Sanna-Maria Karppinen

THE ROLE OF BACH1, BARD1 AND TOPBP1 GENES IN FAMILIAL BREAST CANCER

UNIVERSITY OF OULU; BIOCENTER OULU, UNIVERSITY OF OULU
SANNA-MARIA KARPPINEN

THE ROLE OF BACH1, BARD1 AND TOPBP1 GENES IN FAMILIAL 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 26 June 2009, at 12 noon

OULUN YLIOPISTO, OULU 2009
Karppinen, Sanna-Maria, The role of BACH1, BARD1 and TOPBP1 genes in familial breast cancer
Faculty of Medicine, Institute of Clinical Medicine, Department of Clinical Genetics, University of Oulu, P.O. Box 5000, FI-90014 University of Oulu, Finland; Biocenter Oulu, University of Oulu, P.O. Box 5000, FI-90014 University of Oulu, Finland
Oulu, Finland

Abstract
Approximately 5–10% of all breast cancer cases are estimated to result from a hereditary predisposition to the disease. Currently no more than 25–30% of these familial cases can be explained by mutations in the known susceptibility genes, BRCA1 and BRCA2 being the major ones. Additional predisposing genes are therefore likely to be discovered. This study evaluates whether germline alterations in three BRCA1-associated genes, BACH1 (i.e. BRIP1/FANCI), BARD1 and TOPBP1, contribute to familial breast cancer.

Altogether 214 Finnish patients having breast and/or ovarian cancer were analysed for germline mutations in the BACH1 gene. Nine alterations were observed, four of which located in the protein-encoding region. The previously unidentified Pro1034Leu was considered a possible cancer-associated alteration as it appeared with two-fold higher frequency among cancer cases compared to controls. All the other observed alterations were classified as harmless polymorphisms.

Mutation analysis of the BARD1 gene among 126 Finnish patients having family history of breast and/or ovarian cancer revealed seven alterations in the protein-encoding region. The Cys557Ser alteration was seen at an elevated frequency among familial cancer cases compared to controls (p = 0.005, odds ratio [OR] 4.2, 95% confidence interval [CI] 1.7–10.7). The other alterations appeared to be harmless polymorphisms. To evaluate further the possible effect of Cys557Ser on cancer risk, a large case-control study was performed, consisting of 3,956 cancer patients from the Nordic countries. The highest prevalence of Cys557Ser was found among breast and ovarian cancer patients from BRCA1/BRCA2 mutation-negative families (p < 0.001, OR 2.6, 95% CI 1.7–4.0). In contrast, no significant association with male breast cancer, ovarian, colorectal or prostate cancer was observed.

The current study is the first evaluating the role of TOPBP1 mutations in familial cancer predisposition. The analysis of 125 Finnish patients having breast and/or ovarian cancer revealed one putative pathogenic alteration. The commonly occurring Arg309Cys allele was observed at a significantly higher frequency among familial cancer cases compared to controls (p = 0.002, OR 2.4, 95% CI 1.3–4.2). The other 18 alterations observed were classified as harmless polymorphisms.

Keywords: BACH1, BARD1, BRCA1, breast neoplasms, DNA mutational analysis, genetic predisposition to disease, germ-line mutation, hereditary cancer syndromes, TOPBP1
Acknowledgements

This study was carried out at the Department of Clinical Genetics, University of Oulu and Oulu University Hospital, during the years 2001–2009. I want to sincerely thank all those who participated in this project:

Professor Emeritus Jaakko Leisti and Professor Jaakko Ignatius, the former and the current head of the Department of Clinical Genetics, for giving me the opportunity to work in the field of cancer research and carry out this project at the department. Professor Ignatius is especially thanked for his invaluable advice and suggestions concerning the manuscript of this thesis.

Professor Robert Winqvist, my supervisor, for giving me the opportunity to join his research group and to participate in many projects on such an interesting field of research. He is also thanked for his guidance and support during the project, as well as for providing the good research facilities.

Docent Outi Monni and Docent Minna Tanner, the official reviewers of this thesis, for their criticism and invaluable comments on the manuscript. Anna Vuolteenaho is acknowledged for the language revision of this thesis.

My collaborators and co-authors, Professor Lauri Aaltonen, Docent Bjarni Aagnarsson, Dr Minna Allinen, Docent Adalgeir Arason, Katja Backenhorn, M.Sc., Professor Rosa Barkardottir, Professor Åke Borg, Dr Hannele Erkko, Professor Anne-Marie Gerdes, Dr Mervi Grip, Dr Tarja Ikonen, Dr Gudrun Johannesdottir, Professor Anne Kallioniemi, Professor Juha Kere, Professor Torben A Kruse, Dr Pål Møller, Docent Helmut Pospiech, Dr Katri Pylkäs (née Heikkinen), Dr Katrin Rapakko, Dr Kaarina Reini, Professor Johanna Schleutker, Thomas Sydenham M.Sc., Dr Kirsi Syrjäkoski, Professor Juhani Syväoja, Mads Thomassen, M.Sc. and Jaana Vuosku, Lic.Phil. for their important contribution to this study. Risto Bloigu, M.Sc. and Docent Pentti Nieminen are acknowledged for their help with statistical analyses.

All the former and current group members for creating a friendly working atmosphere in the laboratory and for providing me all kind of help and advice. Especially I want to warmly thank Katri Pylkäs and Katrin Rapakko for all their support and time spend during these years. I owe my special thanks to Katri for her excellent collaboration during the intensive years of research and for sharing many good and bad moments in the lab.

The staff at the Department of Clinical Genetics for all their help and for creating pleasant working atmosphere, and Kati Keränen (née Outila), Arja Tapio and Marika Tähti (née Kujala) for technical assistance. I am grateful to the staff at
the genetics laboratory for their help especially at the very beginning of this project and their expertise as well as helpful and friendly attitude towards my studies.

My family and friends: especially my parents Kaija and Seppo for all their support and help and my brothers Samu-Matti and Petri and their families as well as my wonderful grandma Eila for their encouragement and good times spent together. Finally, I owe my warm thanks to Martti for his care and support, and Saku for bringing so much love and joy to my life.

All the patients as well as their family members who volunteered to participate in this study owe my sincere gratitude.

This study has been financially supported by the University of Oulu, Oulu University Hospital, Foundation for the Finnish Cancer Institute, the Academy of Finland, the Finnish Cultural Foundation, the Maud Kuistila Memorial Foundation, the Cancer Society of Northern Finland, Yliopiston Apteekin rahasto of the University of Oulu, the Tyyni Tani Foundation, the Finnish Cancer Society, the Research and Science Foundation of Farmos, the Ida Montini Foundation and the Finnish Breast Cancer Group, which I gratefully acknowledge.

Oulu, May 2009

Sanna-Maria Karppinen
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>amino acid(s)</td>
</tr>
<tr>
<td>AT</td>
<td>ataxia telangiectasia</td>
</tr>
<tr>
<td>ATM</td>
<td>ataxia telangiectasia mutated gene</td>
</tr>
<tr>
<td>ATR</td>
<td>ataxia telangiectasia and Rad3–related gene</td>
</tr>
<tr>
<td>BACH1/BRIP1/FANCJ</td>
<td>gene for BRCA1-associated C-terminal helicase 1/FANCJ</td>
</tr>
<tr>
<td>BARD1</td>
<td>BRCA1-associated RING domain 1 gene</td>
</tr>
<tr>
<td>BRCA1</td>
<td>breast cancer gene 1</td>
</tr>
<tr>
<td>BRCA2/FANCD1</td>
<td>breast cancer gene 2/gene for Fanconi anaemia subtype D1</td>
</tr>
<tr>
<td>BRCT</td>
<td>BRCA1 carboxy-terminal domain</td>
</tr>
<tr>
<td>CDH1</td>
<td>cadherin 1 gene</td>
</tr>
<tr>
<td>CHK1/CHEK1</td>
<td>checkpoint kinase 1 gene</td>
</tr>
<tr>
<td>CHK2/CHEK2</td>
<td>checkpoint kinase 2 gene</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CSGE</td>
<td>conformation sensitive gel electrophoresis</td>
</tr>
<tr>
<td>C-terminal</td>
<td>carboxy-terminal</td>
</tr>
<tr>
<td>DEAH</td>
<td>Asp-Glu-Ala-His amino acid motif</td>
</tr>
<tr>
<td>DSB</td>
<td>double-strand break</td>
</tr>
<tr>
<td>ESE</td>
<td>exonic splicing enhancer</td>
</tr>
<tr>
<td>FA</td>
<td>Fanconi anaemia</td>
</tr>
<tr>
<td>FA-D1, -J, -N</td>
<td>Fanconi anaemia subtypes D1, J and N</td>
</tr>
<tr>
<td>FANCD1, -J, -N</td>
<td>genes for Fanconi anaemia subtypes</td>
</tr>
<tr>
<td>HDGC</td>
<td>hereditary diffuse gastric cancer</td>
</tr>
<tr>
<td>HNPPC</td>
<td>hereditary nonpolyposis colorectal cancer</td>
</tr>
<tr>
<td>HR</td>
<td>homologous recombination</td>
</tr>
<tr>
<td>HU</td>
<td>hydroxyurea</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase(s)</td>
</tr>
<tr>
<td>kD</td>
<td>kilodalton</td>
</tr>
<tr>
<td>LCL</td>
<td>lymphoplasitoid cell line</td>
</tr>
<tr>
<td>LFS</td>
<td>Li-Fraumeni syndrome</td>
</tr>
<tr>
<td>LOH</td>
<td>loss of heterozygosity</td>
</tr>
<tr>
<td>MLH1, -3</td>
<td>gene for DNA mismatch repair protein MutL homologue 1 and 3</td>
</tr>
</tbody>
</table>
**MRE11** gene for meiotic recombination 11 homologue

**MRN** MRE11, RAD50 and NBS1 complex

**MSH2, -6** gene for DNA mismatch repair protein mutS homologue 2 and 6

**NBS1** Nijmegen breakage syndrome 1 (nibrin) gene

**NLS** nuclear localisation signal

**N-terminal** amino-terminal

**OR** odds ratio

**p53** tumour protein 53

**PALB2/FANCN** gene for partner and localiser of BRCA2/Fanconi anaemia subtype N

**PTEN** phosphatase and tensin homologue gene

**RAD50** gene for DNA repair protein RAD50

**RAD51** DNA repair protein RAD51

**RING** really interesting new gene domain

**SNP** single nucleotide polymorphism

**STK11/LKB1** gene for serine/threonine kinase 11

**TOPBP1** gene for topoisomerase (DNA) II binding protein 1

**TP53** gene for tumour protein 53
List of original publications

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


In addition, some unpublished data are presented.
Contents

Abstract
Acknowledgements 5
Abbreviations 7
List of original publications 9
Contents 11
1 Introduction 13
2 Review of the literature 15
  2.1 Cancer is a complex and multistep genetic disease 15
  2.2 Breast and ovarian cancer 17
    2.2.1 Breast cancer 17
    2.2.2 Male breast cancer 18
    2.2.3 Ovarian cancer 19
  2.3 The inherited basis of cancer 20
  2.4 Inherited susceptibility to breast cancer 21
    2.4.1 General aspects 21
    2.4.2 BRCA1 and BRCA2 – the two major genes for hereditary breast and ovarian cancer susceptibility 23
    2.4.3 Other genes conferring susceptibility to breast cancer 31
    2.4.4 Genetic syndromes associated with an increased risk of breast cancer 38
  2.5 Searching for additional breast cancer susceptibility genes 41
  2.6 BACH1, BARD1 and TOPBP1 are BRCA1-associated proteins 42
    2.6.1 BACH1 – a BRCA1-interacting helicase and Fanconi anaemia protein 42
    2.6.2 BARD1 – a major binding partner of BRCA1 48
    2.6.3 TOPBP1 is a checkpoint protein 54
3 Aims of the study 59
4 Materials and methods 61
  4.1 Subjects (I, II, III, IV) 61
    4.1.1 Breast and/or ovarian cancer patients (I, II, III, IV) 61
    4.1.2 Prostate and colorectal cancer patients (III) 65
    4.1.3 Healthy control individuals (I, II, III, IV) 65
  4.2 DNA extraction (I, II, III, IV) 66
  4.3 Mutation analysis (I, II, III, IV) 66
    4.3.1 Conformation sensitive gel electrophoresis (I, II, III, IV) 66
4.3.2 Minisequencing (III) ................................................................. 67
4.3.3 TaqMan assay (III) ................................................................. 68
4.3.4 Denaturing high-performance liquid chromatography (III) .... 68
4.3.5 Restriction fragment length polymorphism analysis (IV) ..... 69
4.3.6 Direct sequencing (I, II, III, IV) ............................................... 69
4.4 Cell culture (IV) ........................................................................ 70
4.5 Western blot analysis (IV) .......................................................... 70
4.6 Statistical analyses (I, II, III, IV) ............................................... 70
4.7 Ethical issues ............................................................................. 71

5 Results ....................................................................................... 73
5.1 BACH1 germline alterations are rare among Finnish families
with breast and/or ovarian cancer (I) ............................................. 73
5.2 Mutation screening of the BARD1 gene (II, III) ......................... 75
  5.2.1 Prevalence of BARD1 germline alterations among Finnish
       families with breast and/or ovarian cancer (II) ....................... 75
  5.2.2 Investigation of the BARD1 Cys557Ser alteration (II, III) .... 76
5.3 TOPBP1 germline alterations among Finnish families with breast
and/or ovarian cancer (IV) .......................................................... 81
  5.3.1 Genetic analysis of TOPBP1 .................................................. 81
  5.3.2 Functional studies of TOPBP1 (partly unpublished data) ..... 83

6 Discussion .................................................................................. 85
6.1 BACH1 and hereditary predisposition to breast cancer (I) ......... 85
6.2 BARD1 alterations and breast cancer susceptibility (II, III) .... 91
  6.2.1 The BARD1 Cys557Ser variant seems to be associated
       with genetic predisposition to breast cancer (II, III) .......... 94
6.3 TOPBP1 germline alterations do not seem to play a significant
role in breast cancer in Finland (IV) ............................................. 96

7 Summary and concluding remarks ........................................... 99
References .................................................................................... 103
Original publications ................................................................. 137
1 Introduction

Cancer is a complex genetic disease caused by a combination of both genetic and environmental factors. Breast cancer is the main cause of cancer-related death in women worldwide, and more than a million new cases are diagnosed every year. In the general population, the lifetime risk of breast cancer is about 10% by the age of 80 years. Ovarian cancer is much rarer, with a lifetime risk of 1.8%, but it is one of the most lethal cancers. Breast cancer accounts for almost 25% of all female cancers worldwide, while the corresponding proportion for ovarian cancer is 4.6%. Each year in Finland approximately 4,000 women are diagnosed with breast cancer and about 450 with ovarian cancer (Finnish Cancer Registry). The strongest risk factor for developing breast or ovarian cancer is a positive family history of the diseases.

Most breast and ovarian cancers are considered sporadic, but about 5–10% of the cases result from inherited predisposition (Claus et al. 1996, Narod & Foulkes 2004, Parkin 2004). Two major predisposing genes, BRCA1 and BRCA2 (i.e. FANCD1), confer a high disease susceptibility, and account for the majority of families displaying both breast and ovarian cancer or only ovarian cancer, while the mutation rate in families with only breast cancer is only about 20% (Ford et al. 1998). Only a small fraction, about 5–10%, of the remaining familial breast cancer cases are currently attributed to germline mutations in other known susceptibility genes. These genes include CHEK2, BACH1 (i.e. BRIP1/FANCI), PALB2 (i.e. FANCN), RAD50 and NBS1 as well as TP53, PTEN, ATM and STK11 (i.e. LKB1) that are linked to certain rare inherited cancer predisposing syndromes. In conclusion, no more than 25–30% of the familial risk of breast cancer is accounted for by currently known susceptibility genes (Easton 1999, Antoniou & Easton 2006), strongly suggesting that additional predisposing genetic factors remain to be discovered.

Due to genetic heterogeneity, the identification of additional breast cancer susceptibility genes has proved to be difficult. Despite rapid advances in the high-throughput processing of DNA samples, the completion of a comprehensive genetic map in the 1990s and the sequencing of the human genome in 2000, the existence of a third high-penetrance gene still remains elusive. As a consequence, a polygenic model has been proposed to account for the remaining familial aggregation of breast cancer. According to this model, a large number of low- to moderate-penetrance genes together are responsible for the unexplained disease predisposition (Antoniou et al. 2002, Pharoah et al. 2002, Wooster & Weber...
As conventional approaches, including linkage analysis that was used for the mapping of \textit{BRCA1} and \textit{BRCA2}, have turned out to be ineffective when trying to find susceptibility genes of lower penetrance, alternative approaches such as large-scale association studies and population-based case-control studies have been utilised. The isolated Finnish population with potentially fewer disease predisposing genes than in populations of more heterogeneous origin is likely to offer some advantages in the search for additional low- and moderate-penetrance susceptibility genes.

Familial breast cancer is typically associated with germline mutations in different genes in pathways critical to genomic integrity and DNA repair. Additionally, several of the known susceptibility genes encode protein products that associate with and act in similar cellular pathways as \textit{BRCA1} and \textit{BRCA2}. Therefore also other genes involved in these pathways represent good candidates for breast and ovarian cancer predisposition.

This study focuses on assessing the contribution to breast cancer susceptibility of germline mutations in three genes that encode \textit{BRCA1}-associated proteins: 1) \textit{BACH1}, which has been found as a gene defective in the Fanconi anaemia subtype J and encodes a protein that interacts directly with \textit{BRCA1} and contributes to its cellular functions; 2) \textit{BARD1}, which encodes a protein that is the constitutive nuclear partner of \textit{BRCA1} and plays a central role in the regulation of stability, cellular localisation and function of \textit{BRCA1} and 3) \textit{TOPBP1}, which encodes a checkpoint protein that associates with \textit{BRCA1} in the response to DNA damage. The protein encoding regions and exon-intron boundaries of these genes were screened for mutations in patients having a family history of breast and/or ovarian cancer. In order to evaluate the pathogenicity of the observed alterations, their frequencies were also determined in healthy controls. Additionally, the frequency of the potentially disease-related \textit{BARD1} Cys557Ser alteration was further studied in a large cohort consisting of samples of almost 4,000 cancer patients from the Nordic countries.
2 Review of the literature

2.1 Cancer is a complex and multistep genetic disease

Cancer arises from a stepwise accumulation of multiple genetic changes, each conferring one or another type of growth advantage, together leading to the transformation of normal human cells into malignant derivatives. Thus, a normal cell requires multiple changes to its genome before it can take on the role of a cancer cell. The exact number of mutations that confer growth advantage on the cells carrying them in an individual cancer is not well established, but it is likely to be dependent on both the type of alteration and tumour type. It has been suggested that common adult epithelial cancers, such as breast, colorectal and prostate, require 5–7 mutations, while cancers of the haematological system may require fewer. However, some analyses indicate that individual cancers may evolve through mutations in as many as 20 different cancer-associated genes. (rev. in Stratton et al. 2009.) A panel of genes that are mutated in colon cancer has been proposed (Kinzler & Vogelstein 1996), and several genes such as TP53 have been evaluated extensively in a wide range of tumour types. However, not a single human cancer type has yet been described with a comprehensive picture of altered genes. In the future large-scale sequencing of cancer genomes will provide direct estimates of the numbers of mutations in individual cancers and lead to better understanding of cancer progression.

Tumour development is associated with a breakdown of the mechanisms that control cell division and the maintenance of genome integrity. The process from normal cell proliferation towards malignant growth requires several alterations in the physiology of the cell. These events include achievement of autonomous growth signalling, evasion of growth inhibitory signals, avoiding programmed cell death (apoptosis), unlimited potential of cell division, capacity to induce and sustain angiogenesis and ability to invade adjacent tissues and metastasise (Hanahan & Weinberg 2000). Most of these capabilities are acquired through changes in the genome of cancer cells, and the means by which they are obtained vary mechanistically and chronologically. Genome maintenance mechanisms protect cells from cancer development either by correcting damaged DNA with a complex network of repair systems or by removing the genetically damaged cell by apoptosis. Consequently, the likelihood of a single cell to accumulate the required number of alterations to become malignant is very low (Loeb 1991). As
cancer still appears at a substantial frequency in the human population, there must be mechanisms by which cells having permanent DNA damage acquire increased mutability, resulting in tumour progression after accumulation of genetic lesions over several decades.

Two major classes of cellular genes that are involved in tumour development include proto-oncogenes and tumour suppressor genes. Normal cell growth is regulated by a finely tuned balance between these two groups of genes, and the result of an altered molecular signalling is often cancer. The normal activity of a proto-oncogene supports cell proliferation and differentiation. Proto-oncogenes encode many of the proteins in the complex signalling circuitry that enables a normal cell to respond to exogenous growth factors. The functions of proto-oncogenes include protein phosphorylation and transcription regulation. Many of these actions regulate gene expression and cell-cycle dynamics. When abnormally regulated, proto-oncogenes turn into cancer-promoting oncogenes which force cells to grow even in the absence of growth factors. Mutation in one allele of an oncogene is enough to promote uncontrolled growth. Tumour suppressor genes encode proteins whose absence, repression, expression inactivation or mutation promotes oncogenesis. Tumour suppressor genes function as negative growth regulators and often encode proteins involved in cell cycle regulation, induction of apoptosis, cell-cell interactions, DNA repair and the maintenance of genomic integrity. Having lost a function of a certain tumour suppressor gene the cell may loose its ability to properly respond to growth-inhibitory signals and become malignant. The function of tumour suppressor genes was first explained by Knudson’s two-hit model, in which both alleles of the gene must be inactivated to cause cancer (Knudson 1971). In addition to proto-oncogenes and tumour suppressor genes, DNA repair genes are often classified distinctly as the third class of genes involved in the cancer development. There are several pathways in human cells for repairing damaged DNA (rev. in Hoeijmakers 2001), and these involve more than one hundred human gene products, of which numerous are known to be involved in cancer development.

