found in one index patient, but not in the control population, suggesting that it is either a rare benign variant or that it might exert a detrimental effect if it affected splicing or stability of the mRNA. The only amino acid change predicted, Lys274Arg, is also a novel variant, and though its frequency was not significantly elevated in cases compared to that of the controls, we cannot fully exclude the slight possibility that it may represent a low penetrance susceptibility allele. Large-scale association studies would be needed to address this issue.

Indeed, subsequently to performing study I, the SNP rs8170 (Lys279Lys near Lys274), which was detected in 49 of our breast cancer families, was shown to be associated with a slight increase in breast and ovarian cancer risk by large-scale GWAS studies utilizing thousands of cases and controls (Antoniou et al. 2010, Bolton et al. 2010). rs8170 was found to increase the risk for triple-negative breast cancer in the general population 1.25-fold and in BRCA1 mutation carriers 1.27-fold (Antoniou et al. 2010). In our study, we also observed a slightly elevated frequency of this variant in familial breast cancer cases versus controls, but in order to demonstrate a statistically significant association for this kind of low-risk allele, we would have needed much more larger sample sizes, like in these multicenter consortium studies. In the main MERIT40 transcript, this SNP does not result in an amino acid change, but in one of its alternative splice variants the change is predicted to be non-synonymous (Ser281Arg) (Bolton et al. 2010). This variation might have an impact on MERIT40’s cellular functions, but it is not known yet what the roles of MERIT40’s alternative splice variants are.

Unexpectedly to the predicted function of MERIT40 as a TSG, almost half of the
ovarian tumors studied showed copy number gain at the chromosomal region, where \textit{MERIT40} resides, and the gene was significantly overexpressed in both ovarian tumors and cancer cell lines (Bolton \textit{et al.} 2010). Overexpression is rather a property of oncogenes than TSGs, but context-dependent oncogenic functions of \textit{MERIT40} might exist. Alternatively, it is possible that rs8170 is not the main variant responsible for the increased cancer risk, but is only in linkage disequilibrium with it.

A matter that could be of further interest is the observation that all of the amino acid changes predicted in the current study are silent ones, except for Lys274 that is substituted in the most conservative fashion. This notion could reflect the functional essentiality of these amino acid residues. It might be speculated that Lys274 or Lys279 are potential post-translational modification sites, and a switch from Lys to Arg at position 274 could modulate \textit{MERIT40}’s role in DNA damage signaling. Consequently, a better understanding of \textit{MERIT40} functions, and particularly the role of its first 300 residues, would be helpful in evaluating the significance of genetic alterations in \textit{MERIT40} regarding normal cellular functions and human disease.

In conclusion, the present data suggest that severe mutations predisposing to breast cancer are either very rare or absent in the coding region of \textit{MERIT40}. New data from multi-center collaborations, however, show that some \textit{MERIT40} alleles could slightly influence breast cancer risk. It would be interesting to investigate whether large genomic rearrangements or mutations in transcriptional regulatory regions located further away from the coding region could represent alternative and more typical ways of dysregulating \textit{MERIT40} function. On the other hand, it might also be that severe mutations in this gene are poorly tolerated, pointing to the essentiality of \textit{MERIT40} in maintaining genomic integrity.

6.2 \textbf{ABRAXAS is a breast cancer susceptibility gene in the Northern Finnish population (II)}

ABRAXAS is involved in the integrity and DNA damage signaling of a BRCA1 multiprotein complex containing \textit{MERIT40} and RAP80. Thus, mutations interfering with these functions could induce genomic instability and, ultimately, malignancy. In the current study, we screened the \textit{ABRAXAS} gene for mutations, and identified the c.1082G>A (Arg361Gln) missense mutation in 3 out of 125 familial breast cancer patients and in one additional unselected breast cancer patient with a familial background of cancer. Segregation of \textit{ABRAXAS}
c.1082G>A with the cancer phenotype could not be confirmed in the families, due to lack of DNA from most of the family members. However, it is notable that in family BR-0194, the index patient’s granddaughter was diagnosed with neuroblastoma, as variations in \textit{BARD1} contribute to susceptibility of high-risk neuroblastoma as well (Capasso \textit{et al.} 2009), and mutations in \textit{NF1} are also associated with this cancer type (The \textit{et al.} 1993, Martinsson \textit{et al.} 1997, Origone \textit{et al.} 2003). The neuroblastoma patient’s father was positive for the mutation, thus she has 50% chance of being a carrier.

\textit{ABRAXAS} c.1082G>A seems to be a Northern Finnish founder mutation, as it was absent from the Southern Finnish, as well as from Spanish (Osorio \textit{et al.} 2009), Ashkenazi Jewish, Canadian, and Swiss familial breast cancer cases (Novak \textit{et al.} 2009). Of all identified variants in \textit{ABRAXAS}, at present the missense mutation found in the current study appears to be the only one significantly associated with hereditary susceptibility to breast cancer.