Cancer is a common disease caused by a combination of both genetic and environmental factors. In 2002 an estimated 10.9 million people worldwide were diagnosed with cancer and there were 6.7 million deaths from the disease — 12% of deaths worldwide. The proportion of all deaths caused by cancer varies, from 4% in Africa to 23% in Northern America. The geographical distributions of cancer incidence rates are likely explained by the variation in exposure to carcinogens in the external environment and through lifestyle choices as well as
in genetic susceptibility to them. In 2002 the five most common cancers in the world for both men and women were, in terms of new cases, lung (1.4 million), breast (1.2 million), colorectal (1.0 million), stomach (0.9 million) and prostate (0.7 million) cancer. Among women, breast and cervical cancer were the most common malignancies. For men, the most common malignancies worldwide were lung and stomach cancer. (Ferlay et al. 2004.) In Finland, the five most common cancers in the year 2007 among women were breast (4160), rectal (831), uterine (corpus uteri) (789), lung (683) and brain (587) cancer, and among men prostate (4198), lung (1512), rectal (741), bladder (590) and skin (non-melanoma) (575) cancer (Finnish Cancer Registry).

Despite a continuous increase in total cancer incidence during the past fifty years, there has also been significant improvement in the treatment and prevention of the disease. Nonseminomatous germ cell tumours and childhood acute leukaemias were once uniformly fatal, but the cure rates now approach 80% for some subtypes, including widely metastatic nonseminomatous germ cell testicular cancer. Furthermore, some adult leukaemias and lymphomas are curable, and combination chemotherapy for early breast cancer reduces the recurrence rates. Much effort is still needed in the fight against cancer. Preventing the development of a metastatic disease regarding most solid tumours in both children and adults still remains an unsolved issue. Furthermore, many treatments, even when curative, are often associated with toxicity.

2.2 Breast and ovarian cancer

2.2.1 Breast cancer

One in ten of all new cancers diagnosed worldwide is a cancer of the female breast. It is the most common malignancy among women, accounting for about one-fourth of the estimated annual 4.7 million cancer diagnoses in females (Ferlay et al. 2004). In Finland over 4,000 women were diagnosed with breast cancer in 2007 (Finnish Cancer Registry). There is at least a 10-fold variation in the breast cancer incidence rates worldwide, largely as a consequence of a range of socio-economically correlated differences in the population prevalence of several reproductive, hormonal and nutritional factors (rev. in Bray et al. 2004). Currently it appears that environmental rather than genetic determinants are
responsible for most of the observed international and inter-ethnic differences in breast cancer incidence.

In the Western world, the estimated lifetime risk for developing breast cancer is approximately 10% by the age of 80 years in the general population (Feuer et al. 1993). The most significant risk factor is a family history of the disease; breast cancer is approximately twice as common in first-degree relatives of affected individuals as in the general population. Other risk factors for breast cancer predominantly relate to oestrogen exposure. Hormonal and reproductive factors, such as early age at menarche, older age at first birth, short period of breast feeding, nulliparity, late onset of menopause, use of oral contraceptives and hormone replacement therapy have all been associated with an increased risk of breast cancer (ESHRE Capri Workshop Group 2004).

Breast cancer is the main cause of cancer-related death among women globally, responsible for about 375,000 deaths in the year 2000 (Ferlay et al. 2004). The prognosis of breast cancer has improved during the last decades due to more effective screening programmes and methods of treatment. In some high-resource countries, mammographic screening has considerably affected breast cancer diagnosis and mortality. The prognosis depends greatly on the clinical stage at the time of diagnosis. Stage I tumours represent about 40% of all breast cancers, and the 10-year survival rate is 90%, whereas for patients diagnosed with a stage IV tumour, the survival rate is poor, only 5% being alive after 10 years (Finnish Cancer Registry). Most breast cancers are thought to originate from the terminal end ducts or intralobular terminal ducts of the breast, which constitute the most proliferative regions of the breast gland. About 95% of malignant breast tumours are carcinomas, i.e. they originate from the epithelium of the mammary gland. Histologically the most common breast tumours are invasive (infiltrating) ductal (75%) and lobular (15%) carcinomas (Li et al. 2003). Rare histological subtypes include medullar, mucinous, papillary, tubular and cribriform carcinomas and Paget’s disease.

2.2.2 Male breast cancer

Breast cancer occurs also infrequently in males; it comprises less than one percent of all cancers in men. In Finland 16 cases were reported in 2007, representing 0.4% of all breast cancers (Finnish Cancer Registry). Germline mutation in \textit{BRCA2} is the highest known risk factor for male breast cancer. Men with \textit{BRCA2} mutation have a lifetime risk of approximately 10% for developing breast cancer,
which is approximately 80–100 times higher than in the general population (Thompson et al. 2001, Thompson & Easton 2004). The frequency of BRCA2 mutations among different male breast cancer populations studied varies from 4 to 40% (Thorlacius et al. 1996, Friedman et al. 1997). In Finland about 12% of male breast cancers are due to germline mutations in BRCA2 (Syrjäkoski et al. 2004b). Other risk factors for male breast cancer include hormonal imbalance and a family history of breast and ovarian cancer (Weiss et al. 2005). Alterations in hormonal milieu include conditions associated with changes in the oestrogen-testosterone ratio; among which the strongest association is with Klinefelter’s syndrome, increased oestrogen levels and androgen deficiency. The majority of breast tumours in males are either infiltrating ductal carcinomas or intraductal carcinomas, representing 90% of the tumours.

2.2.3 Ovarian cancer

Ovarian cancer is the sixth most common cancer among women worldwide (Ferlay et al. 2004). In Finland 421 ovarian cancer cases were reported in 2007, comprising 3.2% of all female cancers (Finnish Cancer Registry). Like in the context of breast cancer incidence rates, significant geographic and ethnic variations have been observed in ovarian cancer incidence as well. The highest rates are reported for Caucasian women in the industrialised countries of Northern and Western Europe and Northern America. Low rates are found in Africa and Asia. Ovarian cancer is much rarer than breast cancer, with a lifetime risk of 1.8%. The strongest risk factor for developing ovarian cancer is a positive family history of breast or ovarian cancer. Early menarche, nulliparity, infertility and the use of fertility drugs are other known risk factors. In contrast, high parity, lactation, tubal ligation, hysterectomy and use of oral contraceptives have been reported to reduce the risk (Lukanova & Kaaks 2005). Ovarian cancer is the leading cause of death from gynaecologic malignancies and the survival is generally low (Sant et al. 2003). Five-year survival rates for ovarian cancer patients are 70% for stage I and 5% for stage IV disease. Early detection of ovarian cancer is difficult, and the high mortality rate is due to over-representation of stages III and IV at the time of diagnosis. Approximately 90% of ovarian cancers are derived from the epithelium, and are histologically classified into the following subgroups: serous, mucinous, endometrioid, clear cell, Brenner (transitional cell), mixed epithelial and undifferentiated tumours. Stromal and germ cell tumours represent a minority of ovarian cancers.
2.3 The inherited basis of cancer

The hereditary nature of cancer has been well recognised for over 100 years, but the inherited aspects of cancer susceptibility have only become more clearly characterised in the last few decades. Studies have focused on families with several individuals affected with the same type of cancer, characteristic familial constellations of different cancers or an excess of specific forms of cancer. Young age at disease onset, individuals with multiple primary cancers and congenital abnormalities are considered typical features of a hereditary cancer syndrome.

About 5–10% of common cancers are estimated to arise from germline mutations in cancer susceptibility genes. Mapping and isolating of several highly penetrant cancer susceptibility genes has been already completed. These include the \textit{BRCA1} and \textit{BRCA2} genes in breast cancer (Miki \textit{et al.} 1994, Wooster \textit{et al.} 1995) and \textit{MSH2} and \textit{MLH1} genes in colorectal cancer (Fishel \textit{et al.} 1993, Bronner \textit{et al.} 1994). Despite finding high-penetrance genes in many cancers there are genetic components still to be defined, and genetic analysis has shifted towards recognising low- or moderate-penetrance susceptibility genes. New paradigms have been developed for detecting non-Mendelian clustering and the influence of modifier genes and environmental parameters. It may be that even when there clearly is a strong inherited predisposition to certain cancers, the development of cancer is dependent upon secondary factors, which include additional mutations, environmental factors and chance.

Many well-known genetic disorders, including familial adenomatous polyposis (FAP) (rev. in Fostira \textit{et al.} 2007) von Hippel-Lindau disease (VHL) (rev. in Shehata \textit{et al.} 2008) and Li-Fraumeni syndrome (LFS) (Li & Fraumeni 1969), predispose individuals to certain cancers. These syndromic forms of cancer predisposition have been the subject of detailed clinical and molecular studies, and the genes involved in many of these disorders are also well established.

Much is being learned from the studies of cancer-predisposing syndromes and several important characteristics of familial cancer have been identified by examining the families in which they occur. Gene mapping studies of inherited cancers have resulted in the identification of many genes implicated in the initiation of tumour development. For instance, the previously mentioned tumour suppressor genes (see chapter 2.1) were first recognised in studies on inherited forms of cancer.
2.4 Inherited susceptibility to breast cancer

2.4.1 General aspects

Most breast cancers are considered sporadic, but about 5–10% of all cases are associated with family history of the disease (Fig. 1) (Claus et al. 1996, Parkin 2004, Narod & Foulkes 2004). Features suggestive of a hereditary predisposition to breast and ovarian cancer include female breast cancer diagnosed before age 40 years, two or more women diagnosed with breast and/or ovarian cancer, bilateral breast cancer, multiple primary tumours in the same individual and male breast cancer (rev. in Thull & Vogel 2004). Genetic transmission of an autosomal dominant factor responsible for familial breast and ovarian cancer was first reported in the early 1970s (Lynch & Krush 1971). It was not until two decades later that significant progress was made in breast cancer genetics by utilising genetic linkage studies of families with multiple cases of early onset breast cancer, which led to the localisation of a breast and ovarian cancer susceptibility gene, designated BRCA1, to chromosome 17q21 (Hall et al. 1990, Narod et al. 1991). Subsequently in 1994, a genome-wide search using families unlinked to BRCA1 and including male breast cancer, resulted in the localisation of BRCA2 to chromosome 13q12 (Wooster et al. 1994). The BRCA1 gene was cloned in 1994 (Miki et al. 1994), and the BRCA2 gene shortly thereafter (Wooster et al. 1995).

After identifying the BRCA1 and BRCA2 genes it soon became evident that mutations in these two genes do not explain all familial clustering of breast and ovarian cancer (Serova et al. 1997). It has been estimated that BRCA1 and BRCA2 account for 45–50% of the families displaying both breast and ovarian cancer or only ovarian cancer, but only about 20% of the families with only breast cancer (Ford et al. 1998, Anglian Breast Cancer Study Group 2000). Additional breast cancer susceptibility genes have been identified, including PTEN, ATM, TP53, CHEK2, PALB2 and BACH1/BRIP1, but only a small fraction, about 5–10%, of the remaining risk can currently be attributed to these genes. In conclusion, only about 25 to 30% of breast cancer cases can be explained by the currently known susceptibility genes (Easton 1999, Antoniou & Easton 2006) (Fig. 1). It is therefore likely that there are other susceptibility genes that have not yet been identified.
Fig. 1. Breast cancer susceptibility. Inherited breast cancer constitutes five to ten percent of all breast cancer cases. Germline mutations in the two major susceptibility genes, BRCA1 and BRCA2, account for about 20% of the familial cases, and an additional 5–10% of the cases are caused by mutations in other known susceptibility genes. Genetic factors underlying sporadic breast cancer cases are largely unknown. Modified from Bradbury & Olopade 2007.

Since the discovery of BRCA1 and BRCA2, multiple attempts have been made to identify the third major breast cancer susceptibility locus. Several genetic linkage analyses of non-BRCA1/BRCA2 families have suggested various chromosomal regions including 2q, 8p, 9q and 13q potentially harbouring novel breast cancer susceptibility genes (Sobol et al. 1994, Seitz et al. 1997, Kainu et al. 2000, Huusko et al. 2004, Oldenburg et al. 2008). However, these have either been excluded as major predisposing loci by other studies (Rahman et al. 2000, Thompson et al. 2002b), or remain still to be confirmed. The results of these studies have indicated that there are no other highly penetrant allelic variants that predispose to breast cancer. A large number of studies based on conventional histopathology, microarray technology, loss of heterozygosity (LOH) and comparative genomic hybridisation arrays have supported the idea that familial breast cancers that are not associated with mutations in BRCA1 or BRCA2 probably constitute a genetically heterogeneous group (rev. in Wooster & Weber 2003). Subsequently, a polygenic model has been suggested to explain the remaining familial cancers. This hypothesis proposes that the unexplained breast cancer predisposition is due to a combined action of multiple low- and moderate-

The presence of genetic heterogeneity and population-specific effects in breast cancer genetics makes additional susceptibility genes difficult to identify. A fundamental problem in finding lower penetrance genes is that such genes will rarely produce any striking familial patterns involving multiple affected individuals, which is why identifying such genes by conventional approaches (e.g. linkage studies) is not probable. Detection of alleles with lower risk effect, therefore, requires performing large case-control association studies. A suitable approach is to analyse putative candidate genes, being selected on the basis of their biological plausibility, by screening for possible disease-associated mutations.

2.4.2 BRCA1 and BRCA2 – the two major genes for hereditary breast and ovarian cancer susceptibility

Structure and expression

The BRCA1 and BRCA2 genes encode large proteins. BRCA1 is composed of 24 exons, 22 of which are coding, distributed in about 100 kilobases (kb) of genomic DNA on chromosome region 17q21.1 (Miki et al. 1994). The 7.4-kb transcript encodes a protein of 1863 amino acids, with a molecular mass of 220 kilodaltons (kDa). Many alternative spliced transcript variants have been reported for BRCA1 (Lu et al. 1996), but the significance of these with regard to function is not well known. Some of the alternative spliced forms of BRCA1 have been reported to influence its sub-cellular localisation as well as its ability to interact with other proteins, and thus, its physiological functions (Wilson et al. 1997, Thakur et al. 1997). The BRCA2 gene locates to chromosome region 13q12.3 and consists of 27 exons, 26 of which encode a protein of 3418 amino acids, with a molecular mass of 384 kDa (Wooster et al. 1995, Tavtigian et al. 1996). The size of the transcript is 11.2 kb and the entire gene covers approximately 82 kb of genomic sequence (Wooster et al. 1995). Additionally, an alternative spliced BRCA2 transcript lacking exon 3 has been described (Zou et al. 1999).

BRCA1 and BRCA2 do not share significant homology with any other proteins, nor with each other, but both proteins have some characteristic structural features (Fig. 2). Both BRCA1 and BRCA2 contain a large central exon 11, which
encodes approximately 60% and 50% of the protein, respectively. BRCA1 contains two BRCA1 C-terminal (BRCT) domains in its carboxy-terminal region (Koonin et al. 1996). The BRCT motif, a domain of about 80 to 100 amino acids, was first identified in BRCA1 and has subsequently been found in a wide array of proteins involved in DNA repair, DNA recombination, and cell cycle control (rev. in Glover et al. 2004). BRCT domains can bind strand breaks and termini of DNA (Yamane & Tsuruo 1999) and function as phosphopeptide binding sites that can mediate interactions with different target proteins (Rodriguez et al. 2003, Yu et al. 2003, rev. in Glover et al. 2004). The amino terminus of BRCA1 contains RING finger (Miki et al. 1994), a motif found to participate in protein-protein interactions (rev. in Borden 2000) and to have a role in modulating ubiquitination reactions (Lorick et al. 1999). A region of BRCA1 encoded by exon 11 (aa 224-1365) possesses two nuclear localisation signals (NLSs) (Thakur et al. 1997), a DNA-binding domain (Naseem et al. 2006) as well as multiple binding sites for interacting proteins. BRCA2 contains eight conserved BRC repeats, sequence motifs of 30 to 80 amino acids, located in exon 11 (aa 638-2280) (Wooster et al. 1995). The BRC repeats mediate several protein-interactions of BRCA2, including the one with the repair and recombination protein RAD51 (Wong et al. 1997, Katagiri et al. 1998). Through its C-terminal region consisting of a helical domain and three oligonucleotide binding domains BRCA2 is able to bind DNA (Yang et al. 2002). Schematic presentation of the structures of BRCA1 and BRCA2 and binding sites for some of the most important interacting proteins are presented in Figure 2.

BRCA1 and BRCA2 are nuclear proteins and their expression is carefully regulated during the cell cycle (Bertwistle et al. 1997, Ruffner & Verma 1997). For both, the mRNA and protein levels rise at the end of G1, prior to replication onset, and peak during S and G2 phases (Gudas et al. 1996, Vaughn et al. 1996, Bertwistle et al. 1997). Both proteins are ubiquitously expressed in multiple human tissues and show remarkably similar expression profiles. Studies in mice have shown that the expression of BRCA1 and BRCA2 is highest in rapidly proliferating tissues, including ovary, testis, thymus and breast (Miki et al. 1994, Marquis et al. 1995, Rajan et al. 1997). High levels of expression have been observed particularly in mammary epithelial cells during puberty, pregnancy and lactation.
Fig. 2. Schematic structures and selected binding partners of BRCA1 and BRCA2. a) BRCA1 consists of 1,863 amino acids. The RING finger motif, nuclear export signal (NES) (Rodriguez & Henderson 2000), nuclear localisation signals (NLSs) and two BRCT regions are shadowed, and the transcription activation domains (Monteiro et al. 1996, Horwitz et al. 2006) are indicated as bars above the diagram. The approximate regions for selected interacting proteins as well as DNA binding region and clusters for serine and threonine sequences (SCD) are shown as bars below the diagram. b) BRCA2 consists of 3,418 amino acids. The transactivation domain (Milner et al. 1997), eight BRC repeats, helical domain, three oligonucleotide binding domains and NLSs (Spain et al. 1999) are shadowed. The bars below the schematic representation indicate the relative positions of the regions that mediate specific protein interactions.

**Function**

BRCA1 and BRCA2 are multifunctional proteins and both are crucial elements in multiple cellular processes that are responsible for the integrity of the genome both during normal cell cycle and after exposure to DNA damaging agents (rev. in Venkitaraman 2002). Since homozygous *BRCA1* and *BRCA2* mutant mice die early during development, at around on embryonic day 7.5 and 8.5 (E7.5 and
E8.5) respectively, their protein products are definitely essential for embryogenesis (Hakem et al. 1996, Ludwig et al. 1997). Disruption of BRCA1 or BRCA2 also leads to very similar phenotypes. In the absence of functional BRCA1 or BRCA2 proteins cells are sensitive to DNA-damaging agents and exhibit spontaneous and DNA damage-induced chromosomal aberrations (Foray et al. 1999, Moynahan et al. 2001).

The different cellular functions implicated for BRCA1 and BRCA2 are fundamental to all cells, although the roles of each of the two proteins are quite distinct. Both proteins are linked to various processes involved in DNA repair and DNA damage response. These include the repair of double-strand breaks (DSBs) by homologous recombination (HR); an error-free DNA DSB repair pathway, the repair of oxidative damage by transcription-coupled repair (rev. in Chen et al. 1999 and in Venkitaraman 2002) and a possible role of BRCA1 in non-homologous end-joining (rev. in Bau et al. 2006). BRCA1 and BRCA2 have also been proposed to play a role in regulating centrosome amplification (Tutt et al. 1999, Xu et al. 1999), chromatin structure (Fuks et al. 1998, Hu et al. 1999, Bochar et al. 2000) and transcription (Chapman & Verma 1996, Monteiro et al. 1999, Milner et al. 1997). BRCA1 has been implicated in several different checkpoint events (rev. in Deng 2006), and apparently BRCA2 also participates in controlling cell cycle checkpoint (rev. in Lou & Chen 2003). BRCA1 and BRCA2 participate in the Fanconi anaemia pathway. In fact, BRCA2 has been identified as the Fanconi anaemia D1 protein (FANCD1) (Howlett et al. 2002). BRCA2 has also been implicated in the regulation of cytokinesis (Daniels et al. 2004) and the stabilisation of stalled replication forks (Lomonosov et al. 2003) as well as in promoting DNA crosslink repair (Hussain et al. 2003). No known enzymatic activities have been described for BRCA2, whereas BRCA1 displays E3 ubiquitin ligase activity mediated by its RING finger domain (Hashizume et al. 2001). BRCA1 is also associated with X chromosome inactivation (Buller et al. 1999) and has been suggested to play a role in mRNA polyadenylation (Kleiman & Manley 1999, Kleiman & Manley 2001) and in regulating apoptotic cell death (Shao et al. 1996).

The multiple above-mentioned functions of BRCA1 and BRCA2 are mediated through numerous protein-protein interactions that are distinct for each protein (Fig. 2) (rev. in Deng & Brodie 2000, Thompson & Schild 2002, Hohenstein & Fodde 2003, Narod & Foulkes 2004, Yoshida & Miki 2004 and Boulton 2006). BRCA1 re-localises after DNA damage and is a component of several multi-protein complexes that participate in a variety of molecular
functions. In most of these complexes BRCA1 exists as a heterodimer with its major cellular binding partner BARD1 (see chapter 2.6.2). Over 100 proteins have been described to interact with or exist in complexes with BRCA1. For example, BRCA1 associates with the SWI/SNF chromatin remodelling complex (Bochar et al. 2000) and a complex with γ-tubulin, a centrosomal component essential for nucleation of microtubules (Hsu & White 1998). Other BRCA1-associated proteins include BACH1 (Cantor et al. 2001), BARD1 (Wu et al. 1996), TOPBP1 (Greenberg et al. 2006), FANCA (Folias et al. 2002), p53 (Zhang et al. 1998) and ABRA1 (Liu et al. 2007). Several protein-binding partners, though not by far as many as for BRCA1, have also been described for BRCA2, which has been suggested to function as a scaffold protein that helps promote the formation of multiprotein complexes. BRCA2 interacts with the RAD51 and DMC1 recombinases to perform its crucial function in the HR machinery, both in the repair of DSBs and meiosis (Chen et al. 1998b, Thorslund et al. 2007). Examples of other BRCA2-associated proteins are p53 (Marquis et al. 1995), FANCD2 (Hussain et al. 2004), FANCG (Hussain et al. 2003), androgen receptor (AR) (Shin & Verma 2003) and the recently identified PALB2 protein that promotes the nuclear localisation and stability of BRCA2, allowing it to function in DNA repair and checkpoint control (Xia et al. 2006). BRCA1 and BRCA2 also associate with each other (Chen et al. 1998b).

Since the cloning and characterisation of BRCA1 and BRCA2 (Miki et al. 1994, Wooster et al. 1995) there have been numerous attempts to define their biochemical and biological functions as well as their role in breast and ovarian cancers. Although the known functions of BRCA1 and BRCA2 underlie their roles in carcinogenesis, the biological basis of the cancer predisposition in BRCA1 and BRCA2 mutation carriers is currently incompletely understood. It is unclear why mutations in these genes predominantly lead to cancers of the breast and the ovary, even if their functions seem equally important in all other cells of the body as well. Why do tumours arise with the highest frequency in oestrogen-responsive tissues? Several hypotheses have been proposed to explain the mechanism underlying this site-specificity of BRCA1 mutations (rev. in Billack & Monteiro 2005). One of them suggests a correlation between BRCA1 and hormonal actions, as BRCA1 has been reported to bind to and inhibit the action of the oestrogen receptor (Fan et al. 2001, Zheng et al. 2001). An alternate hypothesis suggests that loss of heterozygosity for BRCA1 may be higher in the breast and ovary than in other tissues (rev. in Monteiro 2003).
Prevalence, mutation spectra and cancer risks

It is estimated that \textit{BRCA1} and \textit{BRCA2} may account for no more than 45–50% of all familial risk of breast and ovarian cancer, a proportion substantially smaller than was originally thought. Varying proportions of mutations have been found in different studies mainly due to differences in the sensitivity of the various detection methods used, the complexity of the mutation spectrum and the strictness of the inclusion criteria of the studied families. In general, the results suggest that the two genes account for the majority of high-risk breast cancer families as well as families with a history of both breast and ovarian cancer. Most of the breast cancer families with six or more cases of breast and ovarian cancer are due to mutations in \textit{BRCA1} and \textit{BRCA2}. In contrast, most of the families with less than five cases of breast cancer and no cases of ovarian cancer seem to lack mutations in \textit{BRCA1} or \textit{BRCA2}. (Ford \textit{et al.} 1998.) The proportion of breast and/or ovarian cancer families attributable to \textit{BRCA1} and \textit{BRCA2} also varies between different populations and demonstrates considerable variation in coincidence with ethnic and geographical diversity. For example, the proportion has been reported to be about 35% in Sweden (Håkansson \textit{et al.} 1997, Bergman \textit{et al.} 2005), 48% in Spain (de la Hoya \textit{et al.} 2002, Diez \textit{et al.} 2003), 23% in Netherlands (Verhoog \textit{et al.} 2001) and 40% in North America (Frank \textit{et al.} 1998). In Finland, the corresponding proportion varies from 13 to 21 percent, thus being lower than in many other populations (Vehmanen \textit{et al.} 1997, Huusko \textit{et al.} 1998). The frequency of \textit{BRCA1} and \textit{BRCA2} mutation carriers is estimated to be about 2–7% and 6–10% among unselected breast and ovarian cancer cases, respectively (Claus \textit{et al.} 1996, Peto \textit{et al.} 1999, Syrjäkoski \textit{et al.} 2000, Sarantaus \textit{et al.} 2001).

The number of observed \textit{BRCA1} and \textit{BRCA2} mutations also varies among populations. In countries with ethnically mixed populations, \textit{e.g.} the United Kingdom, Canada and the United States, the range of genetic variation is wide and a large number of different \textit{BRCA1} and \textit{BRCA2} mutations are found. In contrast, accumulation of particular mutations due to founder effect has been identified in geographically or culturally isolated populations, such as the Ashkenazi Jews and Finns (Abeliovich \textit{et al.} 1997, Vehmanen \textit{et al.} 1997, Huusko \textit{et al.} 1998, Sarantaus \textit{et al.} 2000), in which a few mutations account for almost all families attributable to \textit{BRCA1} and \textit{BRCA2}. In Finland 11 founder mutations in the two genes are responsible for 84% of all mutation-positive breast cancer families (Vehmanen \textit{et al.} 1997, Huusko \textit{et al.} 1998). Furthermore, three founder
mutations account for most of the BRCA1/BRCA2 mutation-positive families among Ashkenazi Jews (Abeliovich et al. 1997, Phelan et al. 2002). Founder mutations have also been identified in the Icelandic (Thorlacius et al. 1996), Polish (Gorski et al. 2004) and Dutch (Peelen et al. 1997) populations.

The majority of mutations identified in BRCA1 and BRCA2 are deletions or insertions of a few bases or single base substitutions resulting in premature stop codons and splice-site errors (Breast Cancer Information Core, rev. in Narod & Foulkes 2004 and Rahman & Stratton 1998). In BRCA1 the RING and BRCT repeats are the most frequently mutated sites, and most of the disease-associated mutations result in a truncated protein product with loss of the extreme C-terminus and one or both of the BRCT motifs (Breast Cancer Information Core, Couch & Weber 1996, Szabo et al. 2004b). The majority of cancer predisposing mutations in BRCA2 truncate the protein prior to the carboxy-terminal nuclear localisation signals, resulting in mislocalisation (Spain et al. 1999). Large-scale rearrangements, including insertions, deletions or duplications ranging in size from 510 bp to 37 kb have also been identified in the BRCA1 and BRCA2 genes. The frequencies of large genomic rearrangements vary considerably in different populations. Rearrangements seem to account for a significant proportion of BRCA1 mutations in many populations, as in Holland Australia and Denmark, where they constitute 27–36%, 14.9% and 12.5% of all mutations detected in BRCA1, respectively (Petrij-Bosch et al. 1997, Hogervorst et al. 2003, Woodward et al. 2005, Hansen et al. 2008). In contrast, for example in Finland and Sweden, large deletions constitute only 5.9% and 2.0%, respectively, of the BRCA1 mutations identified (Vehmanen et al. 1997, Huusko et al. 1998, Bergman et al. 2005, Pylkäs et al. 2008). Rearrangements of BRCA2 are more rare and have only been reported with low frequencies (0–8%) in few studies, and they seem to be clustered especially to male breast cancer families (Peelen et al. 2000, Gad et al. 2002, Tournier et al. 2004, Agata et al. 2005, Woodward et al. 2005, Engert et al. 2008, Hansen et al. 2008, Pylkäs et al. 2008). In addition to protein-truncating mutations, a large number of amino acid substitutions have been identified in both BRCA1 and BRCA2. Only a very small proportion of these substitutions have been considered disease-causing, while most have unknown status or have been classified as neutral variants. In fact, all confirmed disease-associated mutations in BRCA2 result in a truncated protein, and no clearly deleterious missense mutations have yet been defined. Characterising the missense alterations as either deleterious or neutral is difficult as the effect of these changes on protein function is often unclear.
Inherited heterozygous mutations in *BRCA1* and *BRCA2* predispose women to high risks of breast and ovarian cancer. The second allele is consistently lost in tumour cells isolated from predisposed individuals, confirming at least in part the classic paradigm for tumour suppressor genes. Early studies estimated the risk among both *BRCA1* and *BRCA2* female mutation carriers to be as high as 80% by the age of 70 for breast cancer, and the risk for ovarian cancer to be more than 40% in *BRCA1* and 27% in *BRCA2* mutation carriers (Ford et al. 1998). A more recent study predicted the risk to be lower by using pooled data from 22 published population studies. The average cumulative risks for breast cancer by the age of 70 were 65% (95% CI 44–78) in *BRCA1* mutation carriers and 45% (95% CI 31–56) in *BRCA2* carriers (Antoniou et al. 2003). The corresponding risk estimates for ovarian cancer by the age of 70 were 39% (95% CI 31–56) and 11% (95% CI 2–19) in *BRCA1* and *BRCA2* mutation carriers, respectively. The relative risk of male breast cancer is elevated for both genes, particularly *BRCA2* (Thompson et al. 2002a, Risch et al. 2006). In addition to breast and ovarian cancer, mutations in *BRCA1* and *BRCA2* also contribute to development of certain other cancers. *BRCA1* mutation carriers are suggested to have greater risk of cancer of the uterus, pancreas, stomach, prostate, testicle and colon (Ford et al. 1994, Thompson et al. 2002a, Risch et al. 2006). *BRCA2* mutations are associated with prostate and pancreatic cancer, as well as melanoma and cancer of the stomach, gall bladder and bile duct (Ford et al. 1994, Breast Cancer Linkage Consortium 1999, Risch et al. 2006).

The risk of developing cancer is not identical for all carriers of *BRCA1* and *BRCA2* mutations. Even the carriers of an identical mutation may have very different outcomes: they may develop different cancers, or the same cancer at different ages, or they may never develop cancer at all. The differences in cancer penetrance associated with *BRCA1* and *BRCA2* mutations suggest that other factors, perhaps both genetic and non-genetic, may modify the risk for developing cancer. In fact, polymorphisms in several genes have been suggested to modify breast and ovarian cancer risk in *BRCA1* and *BRCA2* mutation carriers (discussed in more detail in the last paragraph of chapter 2.4.3). While monoallelic germline mutations in *BRCA1* and *BRCA2* predispose carriers to specific cancers, individuals with biallelic mutations in *BRCA2* result in Fanconi anaemia complementation group D1 (Howlett et al. 2002), which is a rare, recessively inherited chromosomal instability syndrome characterised by skeletal abnormalities, bone marrow failure and cancer predisposition (rev. in Joenje & Patel 2001 and in D'Andrea & Grompe 2003). Biallelic mutations in *BRCA1* have
never been convincingly described in humans and are presumed to be embryonically lethal (rev. in Thompson & Easton 2004 and in Evers & Jonkers 2006).

Despite the fact that BRCA1 and BRCA2 mutations contribute to familial breast and ovarian cancer, it is surprising that somatic mutations in these genes do not seem to play a major role in the tumourigenesis of the sporadic forms of these cancers (rev. in Rahman & Stratton 1998). Instead, several studies have reported that loss of BRCA1 expression through epigenetic mechanisms may contribute significantly to sporadic breast cancer (rev. in Mueller & Roskelley 2003). Additionally, dysfunction of BRCA1 has been observed in sporadic basal-like breast cancers (Turner et al. 2007).

2.4.3 Other genes conferring susceptibility to breast cancer

In addition to BRCA1 and BRCA2, several genes have been reported to confer increased risk of breast cancer. Among those are the CHEK2, RAD50, NBS1, BACH1 and PALB2 genes (Meijers-Heijboer et al. 2002, Heikkinen et al. 2006, Seal et al. 2006, Erkko et al. 2007, Rahman et al. 2007). Additionally, genes that are linked to certain inherited cancer predisposing syndromes are known to cause susceptibility to breast cancer, at least in the context of that particular syndrome. These genes include TP53 (associated with LFS) (Borresen et al. 1992, Easton et al. 1993), PTEN (associated with Cowden’s disease) (Brownstein et al. 1978), ATM (associated with AT) (Vorechovsky et al. 1996, Stankovic et al. 1998), STK11/LKB1 (associated with Peutz-Jeghers syndrome) (Lim et al. 2003), CDH1 (associated with HDGC) (Guilford et al. 1999), as well as several DNA mismatch repair genes (associated with HNPCC) (rev. in Peltomäki 2003 and in Abdel-Rahman et al. 2006), and they are discussed in more detail in chapter 2.4.4. Furthermore, several low-penetrance alleles, including that of FGFR2 (Easton et al. 2007, Hunter et al. 2007, Antoniou et al. 2008), have recently been identified through genome-wide or candidate gene association studies.

Like in the case of BRCA1 and BRCA2, also mutations in TP53 confer a high, greater than ten-fold relative risk of developing breast cancer. Additionally, some instances classify PTEN, STK11 and CDH1 as genes whose mutations confer a high susceptibility to breast cancer, even though others suggest that the relative risk conferred by them is likely to be moderate, ranging from two- to ten-fold. This controversy results from insufficient data. It is known that PTEN, STK11 and CDH1 are all associated with genetic syndromes in which the incidence of breast
cancer is elevated, but the actual risk of each has remained unclear. However, there is currently no evidence that mutations in *PTEN*, *STK11*, *CDH1* or in *TP53* account for a substantial fraction of familial breast cancer in the absence of their respective syndromes. Since the syndromes are rare, the risk of mutations in the associated genes to familial breast cancer is low. Mutations in *CHEK2*, *ATM*, *NBS1*, *RAD50*, *BACH1* and *PALB2* are associated with a moderate, about two- to four-fold risk of developing breast cancer. Low-penetrance alleles confer risks that are elevated less than 1.5-fold. (rev. in Campeau *et al.* 2008 and in Turnbull & Rahman 2008.) Table 1 summarises high- and moderate-penetrance breast cancer susceptibility genes.

**CHEK2**

The identification of the truncating mutation 1100delC in *CHEK2* as a breast cancer predisposing allele (Meijers-Heijboer *et al.* 2002, Vahteristo *et al.* 2002) was the first suggestion that there are variants that confer a lower risk of breast cancer than those conferred by *BRCA1* and *BRCA2*, but are more common in the population. The role of *CHEK2* in breast cancer susceptibility was first indicated by the identification of mutations in a family with Li-Fraumeni syndrome (Bell *et al.* 1999). Multiple subsequent studies have demonstrated that *CHEK2* is a moderate-penetrance breast cancer predisposition gene, and the 1100delC variant confers an about two- to three-fold increased risk of breast cancer among women and a ten-fold increased risk among men (Meijers-Heijboer *et al.* 2002, Vahteristo *et al.* 2002, CHEK2 Breast Cancer Case-Control Consortium 2004, Thompson *et al.* 2006, Weischer *et al.* 2007, Weischer *et al.* 2008). However, marked differences in the prevalence of this allele have been observed within different populations. *CHEK2* 1100delC seems to be most common among Caucasian patients of Northern and Western Europe. In the Netherlands and Finland the allele was present in over 5% of the breast cancer cases with a family history of the disease, compared with a frequency of about 1% in controls (Meijers-Heijboer *et al.* 2002, Vahteristo *et al.* 2002). In contrast, the allele is infrequent for example among North and South Americans (Offit *et al.* 2003, Zhang *et al.* 2008), Australians (Jekimovs *et al.* 2005) and Italians (Caligo *et al.* 2004), and has not been seen at all among breast cancer patients from Spain and Asia (Bellosillo *et al.* 2005, Zhang *et al.* 2008). In addition to 1100delC, several other *CHEK2* variants, such as I157T, have been reported to be associated with breast cancer susceptibility, but compared to 1100delC they seem to have lower impact (Sodha
et al. 2002, Cybulski et al. 2004, Kilpivaara et al. 2004, Bogdanova et al. 2005, Walsh et al. 2006). It has been estimated that CHEK2 may account for 1% of all female breast cancer (Meijers-Heijboer et al. 2002, CHEK2 Breast Cancer Case-Control Consortium 2004). Additionally, CHEK2 variants have also been reported to confer an increased risk of certain other malignancies including prostate, bladder and colorectal cancer (Cybulski et al. 2004, Cybulski et al. 2006, Cybulski et al. 2007, Wasielewski et al. 2008).

**PALB2**

Truncating mutations in *PALB2*, the gene encoding a recently identified BRCA2 binding protein (Xia et al. 2006), have been suggested to confer an about two- to four-fold increased risk of breast cancer (Erkko et al. 2007, Rahman et al. 2007). Furthermore, biallelic truncating mutations in *PALB2* have been reported to result in Fanconi anaemia complementation group N (Reid et al. 2007, Xia et al. 2007). A truncating *PALB2* mutation 1592delT was identified in 2.7% of Finnish familial breast cancer patients and in 0.2% of controls (Erkko et al. 2007). This mutation has not been reported elsewhere and may represent a Finnish founder. Additionally, a French Canadian founder mutation Q775X has been reported (Foulkes et al. 2007). Rahman and colleagues identified truncating mutations of *PALB2* in 1% of the British familial breast cancer patients, being absent in the studied controls (Rahman et al. 2007). Several other studies have also reported rare truncating *PALB2* mutations that are associated with breast cancer susceptibility (Tischkowitz et al. 2007, Cao et al. 2009a, Garcia et al. 2009). In conclusion, it seems that except for founder mutations, *PALB2* alterations are individually rare and may confer only a small fraction of familial risk of breast cancer.

**BACH1**

*BACH1/BRIP1* is another recently identified breast cancer susceptibility gene. It encodes a BRCA1-interacting protein (Cantor et al. 2001) and is discussed in more detail in chapters 2.6.1 and 6.1. Rare truncating mutations in *BACH1* are estimated to confer an about two-fold relative risk of breast cancer (Seal et al. 2006). Similarly to both *BRCA2* and *PALB2*, biallelic mutations in *BACH1* are associated with Fanconi anaemia, specifically with complementation group J (Levitus et al. 2005, Levran et al. 2005, Litman et al. 2005).
RAD50 and NBS1

Mutations in two genes, RAD50 and NBS1, encoding proteins that are part of the MRN complex (composed of MRE11, RAD50 and NBS1), have been associated with an increased risk of breast cancer. Germline mutations 687delT and IVS 3-1G>A in the RAD50 gene have been identified as breast cancer susceptibility alleles in the Finnish population (Heikkinen et al. 2006). The truncating mutation 687delT has been reported in 2.5% of breast cancer cases and in 0.6% of controls, and it has been suggested to be a Finnish founder mutation, as it has not been detected in other Nordic cohorts (Sweden, Norway and Iceland). This RAD50 allele has been reported to confer an approximately four-fold increase in breast cancer risk (Heikkinen et al. 2006). A subsequent study reported the RAD50 687delT allele present in southern Finnish cohort and observed also another, rare truncating variant (Q350X) in RAD50 in one British breast cancer patient (Tommiska et al. 2006). A few studies have linked mutations in NBS1, the gene involved in Nijmegen breakage syndrome, to breast cancer predisposition (Gorski et al. 2003, Steffen et al. 2004, Heikkinen et al. 2006). A rare protein-truncating allele (NBS1 657del5) has been studied in various populations but has only been shown to be associated with breast cancer in the Polish, Byelorussian and German populations (Steffen et al. 2006, Bogdanova et al. 2008). NBS1 alterations have also been suggested to contribute to breast cancer in Northern Finland (Heikkinen et al. 2006). In conclusion, RAD50 and NBS1 mutations appear to be rare and confer only a small proportion of familial breast cancer.

Low-penetrance alleles

Several low-penetrance genes have also been associated with breast cancer. The studies of the Breast Cancer Association Consortium (BCAC) have supported the former results that the D302H variant of CASP8 is associated with a reduced risk of breast cancer (Breast Cancer Association Consortium 2006, Cox et al. 2007). Moreover, the results have also supported the notion that alterations in progesterone receptor (PGR) gene and in TGFB1 are associated with an increased breast cancer risk (Breast Cancer Association Consortium 2006, Cox et al. 2007). Additionally, four recent genome-wide association studies have revealed six novel breast cancer-associated loci (8q, 10q, 16q, 5q, 11p and 6q) (Easton et al. 2007, Hunter et al. 2007, Stacey et al. 2007, Gold et al. 2008). Except for 8q, the rest of these loci contain plausible causative genes (FGFR2, TNRC9, MAP3K1, LSP1,
and ECHD1 and RNF146 respectively). Association of the HMMR gene, encoding a centrosome subunit, with breast cancer has also been reported recently (Pujana et al. 2007).

Since breast cancer risks vary widely among BRCA1 and BRCA2 mutation carriers, the involvement of various modifier alleles that could also act independently as low-penetrance susceptibility alleles has been proposed. To date, several genetic variants have been suggested to modify the cancer risk of BRCA1 or BRCA2 mutation carriers, although many of the observations have not been definitely implicated and await further confirmation (rev. in Thompson & Easton 2004, Antoniou & Easton 2006, Hughes 2008 and Walker et al. 2008). The establishment of the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA) in 2005 provided sufficient power for statistical evaluations of candidate modifier genes (Chenevix-Trench et al. 2007). The results of the CIMBA consortium have confirmed the 135G>C polymorphism in the untranslated region of RAD51 as the first reliable modifier of breast cancer risk among BRCA2 mutation carriers (Antoniou et al. 2007), and have also reported that common breast cancer susceptibility alleles in the FGFR2, TNRC9 and MAP3K1 genes associate with breast cancer risk in BRCA1 and BRCA2 mutation carriers (Antoniou et al. 2008).