The partly cytoplasmic localization of the Arg361Gln mutant \textit{ABRAXAS} is consistent with a deleterious change in a NLS, thus verifying the functional nature of the first of the two putative NLS sequences (Kim \textit{et al.} 2007a). Colocalization of Arg361Gln \textit{ABRAXAS} with DNA damage-induced foci is decreased, presumably because it has reduced nuclear accumulation. In addition, it may also be possible that the alteration adversely affects protein folding. The finding that deletion of the last 159 residues renders \textit{ABRAXAS} cytoplasmic and prevents it from \textit{BRCA1} binding (Kim \textit{et al.} 2007a) is in agreement with the presence of an NLS in this area of the protein, as well as with the importance of the \textit{ABRAXAS} C-terminal segment for proper protein folding. This finding, together with the fact that not all mutant \textit{ABRAXAS} protein was cytoplasmic in our assay, implies that either the second NLS candidate is functional too, or that other members of the \textit{BRCA1} complex are partly capable of importing \textit{ABRAXAS} into the nucleus on their own.

Importantly, based on the central role of \textit{ABRAXAS} in the \textit{BRCA1} complex, it is likely that \textit{BRCA1} and the other components of the complex are also affected by the mutation, compromising DNA damage signaling. We propose that the deregulated genome integrity maintenance function of the \textit{BRCA1} holoenzyme complex underlies breast cancer pathogenesis in the affected patients. We speculate that the c.1082G>A allele might confer at least a moderately increased genetic susceptibility to breast cancer, as suggested by its total absence from healthy female controls and functional defect.
In conclusion, although there had previously been no exclusive evidence for the involvement of ABRAXAS in breast cancer susceptibility, here we have for the first time identified a coding variant with statistically significantly different distribution in cancer vs. control cases. This alteration is likely responsible for an increased breast cancer risk, and is prevalent in a significant 2.4% of the studied Northern Finnish familial breast cancer cases. Based on these novel findings, ABRAXAS is a breast cancer predisposition gene in this population.

6.3 Lack of large genomic rearrangements in the BRIP1 and CHK1 genes in Finnish breast cancer patients (III)

Despite BRIP1’s established role as a FA and moderate penetrance breast cancer susceptibility gene, the role of large BRIP1 deletions in breast cancer susceptibility has not been studied extensively prior to this study. Only one report has been published - during the time our manuscript was in preparation - investigating large genomic rearrangements in BRIP1 in 716 Dutch familial breast cancer patients, also with the MLPA method. However, no evidence was found for rearrangements (Ameziane et al. 2009). All described BRIP1 mutations in breast cancer and FA-J patients to date are of small size, but some of them resulted in truncated proteins (Levitus et al. 2005, Levrann et al. 2005, Litman et al. 2005, Seal et al. 2006, De Nicolo et al. 2008). One of the truncating germline mutations was also found in a prostate cancer patient (Kote-Jarai et al. 2009). Truncating variants could be expected to have similarly devastating effects on protein function as large-scale deletions or coding region disrupting inversions and duplications. In the Finnish population no deleterious BRIP1 variants were found by conventional mutation screening methods, except for the potential disease-related allele Pro1034Leu (Karppinen et al. 2003, Vahteristo et al. 2006). We have now complemented this analysis by excluding the presence of the intronic cryptic splice site mutation IVS11-498A>T that had been identified in a FA-J patient (Levitus et al. 2005, Levrann et al. 2005) and by identifying IVS11del-435–439TCTT in a breast cancer patient. Not being deposited in public databases and with no predictable effect on splicing, its effect is unclear.

CHK1 is a critical maintainer of cell cycle checkpoints and genomic stability, and is also involved in the BRCA1 and FA protein signaling pathways. Although CHK1 is a protein with crucial importance for cell cycle and DNA integrity maintenance control, no mutations in this gene have yet been associated with predisposition to cancer. This is not entirely explainable by its essential nature,
because CHEK2, which is a functionally related breast cancer susceptibility gene, was reported to harbor an extensive germline deletion connected to familial breast cancer (Walsh et al. 2006). Furthermore, ATR, which is another essential protein, phosphorylating CHK1, is biallelically mutated in Seckel syndrome (O’Driscoll et al. 2003). These examples demonstrate that at least some essential genes are able to tolerate severe inactivations, resulting in viable, but diseased phenotypes.