Future results of large-scale studies will show whether the other indicated modifier or other low-penetrance genes associate with increased breast cancer risk. However, the associated risks are likely to be low and the involved genes probably account for a very small proportion of the excess familial risk of breast cancer.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Syndrome associated with gene</th>
<th>Frequency in breast cancer cases</th>
<th>Carrier frequency in population</th>
<th>Breast cancer risk, relative risk (RR) or OR (95% CI)</th>
<th>Protein features/main functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCA1</td>
<td>Hereditary breast and ovarian cancer</td>
<td>BRCA1 and BRCA2 account for about 20% of familial cases</td>
<td>1/860</td>
<td>Risk by 70 years is 65%, RR&gt;10</td>
<td>Involved in the maintenance of genome integrity, DNA repair and damage response.</td>
</tr>
<tr>
<td>BRCA2/FANCD1</td>
<td>Hereditary breast and ovarian cancer, Fanconi anemia D1</td>
<td>2–7% of unselected cases</td>
<td>1/740</td>
<td>Risk by 70 years is 45%, RR&gt;10</td>
<td>Involved in the maintenance of genome integrity and DNA repair (specifically in HR).</td>
</tr>
<tr>
<td>TP53</td>
<td>Li-Fraumeni syndrome</td>
<td>&lt;0.25% of unselected cases</td>
<td>&lt;1/10000</td>
<td>Risk by 45 years is 50–60%, RR&gt;10</td>
<td>A key cellular protein involved in maintaining genomic stability.</td>
</tr>
<tr>
<td>CDH1</td>
<td>Hereditary diffuse gastric cancer</td>
<td>&lt;0.1%</td>
<td>&lt;1/10000</td>
<td>Lifetime risk is 39–52%, RR 2–10</td>
<td>E-cadherin. Functions in the maintenance of cell polarity and in cell-cell adhesion.</td>
</tr>
<tr>
<td>PTEN</td>
<td>Cowden syndrome</td>
<td>&lt;0.1%</td>
<td>&lt;1/10000</td>
<td>Lifetime risk is 25–50%, RR 2–10</td>
<td>Lipid phosphatase that functions as a tumour suppressor and growth regulator.</td>
</tr>
<tr>
<td>STK11</td>
<td>Peutz-Jeghers syndrome</td>
<td>&lt;0.1%</td>
<td>&lt;1/10000</td>
<td>Risk by 70 years is 30–50%, RR 2–10</td>
<td>Serine threonine kinase that inhibits cellular proliferation and controls cell polarity.</td>
</tr>
<tr>
<td>ATM</td>
<td>Ataxia telangiectasia</td>
<td>2.7% (12/443) of familial cases</td>
<td>1/100–500</td>
<td>Risk by 70 years is 23%, RR 2–4</td>
<td>Serine threonine kinase that has a central role in the response to DNA DSBs and in activation of checkpoint proteins.</td>
</tr>
<tr>
<td>Gene (Location)</td>
<td>Syndrome associated with gene</td>
<td>Frequency in breast cancer cases</td>
<td>Carrier frequency in population</td>
<td>Breast cancer risk, relative risk (RR) or OR (95% CI)</td>
<td>Protein features/main functions</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------</td>
<td>---------------------------------</td>
<td>--------------------------------</td>
<td>------------------------------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>CHEK2 (22q)</td>
<td></td>
<td>4.2% (30/718) of familial and 1.9% (201/10860) of unselected cases</td>
<td>1/100–200 Risk by 70 years is 11%, RR 2–3</td>
<td>Serine threonine kinase involved in cell-cycle checkpoint control, DNA repair and DNA damage response.</td>
<td></td>
</tr>
<tr>
<td>BACH1 (17q)</td>
<td>Fanconi anaemia J</td>
<td>0.7% (9/1212) of familial cases</td>
<td>&lt;1/1000 RR 2 (1.2–3.2)</td>
<td>DEAH helicase that interacts with BRCA1. Functions in DNA repair and checkpoint control.</td>
<td></td>
</tr>
<tr>
<td>PALB2/ FANCN (16p)</td>
<td>Fanconi anaemia N</td>
<td>0.9% (18/1918) of unselected Finnish cases</td>
<td>1/500–1000 Risk 2–4, OR 3.9 (1.5–12.1)</td>
<td>BRCA2-binding protein that promotes the localisation and stability of BRCA2, allowing it to function e.g. in DNA repair.</td>
<td></td>
</tr>
<tr>
<td>RAD50 (5q)</td>
<td></td>
<td>1.3% (2/151) of familial and 2.5% (8/317) of unselected Finnish cases</td>
<td>&lt;1/160 OR 4.3 (1.5–12.5) for 687delT</td>
<td>Part of the MRN complex, which is involved in recognition and repair of DNA DSBs.</td>
<td></td>
</tr>
<tr>
<td>NBS1 (8q)</td>
<td>Nijmegen breakage syndrome</td>
<td>0.8% (56/6812) of Central and Eastern European breast cancer cases</td>
<td>&lt;1/1000 OR 2.8 (1.8–4.4) for 657del5</td>
<td>Part of the MRN complex, which is involved in recognition and repair of DNA DSBs.</td>
<td></td>
</tr>
</tbody>
</table>
2.4.4 Genetic syndromes associated with an increased risk of breast cancer

Certain cancer-predisposing syndromes have been observed to increase the risk of breast cancer (Nathanson et al. 2001). These disorders include ataxia telangiectasia (AT), Li-Fraumeni syndrome (LFS), Cowden syndrome, Peutz-Jeghers syndrome, hereditary diffuse gastric cancer (HDGC) and hereditary nonpolyposis colorectal cancer (HNPCC). (rev. in Thompson & Schild 2002, Thull & Vogel 2004, Campeau et al. 2008 and Turnbull & Rahman 2008.)

Ataxia telangiectasia

AT is an autosomal recessive neurodegenerative disorder characterised by cerebellar ataxia, telangiectasias, immunodeficiencies, genomic instability and predisposition to malignancies including childhood leukaemia and lymphoma as well as breast cancer in adult patients (rev. in Lavin & Shiloh 1997, Khanna 2000). AT is associated with homozygous or compound heterozygous mutations in the \( ATM \) gene (Savitsky et al. 1995). Individuals with AT have a 50- to 150-fold excess of cancer risk (Swift et al. 1987).

Heterozygous carriers of \( ATM \) mutations are phenotypically normal but are reported to be at increased risk of developing cancer, especially of the female breast (Table 1), but perhaps also cancer of the stomach, colon, pancreas, bladder and ovary (Swift et al. 1990, Thompson et al. 2005). The first indications that heterozygote \( ATM \) mutation carriers have an increased risk of developing breast cancer came from the observed excess of disease cases among obligate carriers in AT families (Swift et al. 1987, Swift et al. 1991). Several subsequent studies reported that certain heterozygous \( ATM \) mutations may be associated with elevated breast cancer risk (Easton 1994, Vorechovsky et al. 1996, Stankovic et al. 1998, Chenevix-Trench et al. 2002, Thorstenson et al. 2003, Buchholz et al. 2004), while others did not observe this association (FitzGerald et al. 1997, Chen et al. 1998a, Szabo et al. 2004a). Based on several recent estimates, it is evident that heterozygous \( ATM \) mutations are moderate-penetrance susceptibility alleles that confer an approximately two- to four-fold increased risk of breast cancer (Thompson et al. 2005, Renwick et al. 2006). Although fairly common in the general population \( ATM \) mutations probably do not account for a large proportion of breast cancer families (Dapic et al. 2005).
**Li-Fraumeni syndrome**

LFS (Li & Fraumeni 1969) is an autosomal dominant cancer predisposition disorder characterised by the occurrence of several childhood and adult cancers including soft tissue sarcomas, osteosarcomas, breast cancer, adrenocortical carcinoma, leukaemia, brain tumours as well as pancreatic carcinoma (Li et al. 1988, Malkin et al. 1990).

In 1990, Malkin and colleagues (Malkin et al. 1990) linked the Li-Fraumeni syndrome to the $TP53$ tumour suppressor gene, as they were able to show that the majority of families with classical LFS have germline mutations in this gene. Based on current analyses, the prevalence of $TP53$ mutations is almost 80% in LFS and 30–40% in Li-Fraumeni-like syndrome (LFL) (Varley 2003). The risk of breast cancer in a woman with LFS is high (Table 1) and may occur early, often before age 30 (Le Bihan & Bonaiti-Pellie 1994). More than half of the women with germline $TP53$ mutations who survive childhood cancers are estimated to develop breast cancer by the age of 50 years, and the lifetime disease penetrance approaches 100% (Easton et al. 1993). LFS is rare and although highly penetrant, germline mutations in $TP53$ are uncommon in non-Li-Fraumeni breast cancer families and account for less than one percent of early onset familial breast cancer (Borresen et al. 1992, Easton et al. 1993). Therefore, $TP53$ germline mutations are thought to be a rare cause of breast cancer, except in the setting of LFS or LFL.

In 1999, Bell and coworkers suggested that the $CHEK2$ gene might associate with cancer aggregation in families with LFS or LFL (Bell et al. 1999). However, subsequent studies on the cancer spectrum associated with mutations in $CHEK2$ have not confirmed the role of $CHEK2$ as a cause of LFS or LFL (Cybulski et al. 2004, Cybulski et al. 2006, Thompson et al. 2006, Evans et al. 2008).

**Cowden syndrome**

Cowden syndrome, a multiple hamartoma syndrome, is a rare autosomal dominant disorder that has been associated with macrocephaly, skin papillomas, gastrointestinal polyps as well as benign and malignant tumours of the breast, thyroid and endometrium (Starink et al. 1986, Eng 2000). In about 80% of the cases the syndrome is caused by germline mutations in $PTEN$ (Nelen et al. 1996, Liaw et al. 1997, Marsh et al. 1998). Up to 75% of patients have benign or malignant thyroid pathology, and women with Cowden syndrome have a 25–50%
lifetime risk of developing breast cancer (Table 1) and a 10% risk of endometrial cancer (Brownstein et al. 1978, Starink et al. 1986, Eng 2000, Sabate et al. 2006). *PTEN* does not play a role in familial breast cancer outside the context of Cowden syndrome, as no germline mutations in this gene have been observed in women with a family history of breast cancer (Lynch et al. 1997, Tsou et al. 1997, Carroll et al. 1999).

**Peutz-Jeghers syndrome**

Peutz-Jeghers syndrome, an autosomal dominantly inherited disorder, is associated with a variety of malignancies including stomach, colon, pancreatic, breast, lung, small bowel, thyroid, endometrial and uterine cancer (Giardiello et al. 1987, Boardman et al. 1998). Patients have a 47% risk for developing any cancer by the age of 65 years, and the risk of breast cancer in women with this syndrome is around 30% by age 65 (Table 1) (Lim et al. 2003). The syndrome is also characterised by gastrointestinal polyposis and by melanin pigmentation of the buccal mucosa, lips, fingers and toes (rev. in Tomlinson & Houlston 1997). Germline mutations in the *STK11/LKB1* gene have been associated with Peutz-Jeghers syndrome (Hemminki et al. 1998, Jenne et al. 1998). As many as 50 to 70% of the patients with the syndrome have been reported to have germline mutation in this gene (Wang et al. 1999, Westerman et al. 1999, Olschwang et al. 2001, Lim et al. 2003). Like Li-Fraumeni and Cowden syndromes, Peutz-Jeghers syndrome is rare and accounts for only a very small fraction of breast cancer cases. Mutations in *STK11* do not appear to play an important role in sporadic breast cancer development (Bignell et al. 1998) and have not been reported in familial breast cancer outside the context of Peutz-Jeghers syndrome.

**Hereditary diffuse gastric cancer**

HDGC is a rare autosomal dominant cancer-susceptibility syndrome. The estimated lifetime risk of developing gastric cancer in these families is 67% among men and 83% among women (Pharoah et al. 2001, Kaurah et al. 2007). Affected women are also reported to present an excess of lobular breast cancer (Keller et al. 1999, Kaurah et al. 2007). As many as 30–50% of the HDGC cases are caused by germline mutations in *CDH1*, a gene encoding E-cadherin (Table 1) (Guilford et al. 1998, Lynch et al. 2005, Kaurah et al. 2007). The estimated lifetime risk of lobular breast cancer among female *CDH1* mutation carriers is
39–52% (Pharoah et al. 2001, Kaurah et al. 2007). Hence, although there is strong evidence for establishing lobular breast cancer as part of the HDCG syndrome, the association between CDH1 mutations and breast cancer susceptibility outside the context of this disorder still remains unclear (Masciari et al. 2007, Schrader et al. 2008).

**Hereditary nonpolyposis colorectal cancer**

Germline mutations of human mismatch repair genes cause susceptibility to HNPCC, an autosomal dominantly inherited disorder. To date, altogether five DNA mismatch repair genes, MLH1, MSH2, MSH6, PMS2 and MLH3, have been found to be associated with predisposition to the disease (rev. in Peltomäki 2003 and Abdel-Rahman et al. 2006). Defects in these genes cause microsatellite instability (MSI), which is a hallmark of HNPCC tumours. HNPCC is one of the most common cancer syndromes in humans. In addition to colorectal cancer the syndrome is also associated with cancer of the endometrium, stomach, biliary tract, urinary tract and ovaries. An association between breast cancer and HNPCC has been proposed in several studies (Risinger et al. 1996, Boyd et al. 1999, Scott et al. 2001). By contrast, other studies have shown no excess of breast cancer incidence (Siah et al. 2000, Vasen et al. 2001). Therefore, despite a large number of studies, it has remained controversial whether the incidence of breast cancer is increased among families with HNPCC syndrome.

### 2.5 Searching for additional breast cancer susceptibility genes

Putative candidate genes for cancer susceptibility are usually chosen on the basis of their biological plausibility. Protein products of the genes known to be associated with hereditary predisposition to breast cancer are involved in the maintenance of genomic integrity and DNA repair (rev. in Yoshida & Miki 2004 and Ralhan et al. 2007). Therefore, also other genes coding for proteins in these pathways may be involved in breast cancer development, and have for that reason been under an active research in order to find additional susceptibility genes. Moreover, genes encoding proteins that interact with the products of known predisposition genes represent good candidates to investigate. In particular, genes whose products interact with BRCA1 or BRCA2 have been attractive targets of research over the past years. BRCA1/2-associated proteins may also modify the risk of developing cancers in BRCA1/2 mutation carriers. In addition, the genes
involved in other critical biochemical pathways, including carcinogen and steroid hormone metabolism and immune surveillance, have been hypothesised as candidates (Nathanson & Weber 2001). Investigation of certain cancer-predisposing recessive syndromes could also reveal an excess of breast cancer in relatives of affected individuals and lead to the identification of new candidates. *MRE11, RAD51, BARD1* and *ATR* are among the numerous candidate genes that have been studied for breast cancer susceptibility.

As certain disease alleles are known to be enriched in isolated populations, the use of the genetically relatively homogeneous Northern Finnish population should offer advantages in the search for genes associated with familial breast cancer.

### 2.6 BACH1, BARD1 and TOPBP1 are BRCA1-associated proteins

#### 2.6.1 BACH1 – a BRCA1-interacting helicase and Fanconi anaemia protein

BACH1 (BRCA1-associated C-terminal helicase 1), also known as BRCA1-interacting protein 1 (BRIP1), was originally identified through its direct binding to *BRCA1* and was therefore suggested to be a candidate gene for predisposition to hereditary breast cancer (Cantor et al. 2001). More recently, *BACH1* has been identified as *FANCJ*, the gene defective in Fanconi anaemia complementation group J (Levitus et al. 2005, Levrans et al. 2005, Litman et al. 2005). *BACH1* consist of 20 exons, encoding a 130-kDa nuclear protein of 1,249 amino acids (Fig. 3) (Cantor et al. 2001). The C-terminal part of the protein (aa 979-1006) interacts with the highly conserved BRCT repeats of *BRCA1*, while residues 158-175 comprise a putative nuclear localisation signal (Cantor et al. 2001, Cantor et al. 2004). Amino acid residues 1-888 reveal strong homology with the catalytic and nucleotide binding domains of known DEAH helicases (Cantor et al. 2001). *BACH1* is located at chromosome region 17q22 (Cantor et al. 2001), close to *BRCA1* which maps to 17q21 (Hall et al. 1990, Narod et al. 1991). The presence of other breast cancer-associated genes besides *BRCA1* in the long arm of chromosome 17 has been suggested on the basis of allelic imbalance, LOH and comparative genomic hybridisation studies (Plummer et al. 1997, Phelan et al. 1998, Orsetti et al. 2004).
Fig. 3. Schematic structure of the BACH1 protein consisting of 1,249 amino acids. Roman numerals above the diagram show the locations of the seven conserved helicase homology blocks in the DEAH helicase homology region (hatched area). The other known functional domains of the protein and the BRCA1-interacting region are shadowed and their aa positions are displayed. Amino acid positions of the coding region variants observed in this study (I) are marked above the schematic structure.

**BACH1 contributes to many functions of BRCA1 and plays an important role as a DNA helicase**

BACH1 and BRCA1 exhibit similar features in their responses to DNA damage. Their localisation, immediate presence at sites of DNA damage and DNA damage-induced phosphorylation patterns are alike (Peng et al. 2006). Moreover, BACH1 has been shown to support the localisation of BRCA1 to DNA damage foci (Peng et al. 2006). The expression pattern of BACH1 (Cantor et al. 2001) is also similar to that of BRCA1 (Miki et al. 1994, Zabludoff et al. 1996), both being expressed at some level in all tissues and at highest level in testis. As a physiological partner of BRCA1, BACH1 participates in many of its DNA repair and tumour suppressor functions (Cantor et al. 2001). Both proteins are critical for the repair of DSBs by homologous recombination, mitomycin-C resistance and for maintenance of chromosomal integrity (Litman et al. 2005). The interaction between BRCA1 and BACH1 has been shown to be essential for BRCA1-mediated repair of DNA DSBs (Cantor et al. 2001) and for BRCA1 to mediate DNA damage-induced G2/M checkpoint control (Rodriguez et al. 2003, Yu et al. 2003). Additionally, BACH1 is needed for the DNA damage-induced association between BRCA1 and TOPBP1 and in that context for the activation of the S phase checkpoint (Fig. 6) (Greenberg et al. 2006).
BACH1 binds to the BRCT repeats of BRCA1 (Cantor et al. 2001). BRCT domains are important for BRCA1 focus formation (Au & Henderson 2005) and for the ability of BRCA1 to mediate DNA repair (rev. in Jasin 2002), G2/M checkpoint control (Botuyan et al. 2004) as well as transcriptional activation (Monteiro et al. 1996). Moreover, BRCT domains are essential for the tumour suppressor function of BRCA1, as evidenced by the many breast and ovarian cancer associated protein truncation and missense variants locating in these domains (Breast Cancer Information Core, rev. in Glover 2006). Many mutations (e.g. R1699W, S1655F, M1775R and P1749R) in the BRCA1 BRCT domains are known to cause defects in BRCA1 DNA repair function and to prevent the ability of BRCA1 to bind BACH1 (Cantor et al. 2001, Cantor et al. 2004, Clapperton et al. 2004, Shiozaki et al. 2004). These findings have suggested that loss of BRCA1-BACH1 interaction may contribute to tumour progression.

BACH1 works both as a DNA-dependent ATPase and 5'-to-3' DNA helicase (Cantor et al. 2004), and is a member of the DEAH helicase family. DNA helicases represent an important group of enzymes implicated in genetic recombination, transcription, DNA replication and DNA repair, in order to maintain genome stability (rev. in van Brabant et al. 2000). Abnormal function of DNA helicases has been linked to decreased cell viability or disease development (rev. in van Brabant et al. 2000 and in Hickson 2003). Mutations in the human RecQ helicase genes, WRN (Imbert et al. 1996, Yu et al. 1996), BLM (Ellis et al. 1995) and RECQ4 (Kitao et al. 1999), lead to rare autosomal recessive chromosomal instability disorders, referred to as Werner syndrome, Bloom syndrome and Rothmund-Thomson syndrome that are associated with premature ageing and cancer predisposition (rev. in Hickson 2003 and in Hanada & Hickson 2007). In addition, mutations in two DNA helicase-encoding genes, XPB and XPD, have been linked to other genetic disorders, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy (rev. in van Brabant et al. 2000). Thus, the recent identification of BACH1 as a Fanconi anaemia protein J is consistent with the important role of DNA helicases as caretakers of the human genome (Levitus et al. 2005, Levrar et al. 2005, Litman et al. 2005). Several studies have shown the important role of the helicase activity of BACH1 in maintaining genomic stability. A recent study reported that a helicase disrupting mutation in BACH1 led to increased activation of DNA damage checkpoints and genomic instability (Kumaraswamy & Shiekhattar 2007). In addition, the interaction of the helicase domain of BACH1 with mismatch repair complex has been shown to be required for the correction of the interstrand cross-link response.
in FA-J cells independently of BRCA1 (Peng et al. 2007). The identification of familial breast cancer patients with helicase disrupting mutations in BACH1 has suggested that the helicase activity of BACH1 may also contribute to its tumour suppression function (Cantor et al. 2001, Cantor et al. 2004).

**BACH1 in the Fanconi anaemia pathway**

Fanconi anaemia (FA) is a rare recessively inherited disorder characterised by cellular hypersensitivity to DNA cross-linking agents, chromosome instability and increased predisposition to cancers including both haematological malignancies and solid tumours. The syndrome is also associated with developmental abnormalities, bone marrow failure, diverse congenital defects, skeletal abnormalities, retarded growth, abnormal skin pigmentation and microphthalmia. The prevalence of FA is estimated to be 1 to 5 per million, and heterozygote frequency is estimated to be 1 in 300 in Europe and the United States. (rev. in Joenje & Patel 2001 and in D'Andrea & Grompe 2003.) Fanconi anaemia has extensive genetic heterogeneity. To date, 13 complementation groups have been described (FA-A, FA-B, FA-C, FA-D1, FA-D2, FA-E, FA-F, FA-G, FA-I, FA-J, FA-L, FA-M and FA-N), and with the recent identification of the gene for subtype I all the 13 associated genes have now been identified (rev. in Wang 2007). FA proteins have an important role in protecting the genome. They are part of a complex network of interacting proteins that affect DNA stability both during the S phase of the cell cycle and during recognition and repair of damaged DNA. (Surralles et al. 2004, Bagby & Alter 2006.)

The identification of BRCA2 as the gene responsible for a rare and highly cancer-prone form of FA complementation group D1 (Howlett et al. 2002) has been a key finding linking breast cancer susceptibility and the Fanconi anaemia DNA repair pathway. The recent reports showing that biallelic inactivating mutations also in two other breast cancer susceptibility genes, BACH1 and PALB2, cause Fanconi anaemia have further strengthened the link between the FA pathway and breast cancer (Levitus et al. 2005, Levran et al. 2005, Litman et al. 2005, Reid et al. 2007, Xia et al. 2007). The functional linkage has also been demonstrated by the observation that FA proteins and proteins associated with breast cancer co-operate in the DNA damage response pathway (Fig. 4) (rev. in D'Andrea & Grompe 2003). Several studies have been performed to investigate whether heterozygous mutations also in other FA genes may confer an increased risk of breast cancer. The FANCA, -B, -C, -D2, -E, -F and -G genes have been
screened in breast cancer families, but no pathogenic variants have been found (Seal et al. 2003, Garcia et al. 2009). Additionally, epidemiological studies have assessed whether FA heterozygotes are at increased risk of developing cancer. Two recent studies did not find an increase in the overall cancer incidence, but it did find a possible association with prostate cancer and some evidence that FANCC mutations are breast cancer susceptibility alleles (Berwick et al. 2007, Tischkowitz et al. 2008). Earlier studies have failed to find significantly increased cancer risks for FA heterozygotes (Swift et al. 1980, Potter et al. 1983).

The finding that BACH1 is the gene defect in FA-J patients not only links BACH1 to the FA pathway but also highlights that BACH1 function is critical for normal human development as well as for disease prevention. Patients in the FA-J complementation group have several symptoms which are typical of other FA complementation groups, including café-au-lait spots, thumb abnormalities,
kidney defects, microphthalmia, microcephaly, growth delay and aplastic anaemia (Levitus et al. 2004). However, there are some differences in the phenotypes associated with mutations in \textit{BRCA2}/\textit{FANCD1} and \textit{BACH1}/\textit{FANCI}. Biallelic \textit{BRCA2} mutations confer a high risk of childhood solid and haematological cancers (Reid et al. 2005), whereas to date only one cancer has been reported in an individual with FA-J who has biallelic mutations (Levitus et al. 2005, Levrnan et al. 2005, Litman et al. 2005). Furthermore, monoallelic \textit{BRCA2} mutations confer a high risk of breast cancer, whereas monoallelic \textit{BACH1} mutations confer a moderate risk, similar to truncating mutations of \textit{PALB2}/\textit{FANCN} (rev. in Campeau et al. 2008, Turnbull & Rahman 2008). The biological explanations for these differences are unclear.

\textbf{BACH1 is a recently identified breast cancer susceptibility gene}

The first indication that \textit{BACH1} mutations might predispose to breast cancer came from a study where two heterozygous missense mutations resulting in defective helicase activity were detected in individuals diagnosed with breast cancer (Cantor et al. 2001). Subsequently, a number of studies have been performed to evaluate the possible role of \textit{BACH1} in breast cancer susceptibility (discussed in chapter 6.1) (Cantor et al. 2001, Luo et al. 2002, Lei & Vorechovsky 2003, Rutter et al. 2003, Sigurdson et al. 2004, Lewis et al. 2005, Garcia-Closas et al. 2006, Seal et al. 2006, Vahteristo et al. 2006b, Frank et al. 2007, Song et al. 2007, Cao et al. 2009b, De Nicolo et al. 2008, Guenard et al. 2008). Based on these studies it has been suggested that certain rare \textit{BACH1} missense variants may confer susceptibility to breast cancer. Strong evidence indicating a role for \textit{BACH1} in breast cancer susceptibility came from a recent study showing that inactivating truncation mutations of \textit{BACH1} confer a two-fold relative risk of breast cancer (Seal et al. 2006). Subsequently other studies have also reported inactivating truncating mutations. Moreover, a recent study that completed a sequence-level map of chromosomal rearrangements in the genome of the MCF-7 breast cancer cell line identified a truncation mutation in the \textit{BACH1} gene (Hampton et al. 2009). This observation further supports the involvement of \textit{BACH1} in breast cancer progression. Current evidence indicates that \textit{BACH1} is a moderate-penetrance susceptibility gene conferring an about two-fold risk of breast cancer.
2.6.2 BARD1 – a major binding partner of BRCA1

BARD1, a protein of 777 amino acids, was originally found through its interaction with BRCA1 with which it has both structural and functional similarities (Wu et al. 1996). The two proteins share a closely related domain structure: both possess an amino-terminal RING finger motif and two carboxy-terminal BRCT domains (Fig. 2 and Fig. 5). The heterodimer formation between BARD1 and BRCA1 is mediated through sequences encompassing their RING motifs, and the association seems to be important for several tumour suppression functions of BRCA1. The interaction has also been shown to stabilise both proteins. (Wu et al. 1996, Meza et al. 1999.) Unlike BRCA1, BARD1 also contains four centrally positioned ankyrin repeats (ANK) (Fox et al. 2008) that are found in many proteins involved in transcriptional regulation (Sedgwick & Smerdon 1999). Two different BARD1 transcripts have been identified, full-length BARD1 and a splice variant lacking exons two, three (RING finger) and five (Feki et al. 2004). BARD1 orthologues have been identified in Mus musculus (mouse), Rattus norvegicus (rat), Xenopus laevis (frog), Caenorhabditis elegans (worm) and Arabidopsis thaliana (plant).

BARD1, like BRCA1, is a nuclear protein and its tissue distribution is similar to that of BRCA1. Expression of the two proteins is highest in the testis and spleen, tissues in which cells undergo the high rates of proliferation and apoptosis (Miki et al. 1994, Marquis et al. 1995, Ayi et al. 1998). The relative expression of BRCA1 and BARD1 is critical in determining cell fate. When BRCA1 levels in cells are low, BARD1 accumulates in the cytoplasm where it plays a role in initiating programmed cell death (Irminger-Finger et al. 2001, Jefford et al. 2004, Rodriguez et al. 2004). BARD1 regulates the sub-cellular localisation of BRCA1, both by translocating BRCA1 into the nucleus and blocking its exit by covering up the BRCA1 nuclear export signal (Fabbro et al. 2002). BARD1 is expressed in a cell cycle-dependent manner with maximal levels occurring in mitosis and lower levels in G1 and S phases, thereby coinciding with BRCA1 expression both during mitosis and S phase (Hayami et al. 2005, Choudhury et al. 2004). BARD1 is transcriptionally upregulated also in response to DNA damage and hormone signalling (Irminger-Finger et al. 2001). In breast, ovarian and uterine tissues, the expression of BARD1 and BRCA1 is differentially regulated in accordance to the ovulatory cycle (Irminger-Finger et al. 1998). BARD1 levels are increased during the ovulatory cycle, suggesting that it might have a role in the endometrium after ovulation. Oestrogen receptor α (ER-α) has recently been observed to regulate the
expression of BARD1 through an intronic DNA sequence (Creekmore et al. 2007). BARD1 deficiency seems to be deleterious to the cell. Reduced expression of BARD1 in murine mammary epithelial cells results in genomic instability and cellular changes indicating a premalignant phenotype (Irminger-Finger et al. 1998). Moreover, homozygous depletion of BARD1 in mice results in early embryonic lethality (E7.5) and genomic instability (Wu et al. 1996, McCarthy et al. 2003). The phenotype of Bard1-null mice is remarkably similar to that of Brca1-nulls, suggesting functional relationship between the two proteins. Likewise, indistinguishable phenotypes have also been described for nematodes (C. elegans) harboring mutations in the orthologs of BRCA1 and BARD1 (Boulton et al. 2004).

Fig. 5. Schematic structure and selected binding partners of the BARD1 protein. Functional domains (RING finger, nuclear export signal (NES) (Rodriguez et al. 2004), two nuclear localisation signals (NLSs) (Schuchner et al. 2005), four ankyrin repeats and two BRCT regions) are shadowed and their aa positions displayed. Approximate regions for selected interacting proteins and the region crucial for induction of apoptosis are shown as bars below the diagram. Amino acid positions for the coding region variants observed in this study (II) are marked above the schematic structure.
BARD1 functions as a heterodimer with BRCA1 but has also BRCA1-independent functions

BARD1 has been implicated in multiple crucial cellular processes. As a major binding partner of BRCA1, many functions of BARD1 are mediated through association with it. The role of BRCA1-BARD1 heterodimer in BRCA1-mediated tumour suppression is strongly supported by a recent study that observed that inactivation of either Brca1 or Bard1 in mice induces breast carcinomas that are indistinguishable from each other and from those of double Brca1/Bard1-mutant animals (Shakya et al. 2008). The tumours of these mutant mice were similar with respect to their frequency, latency, histopathology and cytogenetic features and were highly reminiscent of the basal-like breast tumours that arise in human BRCA1 mutation carriers.

The BRCA1-BARD1 heterodimer functions as an E3 ubiquitin ligase (Hashizume et al. 2001, Morris & Solomon 2004) that mediates the transfer of ubiquitin to lysine (K) residue(s) in the target substrates (rev. in Starita & Parvin 2006). In addition to the classical K48-linked ubiquitination which targets proteins for proteosomal degradation, the BRCA1-BARD1 heterodimer has been observed to form at least two unconventional linkages in ubiquitins; through K6 (Wu-Baer et al. 2003, Morris & Solomon 2004, Nishikawa et al. 2004) and K63 (Xia et al. 2003, Christensen et al. 2007). K6-linked ubiquitin structures have been observed at sites of DNA repair and replication (Morris & Solomon 2004), indicating that the ubiquitin activity of the BRCA1-BARD1 induces DNA repair pathways. Numerous targets for the ubiquitination of BRCA1-BARD1 have been suggested. For instance, the heterodimer ubiquitinates in vivo a nucleolar phosphoprotein B23, also known as nucleophosmin (NPM1), which interacts with BRCA1-BARD1 to function in a common tumour suppression pathway. Another target of the ubiquitin ligase is histone 2AX (H2AX) implying that BRCA1-BARD1 activity is involved in chromatin modification. Additional putative substrates for BRCA1-BARD1 ubiquitination include BRCA1 itself, FANCD2, CtIP, ER-α, TFIIE, γ-tubulin, HMMR, RPB1, RPB8 and p53. Numerous of these substrates including NPM1, CtIP, RPB8 and TFIIE have not been shown to be targeted to proteosomes, suggesting different fates for these proteins by the BRCA1-BARD1 mediated ubiquitination. Most of the known substrates for ubiquitination of BRCA1-BARD1 have been shown to be ubiquitinated in vitro, but have not been confirmed to be substrates in vivo. (Pujana et al. 2007, rev. in Starita & Parvin 2006 and in Wu et al. 2008.) Cancer-associated mutations (C61G
and C64G) within the RING finger of BRCA1 abolish the ubiquitin ligase activity, demonstrating that the activity is critical for the tumour suppressor function of the BRCA1-BARD1 complex (Wu et al. 1996, Brzovic et al. 2001, Hashizume et al. 2001, Ruffner et al. 2001). Several studies have reported that the mutated residues disrupt the interaction between BRCA1 and BARD1 (Brzovic et al. 2001, Fabbro et al. 2002, Morris et al. 2002, Morris et al. 2006).

BARD1 plays a crucial role in association with BRCA1 in coordinating cellular responses to DNA damage to maintain genomic stability. Following exposure to DNA damage, BRCA1 and BARD1 are phosphorylated and colocalise in nuclear foci with other DNA repair and checkpoint proteins (Jin et al. 1997, Scully et al. 1997). BARD1 and BRCA1 participate as a heterodimer in different supercomplexes to execute diverse cellular functions. The presence of the BRCA1-BARD1 heterodimer implicates an important role for the ubiquitin ligase activity in these complexes. The BRCA1-BARD1 supercomplexes include RNA polymerase II holoenzyme complex (pol II), MRN complex, BRCA2-RAD51 complex, BACH1-TOPBP1 complex and the recently identified RAP80-ABRA1 containing complex. Each complex contributes to different DNA damage responses and performs certain checkpoint and repair functions. Many of the proteins in these complexes are associated with breast cancer, supporting the view that certain pathways of DNA repair and checkpoint control are required for preventing tumourigenesis in breast tissue. BRCA1-BARD1 containing supercomplexes and their functions are presented in Figure 6. (Greenberg et al. 2006, rev. in Greenberg 2008 and in Wu et al. 2008.)

Both BARD1 and BRCA1 are implicated in transcriptional regulation. BRCA1 associates with a number of transcriptional factors. For example, BRCA1 stimulates p53-dependent transcription of p21WAF1/CIP1 (Monteiro et al. 1996, Zhang et al. 1998, Wu et al. 2008). BARD1 has recently been reported to reduce the transcriptional activity of BRCA1 on the p21 and GADD45 promoters (Fabbro & Henderson 2008).
Fig. 6. BRCA1-BARD1 supercomplexes (modified from Wu et al. 2008). Proteins associated with familial breast cancer are indicated in gray. a) In the RNA polymerase II complex the BRCA1-BARD1 heterodimer may act as a transcriptional co-activator. In a model in which the heterodimer acts as a sensor for DNA damage, the heterodimer dissociates from Pol II after certain types of DNA damage. The interaction with CstF-50 may prevent inappropriate mRNA processing at sites of DNA repair. b) The complex containing BRCA1, BARD1, BACH1 and TOPBP1 responds to DNA damage during the S phase of the cell cycle. The interaction causes dissociation of TOPBP1 from the replication point. Interaction with the MSH2-MSH6-MLH1 complex indicates a role in DNA mismatch repair. c) BRCA1-BARD1 associates with the MRN complex and CtIP at sites of DNA damage, in a manner dependent on ATM or CHK2. This complex is required for activation of the G2/M checkpoint and CHK1. The complex also bridges the two DNA ends of DSB, an intermediate for DNA repair pathways. d) In a complex with BRCA2 and RAD51, BRCA1-BARD1 plays an important role in HR repair of DNA DSBs as well as in localisation of the RAD51 recombinase to damaged DNA sites. e) BRCA1-BARD1 interacts with the RAP80-ABRA1 complex, which may help to recruit it to the sites of DNA damage. The complex is also required for DNA damage resistance, G2/M checkpoint control and DNA repair.
The BRCA1-independent functions of BARD1 include its role in p53-dependent apoptotic signalling. BARD1 binds to p53, facilitating its phosphorylation and stabilisation (Irminger-Finger et al. 2001). BARD1 has also been reported to be able on its own to elicit p53-independent cell cycle arrest in G1 phase (Schuchner et al. 2005). Recently, BARD1 has been shown to interact with human papilloma virus (HPV) E6 protein which interacts with p53. These findings suggest that BARD1 may regulate the transcriptional activities of p53 (Yim et al. 2007).

BRCA1-independent function of BARD1 includes also its interaction with an NF-kB co-factor, B cell leukaemia/lymphoma 3 protein (BCL3), thereby possibly modulating the activity of transcription factor NF-kB (Dechend et al. 1999). The NF-kB family of transcription factors are involved in many physiological processes, including inflammation, apoptosis and the regulation of T- and B-cell response, and are overexpressed in many tumours (rev. in Baldwin 2001). BARD1 has also been shown to interact with the Ewing sarcoma gene product (EWS) and the oncogenic fusion protein EWS-FLI1 (Spahn et al. 2002).

**BARD1 in breast cancer predisposition**

Several possible cancer-associated BARD1 alterations have been reported in breast and ovarian cancer patients (Thai et al. 1998, Ghimenti et al. 2002, Ishitobi et al. 2003, Sauer & Andrulis 2005, Stacey et al. 2006, Gorringe et al. 2008). Germline BARD1 alteration have also been reported in a patient diagnosed with breast, ovarian and endometrial cancer, and somatic BARD1 mutations have been observed in breast and endometrial cancer (Thai et al. 1998). These findings have suggested that BARD1 might contribute to the development of human carcinomas.

The role of putative cancer-associated BARD1 alterations in breast and ovarian cancer is discussed in more detail in chapter 6.2. Aberrant BARD1 expression has been observed in some sporadic breast tumours (Yoshikawa et al. 2000), and elevated cytoplasmic expression of aberrant forms of BARD1 has been shown to associate with poor prognostic factors in breast and ovarian cancer but not in lung cancer (Wu et al. 2006). Expression of truncated BARD1 isoforms has also been observed to correlate with advanced stages of ovarian cancer, typically with clear cell carcinoma that is thought to have the worst prognosis of all of the epithelial ovarian cancers (Li et al. 2007). These observations suggest that BARD1 isoforms may play a role in cancer progression.
2.6.3 TOPBP1 is a checkpoint protein

TOPBP1 was originally identified as a protein interacting with DNA topoisomerase IIβ, and like its homologues in yeast (Schizosaccharomyces pombe Rad4/Cut5 and Saccharomyces cerevisiae Dpb11 proteins) and fly (Drosophila melanogaster Mus101 protein) it is involved in DNA damage and replication checkpoint pathways (Yamane et al. 1997, Mäkiniemi et al. 2001, Yamane et al. 2002, Garcia et al. 2005). As a striking structural feature TOPBP1 possesses eight BRCT domains, more than any other known protein (Fig. 7). Proteins with more than two BRCT repeats are known to offer the possibility of bringing together multiple interacting proteins, a situation that is necessary for DNA damage repair responses. Thus TOPBP1 is predicted to act as a scaffold protein coordinating DNA damage response (Garcia et al. 2005). Expression studies have shown that TOPBP1 is required for normal cell survival, since down-regulation of TOPBP1 results in reduced cell viability due to increased apoptosis (Yamane et al. 2002). Reduction of TOPBP1 levels has also been shown to inhibit the damage-inducible phosphorylation of CHK1 and other ATR substrates (Liu et al. 2006). In the absence of TOPBP1 cells have been shown to accumulate spontaneous DNA damage and activate the ATM-CHK2 pathway (Kim et al. 2005).

Fig. 7. Schematic structure of the TOPBP1 protein. The locations of the eight BRCT regions are shadowed, and are numbered above the diagram. The positions of the other known functional domains as well as the approximate interacting regions of the selected binding partners are shown as bars below the diagram. Amino acid positions for the missense variants observed in this study (IV) are marked above the schematic structure.
TOPBP1 associates with BRCA1 and has structural and functional similarities with it

The extended C-terminal part of TOPBP1 harbouring two of the eight BRCT domains shows considerable similarity to the corresponding part of BRCA1, even outside the BRCT regions (Yamane et al. 1997, Mäkiniemi et al. 2001). The expression of both BRCA1 and TOPBP1 is highest in S-phase during which the two proteins colocalise at sites adjacent to replication origins (Mäkiniemi et al. 2001). The localisation patterns of TOPBP1 and BRCA1 also have similarities during late mitosis as well as in meiotic prophase I (Reini et al. 2004). The similarities between TOPBP1 and BRCA1 are further emphasised by the recent observation that TOPBP1 was aberrantly expressed in a significant number of breast carcinomas (Going et al. 2007), suggesting a role for TOPBP1 in breast cancer development. Furthermore, TOPBP1 and BRCA1 interact with ubiquitin ligases EDD/HYD1 and BARD1, respectively, (Wu et al. 1996, Honda et al. 2002), both of which are aberrantly expressed in breast cancer (Thai et al. 1998, Clancy et al. 2003).

In intact cells TOPBP1 associates with ubiquitin ligase HYD1, and ubiquitination leads to degradation of the protein through the proteosome pathway (Honda et al. 2002). DNA damage-induced phosphorylation inhibits TOPBP1 ubiquitination, stabilising the protein, and stimulates its co-localisation with BRCA1 and several other molecules critical for double-strand break DNA repair, including the immediate DNA damage marker H2AX, PML, RAD50, ATM, RAD9, BLM, NBS1 and 53BP1 (Honda et al. 2002, Yamane et al. 2002, Xu et al. 2003). The complex containing TOPBP1, BRCA1-BARD1 and BACH1 responds to DNA damage during the S phase of the cell cycle (Fig. 6). It participates in replication origin and initiation related events critical for the activation of the S phase checkpoint (Mäkiniemi et al. 2001, Greenberg et al. 2006). Furthermore, TOPBP1 and BRCA1 have been shown to possess overlapping functions in G2/M checkpoint regulation (Yamane et al. 2003). There is also evidence that both proteins are involved in DNA double-strand break repair by HR, as TOPBP1 was recently reported to associate with NBS1 and to promote homologous recombination repair (Morishima et al. 2007). Moreover, TOPBP1 and BRCA1 are substrates for the ATM and ATR kinases without which there is no effective DNA damage response (Cortez et al. 1999, Tibbetts et al. 2000, Yamane et al. 2002). Both proteins are also required for activation of CHK1, a downstream substrate of ATR. The involvement of TOPBP1 in CHK1
activation is discussed in more detail in the next chapter. (Yamane et al. 2003, Kumagai et al. 2006, Yoo et al. 2007.)