However, no large deletions or amplifications were uncovered in either of these genes. Thus, to date, no large-size germline BRIP1 deletions have been reported in either FA or hereditary breast cancer. The only known extensive genomic BRIP1 alteration was found in the MCF7 breast cancer cell line, where its translocation resulted in the loss of the last 3 exons (Hampton et al. 2009). It is not known, though, whether this translocation exists in the original cancer specimen or was introduced by cell culturing conditions. Our results indicate that extensive BRIP1 alterations are absent from the Finnish population, or are extremely rare. As even early truncating germline mutations occur in BRIP1 (Levitus et al. 2005, Seal et al. 2006), it is intriguing to hypothesize about the reason for why large-size deletions are not uncovered. We can speculate that maybe an important dosage-sensitive non-coding transcript is transcribed from the gene. Nevertheless, it would be interesting to examine if BRIP1 has big rearrangements in the germline of some FA-J patients, in whom mutation in only one of the alleles could be identified.

As CHK1 modulates the FA pathway and signaling of key breast cancer susceptibility genes from multiple sides, it can be considered an interesting candidate gene for cancer predisposition. As no such alterations have been detected to date in CHK1, if this gene is mutated in any cancer syndromes, it might occur by more subtle mechanisms, such as targeting regulatory elements. Alternatively, germline mutations might not be tolerated in CHK1.

### 6.4 Identification of the first breast cancer family with an extensive deletion in an upstream Fanconi anemia signaling pathway gene, FANCA (IV)

Though large-size FANCA deletions in FA-A patients are known to be pathogenic, their role in breast cancer susceptibility in a monoallelic context has not been well studied. We uncovered an extensive constitutional FANCA deletion in a breast cancer patient and in some of her family members. The deletion was not present in any of the 124 controls, arguing for a potential conferred cancer risk. It was not
observed in an unselected breast cancer cohort of altogether 540 patients either, suggesting that it is rather a private than a common Finnish founder mutation.

The analysis of additional family members of the index patient indicated that possibly some other, currently unknown cancer susceptibility allele might also segregate in this family, as not all tested cancer cases carried the mutation. In particular, patient B92 is indicative of Li-Fraumeni syndrome, although the tumor DNA that was available from this patient was not suitable for TP53 germline mutation testing. The pedigree analysis also indicated that the penetrance of the deletion allele is incomplete. The family contains two healthy carriers, one of whom, however, is still young considering the age of cancer onset of the other family members. Such a partial segregation of the genotype with the disease and incomplete penetrance is a characteristic of moderate penetrance cancer susceptibility genes.

Of interest is the pancreas cancer case who turned out to be an obligate carrier, as mutations in PALB2 (Jones et al. 2009) and BRCA2 (The Breast Cancer Linkage Consortium 1999, Lal et al. 2000, Hahn et al. 2003, Lynch et al. 2005, Couch et al. 2007, Kim et al. 2009) predispose to pancreatic cancer as well. In addition, another upstream FA gene, FANCC, has been associated with hereditary predisposition to pancreatic cancer (Couch et al. 2005). Notably, to date, no large FANCA deletion studies have been conducted in familial pancreatic cancer. Additionally, this FANCA ex1-ex12 deletion carrier family exhibits one prostate cancer case - although his carrier status is unknown -, thus displaying a similar cancer spectrum as proposed for germline BRCA2 mutations (The Breast Cancer Linkage Consortium 1999).

As the observed deletion extends to the promoter region of FANCA, it is highly unlikely that the mutated allele is transcribed. We can speculate that a decreased FANCA expression can affect the function of the FA nuclear core complex (reviewed by Moldovan & D’Andrea 2009), BRCA1 (Folias et al. 2002), the chromatin remodeling BRG1 protein (Otsuki et al. 2001), or the other protein partners by haploinsufficiency. The proposed impact of this deletion is reminiscent of the downstream FA genes that are involved in genetic predisposition to breast cancer and might be in accordance with the continuum model of TSG inactivation (Alimonti et al. 2010). Consequently, it is plausible that FANCA might act as a moderate penetrance breast cancer susceptibility gene, at least if an extensive mutation disrupts its gene function.

The heterozygous carriers of the FANCA ex1-ex12 deletion did not show an increased frequency for spontaneous or MMC-induced chromosomal aberrations,
typical for FA cells. Functional studies on FA heterozygotes treated with MMC or DEB have yielded contradictory results by other research groups (Auerbach & Wolman 1978, Marx et al. 1983, Cervenka & Hirsch 1983, Rosendorff & Bernstein 1988, Mohseni-Meybodi et al. 2007). Studying the functional effect of such heterozygous alterations might, however, be difficult by short-term assays, as genomic instability in these cells could be mild, and only the age-related accumulation of secondary mutations might lead to a detectable phenotype.