Like BRCA1, also TOPBP1 acts as a transcriptional regulator. Both proteins interact with BRG1, a subunit of the SWI/SNF chromatin-remodelling complex (Bochar et al. 2000, Liu et al. 2004) and the human papilloma virus replication/transcription factor E2 (Boner & Morgan 2002, Kim et al. 2003). Interaction of TOPBP1 with the chromatin modification factor BRG1 results in repression of the transcriptional and apoptotic function of E2F1 both during normal growth and after DNA damage (Liu et al. 2004). TOPBP1 also interacts with the protein product of the proto-oncogene ABL1 that is frequently mutated in chronic myelogenous leukaemia and represses its expression (Zeng et al. 2005). In complex with MIZ1 TOPBP1 represses the promoter of oncoprotein MYC (Yoshida & Inoue 2004).

**TOPBP1 is required for both ATR- and ATM-dependent activation of CHK1**

After exposure to DNA-damaging agents or other cellular stresses eukaryotic cells respond by activating cell cycle checkpoints. These checkpoint mechanisms maintain genomic integrity by coordinating cell cycle progression as well as other processes including DNA replication, transcription, apoptosis, senescence and DNA repair. Failure of checkpoint responses leads to DNA damage, genomic instability and tumour development. (rev. in Zhou & Elledge 2000.) The significance of checkpoints to human health is illustrated by a growing list of genes involved, such as ATM, BRCA1, TP53, NBS1 and CHEK2 that are mutated in cancer and cancer-predisposing syndromes (Malkin et al. 1990, Miki et al. 1994, Savitsky et al. 1995, Varon et al. 1998, Meijers-Heijboer et al. 2002). TOPBP1 plays an important role in recruiting key elements of the checkpoint signalling machinery for cell cycle arrest and DNA damage repair. Recently, TOPBP1 has been shown to play an important role as a general activator of CHK1 (Kumagai et al. 2006, Yoo et al. 2007).

Two members of the phosphatidylinositol-3-kinase-related (PIKK) family of proteins, ATM and ATR, are key DNA damage sensors and signal transducers in the checkpoint responses. Although ATM and ATR share common downstream substrates such as p53 and BRCA1 they primarily respond to different stimuli. ATM mainly responds to DNA double-strand breaks, while ATR is primarily activated by replication stress and single stranded (ss) DNA gaps, although it also
participates in double-stranded DNA damage responses. After activation, ATM phosphorylates a wide array of DSB repair and genome surveillance factors including Artemis, NBS1, H2AX, BRCA1, FANCD2, p53, MDC1, BLM and CHK2. Thus, ATM orchestrates the DSB response by phosphorylating substrates required for the G1/S, intra-S and G2/M checkpoints. A key function of ATR involves activation of the downstream checkpoint effector kinase CHK1, which promotes cell survival by blocking the firing of replication origins, preventing entry into mitosis, stabilising stalled replication forks and facilitating DNA repair. ATM has been shown to be required for the activation of CHK1 by ATR in response to DNA DSBs but not to replication stress. (Kumagai et al. 2006, Yoo et al. 2007, rev. in Thompson & Schild 2002, Lavin & Kozlov 2007 and Cimprich & Cortez 2008.)

The complex pathway leading to the activation of CHK1 includes several checkpoint proteins that are recruited to ssDNA in response to replication stress. ATR is recruited to ssDNA via its stable partner ATRIP (ATR-interacting protein), which binds to Replication Protein A (RPA)-coated ssDNA. A second complex, the clamp loader consisting of RAD17 and small subunits 2-5 of the replication factor C (RFC2-5), is independently recruited to ssDNA. Activation of this complex then loads the RAD9-RAD1-HUS1 (9-1-1) clamp complex. (rev. in Cimprich & Cortez 2008 and in Sancar et al. 2004.) The interaction of 9-1-1 complex with TOPBP1 is mediated through RAD9 and is necessary for the activation of ATR. Thus, interaction of RAD9 with TOPBP1 is critical for the initiation of checkpoint signalling. TOPBP1 serves as an activator of the ATR-ATRIP complex and functions as a link between the independently recruited 9-1-1 and ATR-ATRIP complexes. The ATR activation domain of TOPBP1 locates between the sixth and seventh BRCT repeat (Fig. 7). (Kumagai et al. 2006, Burrows & Elledge 2008.) Once activated, ATR-ATRIP phosphorylates its effector kinase CHK1 in addition to its other downstream effectors. The final activating phosphorylation of CHK1 requires the assistance of the mediator protein Claspin. Subsequently, CHK1 phosphorylates key cell-cycle enzymes such as CDC25 and WEE1. (rev. in Sancar et al. 2004.) Recently, TOPBP1 has also been shown to be critical for the ATM-dependent activation of ATR and CHK1 in response to double-stranded DNA breaks (Yoo et al. 2007).
3 Aims of the study

A significant proportion of familial breast cancer cannot be explained by mutations in the two main susceptibility genes \textit{BRCA1} and \textit{BRCA2}, or in the other genes known to be associated with the disease predisposition. Therefore, additional predisposing genes are likely to be discovered. The protein products of the currently known susceptibility genes are involved in DNA damage response and DNA repair, and many of them associate with BRCA1 or BRCA2. Consequently, other genes coding for proteins in these pathways may also be involved in cancer development. In this study a case-control design was used to screen putative candidate genes. The study focused on investigating three genes (\textit{BACH1}, \textit{BARD1} and \textit{TOPBP1}) that all encode proteins involved in the same cellular pathways as BRCA1.

The specific aims of the present study were:

1. To analyse whether germline alterations in the \textit{BACH1} gene that encodes a BRCA1-interacting helicase contribute to breast cancer risk in Finland.
2. To screen the \textit{BARD1} gene encoding the main binding partner of BRCA1 for germline alterations in Finnish patients with family history of breast and/or ovarian cancer and to determine whether possible aberrations are associated with an increased susceptibility to breast cancer.
3. To specifically investigate the prevalence of the \textit{BARD1} Cys557Ser alteration in a large cohort of breast and/or ovarian cancer patients not only from Finland, but also from the other Nordic countries, and to determine whether the variant is associated with other types of cancer.
4. To analyse BRCA1-associated checkpoint protein encoding gene \textit{TOPBP1} for germline alterations among Finnish breast and ovarian cancer patients to assess their possible role in cancer development.
4 Materials and methods

4.1 Subjects (I, II, III, IV)

The total numbers of cancer cases and healthy control individuals in each study (I, II, III, IV) are summarised in Table 2.

Table 2. Summary of subjects included in studies I-IV.

<table>
<thead>
<tr>
<th>Study material</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study material</td>
<td>Study I</td>
</tr>
<tr>
<td>Analysis of</td>
<td>Analysis of</td>
</tr>
<tr>
<td>BACH1</td>
<td>BARD1</td>
</tr>
<tr>
<td>Cys557Ser</td>
<td></td>
</tr>
<tr>
<td>Breast and ovarian cancer samples</td>
<td></td>
</tr>
<tr>
<td>Breast and/or ovarian cancer families</td>
<td>151</td>
</tr>
<tr>
<td>Breast cancer patients without known family history of the disease</td>
<td>-</td>
</tr>
<tr>
<td>Unselected breast cancer patients</td>
<td>-</td>
</tr>
<tr>
<td>Male breast cancer patients</td>
<td>-</td>
</tr>
<tr>
<td>Prostate cancer samples</td>
<td></td>
</tr>
<tr>
<td>Prostate cancer families</td>
<td>-</td>
</tr>
<tr>
<td>Unselected prostate cancer patients</td>
<td>-</td>
</tr>
<tr>
<td>Colorectal cancer families</td>
<td>-</td>
</tr>
<tr>
<td>Healthy control individuals</td>
<td>304</td>
</tr>
<tr>
<td>Total number of samples</td>
<td>455</td>
</tr>
</tbody>
</table>

4.1.1 Breast and/or ovarian cancer patients (I, II, III, IV)

Patients with family history of breast and/or ovarian cancer (I, II, III, IV)

Altogether 151 Finnish breast and/or ovarian cancer families were included in the mutation analysis of the BACH1 gene (I) (Table 2). In all, the entire BACH1 gene was analysed in 214 breast or ovarian cancer patients from these 151 families.

The 151 families included in the study originated from Northern Finland and samples were collected at Oulu University Hospital during the years 1993–1999. Oulu University Hospital collects patients from an extensive geographical region covering a large area of Northern Finland. The families were included on the basis of affected family members, types of cancer present in the family and the
of the families studied, 114 were associated with breast cancer, 29 with breast and ovarian cancer and eight with ovarian cancer (Table 3). Based on the following criteria, the families were subdivided into 76 high, 52 moderate and 23 low cancer risk families (Table 3):

1. High-risk families had three or more cases of breast and/or ovarian cancer diagnosed in the same or adjacent generations, or two cases if at least one of them showed a high risk feature including early disease onset (≤35 years), bilateral breast cancer or multiple primary tumours. Most of the high-risk families displayed three or more cancer cases. Previously, all of the high risk families had been screened for germline mutations in the \textit{BRCA1}, \textit{BRCA2} \cite{Huusko1999}, \textit{CHEK2} \cite{Allinen2001} and \textit{TP53} \cite{Rapakko2001} genes, and 11 families were shown to be \textit{BRCA1} or \textit{BRCA2} mutation positive.

2. Moderate risk families had two cases of breast and/or ovarian cancer in first- or second-degree relatives.

3. Low-risk families had a single case of breast or ovarian cancer diagnosed along with multiple cases of other types of cancer. Four of the patients from 23 low-risk families were diagnosed with ovarian cancer and 19 with breast cancer.

Of the high and moderate cancer risk families described above, 126 and 125 families were selected for the screening of germline alterations in the \textit{BARD1} and \textit{TOPBP1} genes, respectively (II, IV) (Tables 2 and 3). Testing was initially carried out on DNA sample from an index case (defined as the youngest available breast or ovarian cancer patient in the family) and upon detection of a possible disease-related alteration the rest of the family members were tested for the corresponding alteration.

The prevalence of the \textit{BARD1} Cys557Ser allele was analysed in 922 index cases from Nordic breast and/or ovarian cancer families (III) (Tables 2 and 3). Of these patients, 205 originated from Finland, 126 of which were also included in study II, while the rest were from Southern Finland and had been collected at Tampere University Hospital; 157 patients originated from Iceland, 219 from Denmark, 176 from Sweden and 165 from Norway. All index cases were diagnosed with breast or ovarian cancer. Inclusion criteria for the patients were two or more cases of breast and/or ovarian cancer in first- or second-degree relatives. Most of the families displayed three or more cancer cases. Based on mutation screening or haplotyping (Iceland), the division into \textit{BRCA1}/\textit{BRCA2}-
associated (n=250) and non-associated (n=593) subgroups was possible for 843 out of the 922 families studied; for the remaining families the \textit{BRCA1}/\textit{BRCA2} mutation status was unknown. In addition, five of the 119 Swedish and two of the 104 Icelandic familial cases had been shown to be positive for \textit{CHEK2} 1100delC. Of the Nordic families, 494 were known to be associated with breast, 202 with breast and ovarian, and 10 with ovarian cancer (Table 3).

**Table 3. Familial breast and/or ovarian cancer patients analysed in studies I-IV as subdivided by their family history of and risk of breast and/or ovarian cancer.**

<table>
<thead>
<tr>
<th>Study material</th>
<th>Number of studied families</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study I</strong></td>
<td></td>
</tr>
<tr>
<td>Analysis of \textit{BACH1}</td>
<td>Study I</td>
</tr>
<tr>
<td>Families as determined by their family history of breast and/or ovarian cancer</td>
<td></td>
</tr>
<tr>
<td>Families with breast cancer</td>
<td>114</td>
</tr>
<tr>
<td>Families with breast and ovarian cancer</td>
<td>29</td>
</tr>
<tr>
<td>Families with ovarian cancer</td>
<td>8</td>
</tr>
<tr>
<td><strong>Study II</strong></td>
<td>Analysis of \textit{BARD1}</td>
</tr>
<tr>
<td>Families as determined by their risk of breast and/or ovarian cancer</td>
<td></td>
</tr>
<tr>
<td>High-risk families</td>
<td>76</td>
</tr>
<tr>
<td>Moderate-risk families</td>
<td>52</td>
</tr>
<tr>
<td>Low-risk families</td>
<td>23</td>
</tr>
</tbody>
</table>

\(^1\) Includes only samples for which the particular classification is known. \(^2\) One family with history of breast cancer was affected with ovarian cancer during the project and was reclassified for study IV, NA not available.

**Breast cancer patients without known family history of the disease (II, IV)**

A total of 188 cancer patients without known family history of the disease (sporadic cases) were included for the screening of possible disease-related alterations (II, IV) (Table 2). The occurrence of \textit{BARD1} Cys557Ser was tested in all of the sporadic cases (II) and \textit{TOPBP1} Arg309Cys in 187 of the cases (IV). The exclusion of family history of breast and ovarian cancer of each patient was performed through a personal questionnaire. Samples originated from the same geographical region as the Northern Finnish breast and/or ovarian cancer families and were collected at Oulu University Hospital during the years 1993–1999.
Breast cancer patients unselected for family history (III)

Altogether 1984 DNA samples from a consecutive series of newly diagnosed breast cancer patients from Finland (616 samples), Iceland (1142 samples) and Denmark (226 samples) were analysed for the prevalence of the BARD1 Cys557Ser allele in study III (Table 2). These unselected cohorts were recruited without selection for or against family history of breast cancer. Blood samples from the Finnish breast cancer patients were collected between the years 2000–2004 at Oulu University Hospital (317 samples; the mean age at diagnosis was 58 years, ranging from 29 to 95 years) or 1986–1994 at Tampere University Hospital. Oulu University Hospital gathers patients from a large area comprising Northern Finland, while patients collected at the Tampere University Hospital originate from Southern Finland. The Icelandic breast cancer patients were recruited at the University Hospital of Reykjavik and through the cancer registry during the years 1987–2004 (the mean age at diagnosis was 58 years, ranging from 18 to 98 years). The Danish breast cancer cases were diagnosed during the years 1994 and 1995 at Odense University Hospital.

Male breast cancer patients (III)

The prevalence of the BARD1 Cys557Ser allele was analysed in a population-based collection of 128 Finnish male breast cancer patients (III) (Table 2). The samples were collected at Tampere University Hospital from all available patients diagnosed in Finland between 1967 and 1996 and represent about 50% of the male breast cancer cases diagnosed during those 30 years in Finland. The mean age at diagnosis was 65.4 years (ranging from 30 to 94 years). All of the male breast cancer patients had previously been screened for Finnish BRCA2 founder mutations (Syrjäkoski et al. 2004b) and 109 of the samples for CHEK2 1100delC (Syrjäkoski et al. 2004a). Ten patients had BRCA2 mutation and one carried CHEK2 1100delC. Eleven male breast cancer patients had a first-degree relative diagnosed with breast cancer, one with ovarian cancer and two patients had both breast and ovarian cancer in the family.
4.1.2 Prostate and colorectal cancer patients (III)

Altogether 613 unselected prostate cancer patients as well as 121 prostate cancer families and 188 colorectal cancer families from Finland were included in study III in order to analyse the prevalence of the BARD1 Cys557Ser allele (Table 2).

A population-based cohort of 613 unselected consecutive Finnish prostate cancer patients was collected at Tampere University Hospital. These patients were diagnosed with cancer during 2000–2001 and the mean age at diagnosis was 69.5 years (ranging from 45 to 93 years). Prostate cancer families included two or more first- or second-degree relatives affected with prostate cancer. Samples from the youngest affected patient available in each of the 121 families were used for the analysis and the mean age at cancer diagnosis for the index patients was 64.8 years (ranging from 44 to 86). CHEK2 1100delC had been analysed in all of the familial prostate cancer patients and in 298 of the unselected prostate cancer patients. Four familial cases and five unselected cases carried CHEK2 1100delC.

Inclusion criteria for the colorectal cancer families were malignancy in two or more first-degree relatives. The mean age at colorectal cancer diagnosis for the analysed index patients was 66.1 years (ranging from 27 to 90 years). All of the familial colorectal cancer patients had previously been studied for the occurrence of CHEK2 1100delC and for microsatellite instability (MSI), and individuals positive for MSI had been screened for mutations in MLH1 and MSH2 genes. One patient carried CHEK2 1100delC and one had MLH1 splice-site mutation.

4.1.3 Healthy control individuals (I, II, III, IV)

The 304 anonymous Finnish control individuals analysed for the BACH1 gene (I) originated from the same geographical region as the familial breast cancer cases and were collected during 1998–2000 at Oulu University Hospital (Table 2).

A total of 1,018 and 697 individuals served as geographically matched Finnish population controls to verify the frequency of the variants alleles in the BARD1 and TOPBP1 genes in studies II and IV, respectively (Table 2). DNA samples from these 1,018 anonymous Finnish voluntary cancer-free blood donors were obtained from the Finnish Red Cross Blood Transfusion Service in Oulu representing the northern area of Finland. Samples were collected in 2002 and the mean age was 41 years (ranging from 18 to 66 years).

In study III, population-based cohorts of altogether 3,591 anonymous cancer-free individuals were used to assess the frequency of BARD1 Cys557Ser in
healthy Nordic populations (Table 2). DNA from healthy individuals originated from Finland (1,992 samples), Iceland (931 samples), Denmark (371 samples) and Sweden (297 samples). Finnish controls consisted of DNA samples from anonymous, voluntary and cancer-free blood donors obtained from the Finnish Red Cross Blood Centre in Tampere (358 samples), Turku (352 samples), Kuopio (264 samples) and Oulu (1,018 samples which were also used in study II) representing central, west-southern, eastern and northern areas in Finland, respectively. Finnish male controls consisted of 440 DNA samples from male blood donors (age ranging from 18 to 65 years) obtained from the Finnish Red Cross Blood Centre in Tampere (176 samples), Turku (176 samples) and Kuopio (88 samples). The Icelandic control group was a combination of three randomly selected sample sets consisting of 347 individuals participating in an Icelandic National Diet Survey in 1990 (age ranging from 15 to 74 years), 479 individuals obtained from the Icelandic genealogical database during 2003–2004 (age ranging from 18 to 75 years) and 105 individuals collected through the Icelandic Blood Bank and the Icelandic Heart Association during the years 1988–1990 (age ranging from 20 to 77 years). Danish controls were cancer-free and anonymous blood donors collected during the years 1994–1995 and in 2002 at Odense University Hospital. The Swedish control group contained military conscripts aged approximately 18 years, representing a geographically matched unselected male population.

4.2 DNA extraction (I, II, III, IV)

Genomic DNA from peripheral blood lymphocytes was extracted either by a standard phenol-chloroform method (Sambrook et al. 1989) or using the Puregene D-50K purification kit (Gentra) (I, II, III, IV).

4.3 Mutation analysis (I, II, III, IV)

4.3.1 Conformation sensitive gel electrophoresis (I, II, III, IV)

Mutations in the protein encoding regions and exon-intron boundaries of the \textit{BACH1}, \textit{BARD1} and \textit{TOPBP1} genes were searched for by conformation sensitive gel electrophoresis (CSGE) (I, II, III) (Ganguly et al. 1993, Körkkö et al. 1998). In addition, CSGE was used in study IV to detect the \textit{BARD1} Cys557Ser
alteration in Finnish and part of the Icelandic breast and/or ovarian cancer patients and controls as well as in colorectal cancer patients. CSGE has been developed for heteroduplex analysis to screen large genes for the presence of single-base and larger mismatches. The method is based on the theory that mildly denaturing conditions emphasise the conformational changes produced by mismatches in the DNA double-helix leading to increase in the differential electrophoretic migration of heteroduplex and homoduplex molecules.

AmpliTaqGold DNA polymerase (Applied Biosystems) was used for amplification of genomic DNA, except for the exon 3 of the BARD1 gene that was amplified using Vent high fidelity DNA polymerase (New England BioLabs) due to a long poly(T) repeat in the sequence. After PCR amplification, PCR products were denaturated at 98°C for 5 minutes after which formation of DNA heteroduplexes was allowed at 68°C for 30 minutes. Samples were electrophoresed using mildly denaturing CSGE-gels (25% acrylamide, 10% ethylene glycol, 15% formamide) at 400 V for 20–24 hours depending on the migration of different fragments. Gels were stained with ethidium bromide and results were visualised with a UV illuminator.

Oligonucleotides were designed using the Primer3 software (http://frodo.wi.mit.edu/) utilising sequence data obtained from the publicly available databases maintained by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/). The following sequence data were used: for TOPBP1 NT_022393 and NM_007027.2; for BACH1 AC060798, NT_010740 and NM_032043; for BARD1 AH005849, NT_005486 and NM_000465.

4.3.2 Minisequencing (III)

Solid-phase minisequencing (Syvänen et al. 1993, Syvänen 1998) was used to assess the frequency of the BARD1 Cys557Ser alteration in male breast cancer patients, prostate cancer patients and male controls (III). The method allows accurate and sensitive quantitation of two sequences that differ from each other by one nucleotide. The solid-phase minisequencing is based on extension of a detection primer that anneals immediately adjacent to a variable nucleotide position of an affinity-captured amplified template, using a DNA polymerase and a single, labelled nucleoside triphosphate.

Genomic DNA was amplified with one normal and one biotinylated primer. The PCR product was then attached to the streptavidin-coated microtitre well with biotin-streptavidin bond. After the wash the other strand was removed by
alkaline denaturation. Detection primers were applied to the wells with $^3$H-labelled nucleotides, and after the incubation and wash the counts were measured by beta-counter. Two wells were used for each sample, one for recognising normal and the other for recognising mutated nucleotide. The primers used for PCR were: forward 5'-CTG CGG CCT GTC GAT TAT ACA G-3' and reverse 5'-GGT GAA TGT AAG CAA TTG GTC-3' (5'-biotin). The minisequencing primer was 5'-GAA TCA TCC TCA GCT AGC CAC T-3'.

4.3.3 TaqMan assay (III)

For the detection of the BARD1 Cys557Ser variant in Danish samples in study III, a TaqMan SNP (single nucleotide polymorphism) genotyping assay (TaqMan® MGB, Applied Biosystems) was employed. This hybridisation-based allelic discrimination assay utilises the 5' nuclease activity of Taq polymerase in conjunction with fluorogenic TaqMan probes to detect a fluorescent reporter generated during or after PCR reactions.

One pair of TaqMan probes and one pair of PCR primers were used for genotyping. DNA was amplified with the use of the following primers: forward 5'-TAT TGC TGC TAC CAG AGA AGA ATG A-3' and reverse 5'-ATG TCCCAA AGC TAA ATC CAT ACT TAC-3'. The genotypes were determined using the following fluorogenic allele-specific probes: 6FAM-5'-CTA GCC ACT GCT CAG-3'-MGB for the wild-type allele and VIC-5'-CTA GCC ACT CCT CAG-3'-MGB for the variant allele. Following amplification, samples were analysed on an ABI Prism 7700 Sequence Detection System. End point results were acquired using Sequence Detection System software (Applied Biosystems).

4.3.4 Denaturing high-performance liquid chromatography (III)

Swedish and Norwegian samples were analysed for the BARD1 Cys557Ser allele by denaturing high-performance liquid chromatography (DHPLC) (III). The method is based on separating heteroduplex molecules from homoduplexes by ion-pair reverse-phase liquid chromatography according to their differences in melting behaviour.

A fragment of 272 bp including BARD1 exon 7 and codon 557 was amplified from genomic DNA by a PCR protocol enabling heteroduplex formation using the following primers: forward 5'-GGA TGT TGA AAA ATT GAG TCG AG-3' and reverse: 5'-AAG CAA TTG GTC AAA TGG AAA-3'. DHPLC analysis was
performed on a WAVE DNA Fragment Analysis System (Transgenomic). During the analysis, partially heat-denatured double-stranded DNA fragments interact in a length- and sequence-specific manner with a non-porous poly(styrene-divinylbenzene) matrix followed by elution of the DNA from the matrix by a linear acetonitrile gradient. The optimal temperatures and gradient conditions were determined using WAVEMAKER™ Software (Transgenomic).

4.3.5 Restriction fragment length polymorphism analysis (IV)

In study IV, PCR followed by restriction fragment analysis was utilised for detection of individuals homozygous for the variant Cys309 allele of the \textit{TOPBP1} gene. The identification of wild-type (C/C) and variant (T/T) homozygotes at position 1010 was based on amplification of a 383-bp DNA fragment followed by digestion with \textit{TaI} enzyme (Fermentas).

Genomic DNA was amplified using the following primers: forward 5'-CAC CCC TGA ATC TTA CTT TGG A-3' and reverse 5'-GAG CAA GCC CAA CTC TGA AG-3'. After amplification the products were digested with \textit{TaI} according to conditions recommended by the manufacturer. In the wild-type allele, the restriction site is absent and no digestion occurs, while in the variant homozygote, the site is present and the PCR product is digested into two fragments (251 bp and 132 bp). In heterozygotes (309 C/T) all three possible fragments appear. Positive control was included in each analysis to validate the results. Digestion products were resolved on 3% ethidium bromide –stained MetaPhor agarose gels (Cambrex) and visualised under UV illumination.

4.3.6 Direct sequencing (I, II, III, IV)

Direct sequencing was used to assess the frequency of \textit{BARD1} Cys557Ser in part of the Icelandic breast and/or ovarian cancer cases and controls (III) as well as to confirm all positive results from studies I-IV identified by different mutation screening methods.

Direct sequencing of independently amplified and purified templates was performed according to the manufacturer’s instructions. Finnish samples were analysed with the Li-Cor IR\textsuperscript{2} 4200-S DNA Analysis system (Li-Cor) using the SequiTherm EXCEL\textsuperscript{TMII} DNA Sequencing Kit-LC (Epicentre Technologies) or ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Danish samples were analysed using an ABI377 automated sequencer, Swedish and Norwegian samples
with an ABI Prizm® 3100 Genetic Analyzer and Icelandic samples with an ABI 3730 capillary sequencer (Applied Biosystems).

4.4 Cell culture (IV)

Two wild-type and four TOPBP1 Arg309Cys heterozygous Epstein Barr Virus (EBV) transformed lymphoblastoid cell lines (LCLs) were derived from patients diagnosed with breast cancer. Cells were cultured in RPMI 1640 medium (Gibco) supplemented with 20% foetal bovine serum, 1% L-glutamine and gentamycin (10 μg/ml) at 37°C in a 5% CO₂/95% air atmosphere. The cells were centrifuged, washed twice with PBS and the pellet was suspended with 2.5 x Laemmli buffer and sonicated briefly. The protein concentrations were measured using the RC DC protein assay kit (BioRad).

4.5 Western blot analysis (IV)

Samples containing 25 μg of protein were separated by SDS-PAGE (10% gels), after which the gels were stained with Sypro Orange (BioRad) to verify equal loading and electroblotted onto PVDF membranes (Millipore). After blocking with 5% non-fat dry milk in tris-buffered saline (TBS) containing 0.1% Tween-20 for 1 h, membranes were probed with the rabbit polyclonal anti-TOPBP1.2 antibody (1.5 μg/ml) (Utio et al. 1995, Mäkiniemi et al. 2001) and further with horseradish peroxidase-conjugated goat anti-rabbit antibody (Jackson ImmunoResearch Laboratories). Secondary antibody was used at dilutions of 1:10000 to 1:15000. The signals were detected by chemiluminescence, using the SuperSignal West Femto detection kit (Pierce). ImageQuant 5.2 (Amersham Lifesciences) was used for quantification of densitometrically scanned Western blot signals. Analyses were performed in triplicate.

4.6 Statistical analyses (I, II, III, IV)

Statistical significance of the observed differences in mutation frequencies between cancer patients and control individuals as well as between different groups of breast and ovarian cancer patients were evaluated by Pearson’s Chi-Square or Fisher’s exact test (I, II, III, IV). Mann-Whitney’s U-test was used to compare mean ages of disease onset between alteration carriers and non-carriers (II, III). Analyses were carried out with SPSS versions 12.0 and 14.0 (SPSS Inc.)
for Windows. In study III, the Mantel-Haenszel method was utilised to estimate further the pooled odds ratios for the combined BARD1 Cys557Ser frequencies from different countries. Data were analysed using STATA version 8.0 (StataCorp LP) for Windows. All p-values were two-sided and p-values < 0.01 were considered statistically significant as suggested by Houlston and Peto (2003) for the association studies for potential low-penetrance alleles. Confidence Interval Analysis (CIA) version 2.0.0.41 for Windows was used for calculating confidence interval values for particular control groups in study III.

4.7 Ethical issues

All participating individuals have given an informed consent for using their pedigree data and blood specimens for a study on cancer susceptibility gene mutations. Family members were only contacted with the permission of the index case. Each DNA sample was labelled with a research code and samples were handled anonymously during the analyses. Results from the analyses were only used for research purposes and no information was given to the participating individuals.

Appropriate research permissions to perform the study have been obtained from the ethical committees of each country and the participating University Hospitals. The study has been performed under approval from the Ministry of Social Affairs and Health in Finland.
5 Results

5.1 BACH1 germline alterations are rare among Finnish families with breast and/or ovarian cancer (I)

In total, 214 patients from 151 Finnish breast and/or ovarian cancer families were analysed for germline alterations in the whole coding region and splice site boundaries of the BACH1 gene. The analysis revealed altogether nine germline alterations, six of which were novel. Five of the alterations found were intronic and four were located within the protein-encoding region (Table 4). The locations of the identified coding region alterations are shown in Figure 3.

All the intronic sequence variants found in the present study were novel, i.e. previously unreported. On the basis of the likely absence of effect on protein function, they were considered not to associate with risk of breast or ovarian cancer and were therefore not additionally investigated.

Two of the coding region variants resulted in amino acid substitutions and two were silent (Table 4). Three of these alterations, two silent (Glu879Glu and Tyr1137Tyr) as well as one missense (Pro919Ser), had previously been reported as polymorphisms (Cantor et al. 2001, Luo et al. 2002). As these variants were found at high frequencies among the 151 index cases studied (9.3%, 50% and 47.7%, respectively), and in addition, the previously reported allele frequencies were similar to those obtained in the present study, their prevalence was not determined among controls. The novel missense variant 3101C>T resulted in a substitution of proline to leucine at codon 1034 (Pro1034Leu) and was present in only one patient (1/214) among the breast and ovarian cancer cases studied, giving a frequency of 0.66% (1/151) among the families studied. The role of this missense alteration was additionally assessed by examining its segregation with cancer and by determining its prevalence among 304 anonymous blood donors.

The novel BACH1 Pro1034Leu alteration was found in one family having a history of different cancers. In addition to the index case diagnosed with ovarian cancer at age 37, the variant was also seen in her maternal half-sister diagnosed with colorectal cancer at age 58. The family also revealed a strong history of other gynaecological cancers; the sister of the proband had been diagnosed with squamous cervical uterine carcinoma in situ at age 38, the mother with squamous cervical uterine carcinoma at unknown age and the maternal aunt with uterine cancer at age 28. Unfortunately, DNA was not available from any other members...
of this family to test for co-segregation of the alteration with cancer. In addition, due to the low number of ovarian cancer cases studied, the possible association between the variant and ovarian cancer could not be determined.

Table 4. Sequence alterations in the protein encoding regions of the BACH1 gene among Finnish families with breast and/or ovarian cancer.

<table>
<thead>
<tr>
<th>Nucleotide change</th>
<th>Effect on protein</th>
<th>Frequency in families¹ (n=151)</th>
<th>Frequency in controls¹ (n=304)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putative disease-associated alteration²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3101C&gt;T</td>
<td>Pro1034Leu</td>
<td>0.66% (1/151)</td>
<td>0.33% (1/304)</td>
<td>Not observed in other studies.</td>
</tr>
<tr>
<td>Polymorphisms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ For heterozygotes, ² Based on the effect on protein and on the difference in allele prevalence between patients and controls, ND not determined, n number of samples.

The Pro1034Leu alteration was also found in one of the 304 anonymous blood donors (0.3%) whose cancer history could not be obtained. Although the frequency of 0.7% (1/151) among the 151 index patients was over two-fold higher than in controls, the difference did not reach statistical significance (p=0.6). It should be noted, however, that the alteration seems to be very rare and the low number of samples studied does not give sufficient power for statistical analyses.
5.2 Mutation screening of the *BARD1* gene (II, III)

5.2.1 Prevalence of *BARD1* germline alterations among Finnish families with breast and/or ovarian cancer (II)

The mutation analysis of the *BARD1* gene in a total of 126 index cases from Finnish breast and/or ovarian cancer families revealed altogether 17 germline alterations (Table 5). Seven of the changes located in the coding region and ten were intronic. Four of the coding region alterations were missense and three were silent substitutions (Table 5, Fig. 5). The silent Gly203Gly was the only novel coding region alteration observed; the others have previously been identified in other studies (Thai *et al.* 1998, Ghimenti *et al.* 2002, Ishitobi *et al.* 2003) or have been described in the SNP database. Of the intronic alterations (unpublished data), eight were novel.

Of the genomic *BARD1* changes observed, Cys557Ser occurred more frequently among familial cancer cases than healthy controls and appeared to be associated with cancer predisposition ($p=0.005$). The other six coding region variants were considered neutral polymorphisms, as most of them also displayed similar allele frequencies in controls, occurred in the SNP database, or had been reported not to be disease-associated in other studies (Thai *et al.* 1998, Ghimenti *et al.* 2002, Ishitobi *et al.* 2003). In addition, all intronic changes were considered to be apparently harmless (Table 5).
Table 5. Germline alterations in the BARD1 gene among Finnish breast and/or ovarian cancer cases and healthy controls.

<table>
<thead>
<tr>
<th>Location</th>
<th>Nucleotide change</th>
<th>Effect on protein</th>
<th>Frequency in familial cases</th>
<th>Frequency in sporadic cases</th>
<th>Frequency in controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon region variants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Putative disease-associated alteration² Exon 7</td>
<td>1743G&gt;C</td>
<td>Cys557Ser</td>
<td>5.6% (7/126)</td>
<td>2.7% (5/188)</td>
<td>1.4% (14/1018)</td>
</tr>
<tr>
<td>Polymorphisms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 4</td>
<td>682A&gt;C</td>
<td>Gly203Gly</td>
<td>0.8% (1/126)</td>
<td>ND</td>
<td>0% (0/350)</td>
</tr>
<tr>
<td>Exon 4</td>
<td>1126G&gt;C</td>
<td>Thr351Thr</td>
<td>21% (27/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Exon 4</td>
<td>1207C&gt;G</td>
<td>Ser378Arg</td>
<td>48% (61/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Exon 6</td>
<td>1591C&gt;T</td>
<td>His506His</td>
<td>9.5% (12/126)</td>
<td>ND</td>
<td>16% (24/150)</td>
</tr>
<tr>
<td>Exon 6</td>
<td>1592G&gt;A</td>
<td>Val507M et</td>
<td>56% (70/126)</td>
<td>ND</td>
<td>47% (70/150)</td>
</tr>
<tr>
<td>Exon 10</td>
<td>2045C&gt;T</td>
<td>Arg658Cys</td>
<td>3.2% (4/126)</td>
<td>ND</td>
<td>3.0% (9/300)</td>
</tr>
<tr>
<td>Intron region variants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IVS1+46</td>
<td>A&gt;C</td>
<td>None</td>
<td>28% (35/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS2+59-60</td>
<td>delAT</td>
<td>None</td>
<td>8.7% (11/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS4+29</td>
<td>C&gt;G</td>
<td>None</td>
<td>7.9% (10/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS5-19</td>
<td>A&gt;G</td>
<td>None</td>
<td>46% (58/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS8-20</td>
<td>A&gt;G</td>
<td>None</td>
<td>0.8% (1/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS8-60</td>
<td>T&gt;C</td>
<td>None</td>
<td>2.4% (3/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS9-69a</td>
<td>T&gt;C</td>
<td>None</td>
<td>2.4% (3/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS9-77a</td>
<td>A&gt;G</td>
<td>None</td>
<td>2.4% (3/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS9+98</td>
<td>C&gt;T</td>
<td>None</td>
<td>0.8% (1/126)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>IVS10-82-85</td>
<td>delTTCA</td>
<td>None</td>
<td>3.2% (4/126)</td>
<td>ND</td>
<td>3.0% (9/300)</td>
</tr>
</tbody>
</table>

¹ For heterozygotes, ² Based on the effect on protein and on the difference in allele prevalence between patients and controls, ³ Always observed together, ND not determined, IVS intervening sequence.

5.2.2 Investigation of the BARD1 Cys557Ser alteration (II, III)

BARD1 Cys557Ser among Finnish families with breast and/or ovarian cancer (II)

In the preliminary screening in study II, 5.6% (7/126) of the index cases from the Finnish breast and/or ovarian cancer families studied were found heterozygous for the Cys557Ser allele (Table 5). The alteration was also seen in 1.4% of the cancer-free controls (n=1018), but with a significantly lower incidence (p=0.005, OR 4.2, 95% CI 1.7–10.7). The prevalence among breast cancer cases who had no known family history of the disease was 2.7% (5/188), the frequency being
slightly, but not significantly higher than among control subjects (*p*=0.2, OR 2.0, 95% CI 0.7–5.5).

All seven index patients carrying Cys557Ser belong to a subgroup of 94 families displaying breast but not ovarian cancer, giving a prevalence of 7.4% among these families (*p*=0.001, OR 5.8, 95% CI 2.3–14.7). Furthermore, all index patients exhibiting the Cys557Ser allele were from *BRCA1/BRCA2* mutation-negative families. To investigate whether the Cys557Ser segregates with breast and/or ovarian cancer in the seven families in which it was observed, the mutation status of other available affected or unaffected family members was analysed. In one of the families, the Cys557Ser allele was seen in a woman with breast cancer at age 75, and in her two daughters diagnosed with breast cancer at ages 49 and 53. The third daughter did not show the missense alteration and was cancer-free at age 63. In addition, the sister of the proband developed uterine cancer at the age of 40, but no DNA was available for the testing of mutation status. Unfortunately, in the remaining six families displaying the Cys557Ser alteration, only a small number of additional DNA samples were available for analysis. Thus, despite a history of breast cancer, and in part of the families other cancers as well, the segregation of Cys557Ser with the disease could not be conclusively determined.

In the 126 families studied, the mean age at breast cancer diagnosis in Cys557Ser carriers was 56.8 years (ranging from 42 to 75 years) compared to 51.9 years in non-carriers and 48.2 years in patients belonging to *BRCA1* and *BRCA2* families. The differences between ages of onset were not statistically significant.

**BARD1 Cys557Ser among Nordic families with breast and/or ovarian cancer (III)**

The *BARD1* Cys557Ser alteration was the only possible disease-related alteration observed in screening of the *BARD1* gene in study II. Its association with cancer was therefore further analysed in study III, which comprised an analysis of altogether 3,956 cancer patients and 3,591 controls from the Nordic countries. The cancer material studied included 2,906 breast and/or ovarian, 734 prostate, 188 colorectal and 128 male breast cancer cases. Results of the mutational analysis are shown in Table 6. During study III, a novel *BARD1* variant Ser558Pro, locating next to Cys557Ser was observed in the Norwegian material. However, due to lack of controls, its possible association with the risk of breast cancer remained unclear.
Table 6. Prevalence of the \textit{BARD1} Cys557Ser among Nordic patients with breast and/or ovarian cancer and healthy controls.

<table>
<thead>
<tr>
<th>Subgroup and study population</th>
<th>Frequency</th>
<th>OR (95% CI)</th>
<th>(p)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>2.7% (98/3591)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unselected cases</td>
<td>4.4% (88/1984)</td>
<td>1.7 (1.23 to 2.22)</td>
<td>0.001</td>
</tr>
<tr>
<td>Familial cases</td>
<td>5.2% (39/757)</td>
<td>1.9 (1.32 to 2.83)</td>
<td>0.001</td>
</tr>
<tr>
<td>\textit{BRCA1}/\textit{BRCA2} mutation-positive families(^1)</td>
<td>2.2% (5/223)</td>
<td>0.8 (0.33 to 2.03)</td>
<td>0.66</td>
</tr>
<tr>
<td>\textit{BRCA1}/\textit{BRCA2} mutation-negative families(^1)</td>
<td>6.8% (31/455)</td>
<td>2.6 (1.72 to 3.95)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Family history of breast and ovarian cancer or only ovarian cancer(^1,2)</td>
<td>3.8% (6/159)</td>
<td>1.4 (0.60 to 3.24)</td>
<td>0.45</td>
</tr>
<tr>
<td>Family history of breast cancer (without ovarian cancer)(^1)</td>
<td>6.8% (27/396)</td>
<td>2.6 (1.68 to 4.05)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

\(^1\) Includes samples for which the particular classification is known, \(^2\) Of these cases, ten displayed family history of only ovarian cancer and none of them carried the Cys557Ser allele.

The \textit{BARD1} Cys557Ser alteration was identified in 5.2% (39/757) of the index cases in families with breast and/or ovarian cancer from Finland, Iceland, Denmark and Sweden and in 2.7% (98/3591) of control individuals (Table 6). This indicates an about two-fold increased risk of familial cases \((p=0.001, \text{ OR } 1.9, 95\% \text{ CI } 1.3–2.8)\). The carrier frequencies varied greatly between different populations. In Finland, 4.9% of familial breast and/or ovarian cancer cases were carriers, the frequency being significantly higher than in controls \((p=0.004, \text{ OR } 3.1, 95\% \text{ CI } 1.5–6.5)\). Although the prevalence of heterozygotes among families with breast and/or ovarian cancer from Sweden and Iceland was almost two-fold higher than in control subjects, the differences were not statistically significant \((p=0.3 \text{ and } p=0.1, \text{ respectively})\). The frequency in Norwegian families was similar to other Nordic countries. However, no control samples from Norway were available for comparison, and therefore the Norwegian material was not included in the combined figures. The elevated incidence both among unselected cases and controls from Denmark indicated a high carrier frequency of the variant in the Danish population. Additionally, unlike in the other Nordic populations, the prevalence in familial cases was slightly lower compared to controls or sporadic cases, and statistical significance was not observed between any of these groups.