In summary, the involvement of upstream FA genes in breast cancer susceptibility has been questionable. The exclusion of these genes as breast cancer susceptibility factors has largely been based on epidemiological studies and on one mutation screening study in familial breast cancer, analyzing only a total of 88 families (Seal et al. 2003). In the current study, we report for the first time an extensive heterozygous FANCA deletion in familial breast cancer, proposing that large FANCA deletions might have a bigger impact on cancer risk than small mutations. However, the possible involvement of conventional FANCA mutations – especially those disrupting reading frame, splicing or transcript stability – would be also worthwhile to be assessed in Finnish cancer patients. Our results indicate a need for a more extensive evaluation of the role of other upstream FA pathway genes in hereditary susceptibility to breast and pancreatic cancer.
7 Concluding remarks

In the studied Northern Finnish familial breast cancer patient cohort consisting of 125 samples the following breast cancer susceptibility genes had previously been found to be mutated: \textit{BRCA1} and \textit{BRCA2} (six and five cases respectively, Huusko \textit{et al.} 1998), \textit{TP53} (one case, Huusko \textit{et al.} 1999), \textit{PALB2} (three cases, Erkko \textit{et al.} 2007), \textit{RAD50} (two cases, Heikkinen \textit{et al.} 2003), \textit{BARD1} (six cases, Karpinnen \textit{et al.} 2004), \textit{ATM} (two cases, Allinen \textit{et al.} 2002b), \textit{CHEK2} (seven cases, Allinen \textit{et al.} 2001 and unpublished data), furthermore one family had mutation in \textit{RAP80} (Nikkilä \textit{et al.} 2009). These together account for 26.4% of the studied Northern Finnish familial breast cancer cases.

In search for additional genes that could explain breast cancer clustering in this population, we screened five additional breast cancer predisposition gene candidates. The major observations of this study regarding the role of the five genes are the following, together with conclusions from the literature:

1. Severe mutations in \textit{MERIT40} are rare or absent in familial breast cancer patients, but it might be a low penetrance breast cancer predisposition gene.
2. \textit{ABRAXAS} is a new breast cancer susceptibility gene, presumably with moderate or even high penetrance. The recurrent c.1082G>A (Arg361Gln) missense variant seems to be a Northern Finnish founder mutation, and explains 2.4% of the studied familial breast cancer cases.
3. Exonic deletions or amplifications affecting the \textit{BRIP1} and \textit{CHK1} genes seem not to contribute to hereditary breast cancer susceptibility in Finland.
4. 1% of the Northern Finnish breast cancer families carried an extensive germline \textit{FANCA} mutation. Large \textit{FANCA} deletions might contribute to breast cancer susceptibility, potentially in combination with other germline mutations. This observation needs to be confirmed by additional studies.

Thus, altogether, mutations have been found in 29.6% of the studied Northern Finnish breast cancer families. These mutations did not co-occur with other mutations, except in one case, where a \textit{BRCA1} and \textit{RAD50} pathogenic alteration was found together in the same individual. All the analyzed mutations occurred in DNA repair/genome integrity maintenance genes playing key roles in BRCA/FA gene signaling, underscoring their role in hereditary predisposition to breast cancer. With the advance of molecular biological methods, for instance next generation gene sequencing, it will be possible to pinpoint the genetic defect of each cancer patient individually, thus we will have a broader insight on the
genetic defect of the remaining large portion of familial breast cancer cases with currently unknown germline mutations. Alternatively, it may be possible that some of these cancer cases arise due to a yet unknown cellular function.
References


Original publications


Reprinted with permission from Springer Science and Business Media (I, III) and Elsevier (IV).

Original publications are not included in the electronic version of the dissertation.
1077. Reponen, Jarmo (2010) Teleradiology—changing radiological service processes from local to regional, international and mobile environment


1081. Alahuhta, Maija (2010) Type 2 diabetes: risk factors of individuals with working-age on weight management and lifestyle changes

1082. Hurskainen, Merja (2010) The roles of collagens XV and XVIII in vessel formation, the function of recombinant human full-length type XV collagen and the roles of collagen XV and laminin α4 in peripheral nerve development and function

1083. Rasi, Karolina (2010) Collagen XV as a matrix organizer: its function in the heart and its role together with laminin α4 in peripheral nerves


1085. Mäkelä, Kari Antero (2010) The roles of orexins on sleep/wakefulness, energy homeostasis and intestinal secretion


1088. Levy, Anna (2011) Lääkentiset asiantuntijat ja eräsulostus on valhiessa: Lääkentiset asiantuntijat yleisorientaatosit, käsitteet hyvin lääkäriin ominaisuuksista ja potilastyylin hahmotumisesta

1089. Miettinen, Johanna (2011) Studies on bone marrow-derived stem cells in patients with acute myocardial infarction


1091. Stefanius, Karoliina (2011) Colorectal carcinogenesis via serrated route

1092. Männistö, Tuja (2011) Maternal thyroid function during pregnancy: effects on pregnancy, peri- and neonatal outcome and on later maternal health

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/
Szilvia Solyom

BRCA/FANCONI ANEMIA PATHWAY GENES IN HEREDITARY PREDISPOSITION TO BREAST CANCER