During the screening process, two individuals homozygous for the variant allele were found, one from Denmark and the other from Iceland. The Danish woman was diagnosed with breast cancer at age 51. Her mother was heterozygous
for the allele and developed breast cancer at age 56, whereas her sister did not display the alteration and was cancer-free at age 57. The Icelandic woman was affected with breast cancer at age 47 and belonged to the group of unselected breast cancer cases. Her family history of cancer was unknown. To analyse segregation of Cys557Ser with the disease in the other families studied, available affected and unaffected family members were tested for mutation status. The occurrence of the Cys557Ser allele in additional family members varied considerably between the families studied, and although many patients did have the alteration, both healthy carriers and affected non-carriers were observed.

Consistently with the observations in study II, the age at breast cancer diagnosis was found to be similar for Cys557Ser mutation carriers and non-carriers among all familial breast and/or ovarian cancer cases (50.6 vs 50.0, \( p=1 \)), as well as among both Finnish (51.4 vs 58.1 years, \( p=0.1 \)) and Icelandic (53.9 vs 58.1 years, \( p=0.5 \)) unselected breast cancer cases.

**BARD1 Cys557Ser is found particularly in families without cases of ovarian cancer and in BRCA1/BRCA2 mutation negative families (III)**

To evaluate further the prevalence of Cys557Ser among breast and/or ovarian cancer cases, the study cohort was subgrouped according to family history of breast, breast and ovarian or ovarian cancer, as well as according to whether or not families displayed BRCA1 and BRCA2 mutations. Of these subgroups, the highest frequencies of the BARD1 Cys557Ser allele were seen among patients from BRCA1/BRCA2 mutation-negative families and also from families lacking any cases of ovarian cancer (Table 6). The pooled odds ratios for the combined figures were further estimated using the Mantel-Haenszel method.

The overall prevalence of Cys557Ser carriers among patients without family history of ovarian cancer was significantly higher than among controls (6.8% vs 2.7%, \( p=0.000009 \), OR 2.6, 95% CI 1.7–4.0). The combined odds ratio calculated with the Mantel-Haenszel method was 1.8 (\( p=0.01 \), 95% CI 1.2–2.9). In contrast, the frequency among patients with a family history of both breast and ovarian or only ovarian cancer did not reach statistical significance (3.8% vs 2.7%, \( p=0.5 \)). In addition, a strong association was found in patients without BRCA1/BRCA2 mutations in the family. Within this subgroup, 6.8% carried the BARD1 variant compared to 2.7% in controls (\( p=0.000003 \), OR 2.6, 95% CI 1.7–4.0). The Mantel-Haenszel approximation also showed a highly significant association (\( p=0.007 \), OR 1.8, 95% CI 1.2–2.8). In contrast, only 2.2% of the cases from
BRCA1/BRCA2 mutation-positive families carried the allele, the frequency not differing from that of controls \((p=0.7)\), but being significantly lower than in the families without BRCA1/BRCA2 mutations \((p=0.01, \text{OR } 3.2, 95\% \text{ CI } 1.2–8.3)\). The corresponding Mantel-Haenszel odds ratio was 3.4 \((p=0.01, 95\% \text{ CI } 1.3–9.6)\).

BARD1 Cys557Ser seems not to associate with other cancers (III)

Since the results from study II gave evidence that BARD1 Cys557Ser might associate with breast cancer but not with ovarian cancer, the evaluation of the possible association of the alteration with other types of cancer was included in study III.

In order to test whether the Cys557Ser variant also associates with breast cancer in men or with other types of cancer, its prevalence was determined in 128 unselected male breast and 188 familial colorectal cancer cases, as well as in 121 familial and 613 unselected prostate cancer cases from Finland. 2.3% of both the males with breast cancer and male controls carried the alteration, indicating no malignancy association. A significantly increased frequency of Cys557Ser was not detected among patients with colorectal or prostate cancer, either. The prevalence among Finnish familial (2.5%) as well as in Finnish unselected (2.6%) prostate cancer patients was similar to that among male controls. Equal frequencies were also observed between familial colorectal cancer patients (1.6%) and population-based controls (1.6%).
5.3 **TOPBP1** germline alterations among Finnish families with breast and/or ovarian cancer (IV)

### 5.3.1 Genetic analysis of TOPBP1

The present study examined for the first time the putative impact of **TOPBP1** on breast cancer.

The analysis of 125 Finnish breast and/or ovarian cancer families revealed altogether 19 germline alterations in the **TOPBP1** gene. Ten of the observed alterations occurred in exon regions (Table 7), three of which were novel, while the rest have been previously reported in the SNP database. All missense variants were located outside of the functional BRCT domains. The locations of the identified amino acid changes are shown in Fig. 7. Four novel and five previously known variants were observed in intron regions. All alterations were assessed for possible effects on splice site consensus sequences (http://www.fruitfly.org/seq_tools/splice.html), and coding sequence variants were tested using the ESEfinder 2.0 program (http://rulai.chsl.edu/tools/ESE/) to ascertain whether the observed variants fell within predicted exonic splicing enhancer (ESE) sequences and whether they would affect ESE functions.

This study revealed one possibly disease-associated variant in exon eight of **TOPBP1**. The incidence of this novel Arg309Cys alteration was significantly elevated among the familial cancer cases when compared to healthy controls ($p=0.002$, OR 2.4, 95% CI 1.3–4.2) (Table 7). The nucleotide substitution leading to Arg309Cys showed no effect on binding sites of exonic splicing enhancers, but was located within a consensus splicing sequence and predicted to affect the consensus acceptor site. The frequency of Arg309Cys in breast cancer patients without known family history of the disease (9.1% [17/187]) was not significantly higher than in controls ($p=0.341$, OR 1.3, 95% CI 0.7–2.4). The other exonic missense variants as well as all of the intronic changes observed were seen at similar frequencies in both cases and controls and were classified as harmless polymorphisms. Moreover, although the novel A to G transition at position 4336 was absent from the controls, it resulted in a synonymous change (Ser1417Ser) and did not have any effect on predicted splicing consensus sequences or ESE motifs, and was therefore unlikely to be a pathogenic alteration. Consequently, apart from Arg309Cys, none of the other **TOPBP1** variants observed appeared to relate to cancer susceptibility and were therefore not studied further.
Table 7. Germline alterations in the TOPBP1 gene among Finnish breast and/or ovarian cancer cases.

<table>
<thead>
<tr>
<th>Nucleotide change</th>
<th>Effect on protein</th>
<th>Frequency in familial cases</th>
<th>Frequency in controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>1010C&gt;T</td>
<td>Arg309Cys</td>
<td>15.2% (19/125)</td>
<td>7.0% (49/697)</td>
</tr>
<tr>
<td>Polymorphisms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1454A&gt;C</td>
<td>Lys457Gln</td>
<td>33.6% (42/125)</td>
<td>39.2% (116/296)</td>
</tr>
<tr>
<td>1996T&gt;C</td>
<td>Val637Val</td>
<td>20.8% (26/125)</td>
<td>21.1% (19/90)</td>
</tr>
<tr>
<td>2535C&gt;T</td>
<td>Ser817Leu</td>
<td>26.4% (33/125)</td>
<td>27.7% (155/559)</td>
</tr>
<tr>
<td>3097G&gt;A</td>
<td>Val1004Val^{2}</td>
<td>3.2% (4/125)</td>
<td>3.5% (11/311)</td>
</tr>
<tr>
<td>3123G&gt;A</td>
<td>Arg1013Gln</td>
<td>12.0% (15/125)</td>
<td>10.9% (34/311)</td>
</tr>
<tr>
<td>3210A&gt;G</td>
<td>Asn1042Ser</td>
<td>36.0% (45/125)</td>
<td>38.2% (34/89)</td>
</tr>
<tr>
<td>4163C&gt;T</td>
<td>Leu1360Leu</td>
<td>35.2% (44/125)</td>
<td>34.1% (31/91)</td>
</tr>
<tr>
<td>4336A&gt;G</td>
<td>Ser1417Ser</td>
<td>0.8% (1/125)</td>
<td>0% (0/362)</td>
</tr>
<tr>
<td>4432A&gt;G</td>
<td>Ser1449Ser</td>
<td>22.4% (28/125)</td>
<td>22.0% (20/91)</td>
</tr>
</tbody>
</table>

1 For heterozygotes, 2 Based on the effect on protein and on the difference in allele prevalence between patients and controls, 3 Always observed together with Arg1013Gln.

TOPBP1 Arg309Cys was present in 19 of 125 familial index cases. To study further the role of Arg309Cys in breast cancer predisposition, available affected and unaffected family members were tested for mutation status. Additional samples were obtained from only eight of 19 families. From four families one additional sample from a breast cancer patient was available, and in all of these families also the other women diagnosed with breast cancer carried the Arg309Cys allele. In the remaining four families, both affected non-carriers and healthy carriers were observed. Due to the small number of additional DNA samples, the segregation of Arg309Cys with the disease in most of the families having an Arg309Cys carrier remained unclear. However, incomplete segregation pattern was observed in some of the families. None of the index cases carrying TOPBP1 Arg309Cys belonged to the ten BRCA1 or BRCA2 mutation-positive families. Furthermore, the prevalence of Arg309Cys appeared to be similar among patients having either a family history of breast or breast and/or ovarian cancer (18.2% [6/33] vs 14.1% [13/92], \( p=0.578 \)). The distribution of carriers between high- and moderate-risk families was also equal.

In order to identify individuals homozygous for the variant allele (1010T/T), restriction enzyme analysis was performed in familial and control groups. In the familial material one T/T homozygous individual diagnosed with breast cancer at
age 45 was observed. The only additional sample from the same family was from a paternal aunt who had breast cancer at age 52 and tested heterozygous for the alteration. No T/T homozygotes were found among the 697 controls.

5.3.2 Functional studies of TOPBP1 (partly unpublished data)

To evaluate the possible functional consequences of the Arg309Cys substitution, TOPBP1 protein levels and checkpoint activation were studied in two wild-type and four Arg309Cys heterozygous lymphoblastoid cell lines (Fig. 8).

The characterisation of TOPBP1 protein levels in wild-type and Arg309Cys heterozygous LCLs showed that TOPBP1 expression was not reduced in LCLs derived from Arg309Cys carriers, indicating that Arg309Cys has no effect on TOPBP1 protein levels (Fig. 8 A & B). The Western blot analysis did not show any aberrant-sized protein products, either. These observations implicated that the variant appears not to have any effect at the mRNA level. In order to investigate whether wild-type and heterozygous LCLs differ in their response to replication arrest, the ability to phosphorylate p53 on Ser392 and CHK1 on Ser345 was measured after hydroxyurea (HU) treatment (unpublished data). The phosphorylation levels of both p53 and CHK1 were similar between control and Arg309Cys heterozygous LCLs, as were the levels of the corresponding proteins (Fig. 8 C & D). The data presented also demonstrate that during the course of treatment with HU, phosphorylation of both p53 and CHK1 occurred at the same time points in wild-type and Arg309Cys heterozygous cell lines (Fig. 8 C & D), indicating that the amino acid change does not cause a delay in the replication block response. For all functional tests, the mean value of analysis in triplicate was calculated. Unfortunately, no cell line from the patient homozygous for the 1010T allele was available.
Fig. 8. Characterisation of cell lines heterozygous for the TOPBP1 Arg309Cys alteration (unpublished data). A) Arg309Cys has no effect on TOPBP1 protein levels. Western blot analysis was performed for two control cell lines and four heterozygous Arg309Cys mutant cell lines (marked with *). B) Graphic presentation of TOPBP1 protein levels. The average value of densitometric analysis of three independent westerns is shown. The intensity of TOPBP1 signal was quantified by using the ImageQuant program, and the level observed in control samples was taken as 100%. C) and D) Checkpoint activation after replication block. Control lymphoblasts and lymphoblasts harbouring the TOPBP1 Arg309Cys alteration were treated with 1 mM hydroxyurea and analysed for CHK1 and phospho-CHK1 levels (C) and p53 and phospho-p53 levels (D) by Western blot analyses. For all experiments, β-tubulin was used as loading control.
6 Discussion

6.1 BACH1 and hereditary predisposition to breast cancer (I)


Strong evidence of the involvement of BACH1 mutations in breast cancer has come from a study indicating that monoallelic truncating mutations in BACH1 are breast cancer susceptibility alleles conferring a two-fold relative risk of breast cancer (Seal et al. 2006). The study identified altogether six different truncating mutations (Table 8), five of which (141delC, 2392C>T, IVS17+2insT, 2008insT and 2255delAA) were present in 0.7% (9/1,212) of individuals with breast cancer from BRCA1 and BRCA2 mutation negative families, and two (2392C>T and 2108delAinsTCC) were identified among 2,081 controls with a frequency of 0.1%. Each of the mutations was rare, being present in only one to five carriers among the 1,212 breast cancer patients, giving frequencies of 0.08 to 0.4%. Three of the truncating mutations were the same as those previously identified as biallelic in patients with Fanconi anaemia complementation group J (Table 8) (Levitus et al. 2005, Levrant et al. 2005, Litman et al. 2005). Recently, another study observed a deletion of four nucleotides in exon 20 of BACH1 (2992-2995delAAGA) resulting in a premature stop at codon 1057. This amino acid change partially encompasses the BRCA1 binding domain of BACH1 and interferes with the stability of BACH1 and with its ability to interact with BRCA1 (De Nicolo et al. 2008). Previously, a BACH1 variant 3401delC that truncates the protein by 100 amino acids was observed in one family with history of breast cancer (Lewis et al. 2005). Table 8 summarises the reported breast cancer-associated truncating mutations in BACH1.

In addition to truncating mutations, several missense variants have been reported in BACH1 that may associate with breast cancer risk (Table 9) (Cantor et al. 2001, Rutter et al. 2003, Sigurdson et al. 2004, Lewis et al. 2005, Seal et al. 2006, Vahteristo et al. 2006b, Cao et al. 2009b, De Nicolo et al. 2008, Guenard et al. 2008). Indeed, the first indication of a potential association between germline
BACH1 mutations and breast cancer development came from a study that identified two BACH1 germline missense mutations (Pro47Ala and Met299Ile) in two early-onset breast cancer patients but not in any of the 200 controls (Cantor et al. 2001). Both of the observed mutations located in the helicase domain; Pro47Ala in a highly conserved residue within the nucleotide-binding domain and Met299Ile in a helicase homology region, and both were shown to disturb BACH1 protein function by altering both ATPase and helicase activity. Pro47Ala was also associated with BACH1 protein destabilisation (Cantor et al. 2001). A later study observed the P47A alteration in four of 1,212 affected individuals (0.3%) but also in four of 2,081 controls (0.2%), suggesting that despite the deleterious effect on BACH1 function, it probably does not confer a risk of breast cancer (Seal et al. 2006). The Met299Ile alteration has not been reported by others. Additionally, although the BACH1 Arg173Cys variant, locating in a putative nuclear localisation signal, has been reported to impair the nuclear localisation of BACH1, it does not seem to associate with breast cancer susceptibility (Luo et al. 2002, Lei & Vorechovsky 2003, Rutter et al. 2003, Lewis et al. 2005, Seal et al. 2006, Guenard et al. 2008). For the rest of the reported disease-associated alterations, the possible effect on BACH1 protein function is unknown or based on different computational predictions. Functional studies are therefore warranted to resolve their potential significance in cancer.

The reported, possibly breast cancer-associated missense alterations in BACH1 are summarised in Table 9. Two missense variants Val193Ile and Leu195Pro have been suggested to associate with breast cancer susceptibility in several studies (Rutter et al. 2003, Lewis et al. 2005, Vahteristo et al. 2006b), while other studies have not found the association for either of the variants (Cantor et al. 2001, Seal et al. 2006, Guenard et al. 2008). In the present study, the Pro919Ser alteration was found in 50% of the cancer cases. This amino acid change has been associated with a risk of pre-menopausal breast cancer in one later study (Sigurdson et al. 2004), but it has not been considered to be disease-associated in two other cohorts of pre-menopausal breast cancer (Table 9) (Vahteristo et al. 2006b, Frank et al. 2007). Our results are consistent with the other studies comprising breast cancer cases of all ages, in which P919S has been found either with similar frequencies as in our study or in equal frequencies in both cancer cases and controls, and considered to be a non-disease associated polymorphism (Cantor et al. 2001, Luo et al. 2002, Rutter et al. 2003, Seal et al. 2006, Vahteristo et al. 2006a, Cao et al. 2009b, Guenard et al.
Thus at present, there is no evidence for Pro919Ser to be considered in breast cancer susceptibility.

The novel Pro1034Leu alteration identified in one family in the present study (Table 9) changes proline to leucine at codon 1034 and locates at the C-terminal region of BACH1. This non-conservative amino acid change seems to be rare and has not been observed in other studies. Originally this amino acid position was thought to locate at BRCA1 binding domain of the BACH1 protein (Cantor et al. 2001), but more recently it was mapped outside of the more narrowly defined interaction domain (Cantor et al. 2004). The alteration also resides outside of the BACH1 helicase domain and does not disrupt ATPase or helicase function (Cantor et al. 2004). Although it does not compromise BACH1 enzymatic function, it remains to be determined whether it is associated with loss of any other functions including its ability to participate in DNA damage response. Proline and leucine are biochemically different amino acids, both are non-polar, but while proline is a cyclic amino acid, leucine has a branched hydrocarbon side chain. Although the Pro1034Leu alteration was also found in one of the 304 anonymous population controls of unknown cancer history, the effect on cancer risk cannot be totally excluded. Larger population-based studies with breast and ovarian cancer patients might reveal whether this variant could be a very rare cancer-associated alteration.

Most of the disease-associated BACH1 missense variants have been observed only in a single breast cancer case comprising less than 1.5% of the cases and absent from controls. Due to the rareness of these variants, the difference in frequencies between affected individuals and controls is unlikely to reach statistical significance. A recent study found a marginally significant association of two SNPs (rs4988344 and rs2191249) in introns 5 and 14 of the BACH1 gene with ovarian cancer (Song et al. 2007). Furthermore, several BACH1 missense alterations not included in Table 9 have been reported that would also need further investigation. These include variants that have been observed to be present in a single case and absent from controls and predicted not to have any effect on protein function as well as variants whose frequencies have not been studied in controls (Seal et al. 2006, Cao et al. 2009b).

Rare truncating mutations in BACH1 are estimated to confer about two-fold relative risk of breast cancer (Seal et al. 2006). The available data indicate that the majority of BACH1 missense variants are not associated with a risk of breast cancer comparable to that conferred by truncating variants. However, the biological significance of most of the identified sequence alterations has not been
resolved, and it is possible that specific rare missense variants confer susceptibility to breast or ovarian cancer. Overall, mutations in \textit{BACH1} seem to be very rare and may account for only a very small proportion of familial breast cancer. Since we found no clearly pathogenic \textit{BACH1} mutations among the 151 Finnish breast and/or ovarian cancer families studied, the present study indicates that germline alterations in \textit{BACH1} do not seem to play any significant role in familial breast and/or ovarian cancer in the Finnish population. These results were supported by another Finnish study, which was unable to detect any disease-related mutations in \textit{BACH1} among \textit{BRCA1}- and \textit{BRCA2}-negative breast cancer families (Vahteristo \textit{et al.} 2006b). No truncating \textit{BACH1} mutations have been identified in Finnish breast cancer families. However, since the disease-causing alterations in \textit{BACH1} seem to be very rare among the populations studied, the possibility cannot be excluded that cancer-associated mutations could also be present in the Finnish population.
Table 8. Summary of the reported breast cancer-associated truncating mutations in the BACH1 gene.

<table>
<thead>
<tr>
<th>Nucleotide change</th>
<th>Effect on protein</th>
<th>References classifying as possibly pathogenic</th>
<th>Frequency(^1) in references classifying as possibly pathogenic</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>141delC</td>
<td>Premature truncation</td>
<td>Seal et al. 2006</td>
<td>0.1% (1/1212)</td>
<td>0% (0/2081)</td>
</tr>
<tr>
<td>2008insT</td>
<td>Premature truncation</td>
<td>Seal et al. 2006</td>
<td>0.1% (1/1212)</td>
<td>0% (0/2081)</td>
</tr>
<tr>
<td>2108delAinsTCC</td>
<td>Premature truncation</td>
<td>Seal et al. 2006</td>
<td>0% (0/1212)</td>
<td>0.05% (1/2081)</td>
</tr>
<tr>
<td>2255delAA</td>
<td>Lys752fsX764</td>
<td>Seal et al. 2006</td>
<td>0.1% (1/1212)</td>
<td>0% (0/2081)</td>
</tr>
<tr>
<td>2392C&gt;T</td>
<td>Arg798X</td>
<td>Seal et al. 2006</td>
<td>0.4% (5/1212)</td>
<td>0.05% (1/2081)</td>
</tr>
<tr>
<td>IVS17+2insT</td>
<td>Exon 17 or 18 skipped</td>
<td>Seal et al. 2006</td>
<td>0.1% (1/1212)</td>
<td>0% (0/2081)</td>
</tr>
<tr>
<td>2992delAAGA</td>
<td>Glu998fsX1057</td>
<td>De Nicolo et al. 2008</td>
<td>2.0% (1/49)</td>
<td>0% (0/50)</td>
</tr>
<tr>
<td>3401delC</td>
<td>1134fsX1149</td>
<td>Lewis et al. 2005</td>
<td>1.3% (1/75)</td>
<td>0% (0/93)</td>
</tr>
</tbody>
</table>

\(^1\) For heterozygotes, fs frameshift.
Table 9. Summary of the reported possibly breast cancer-associated missense alterations in the BACH1 gene.

<table>
<thead>
<tr>
<th>Nucleotide change</th>
<th>Effect on protein</th>
<th>References classifying as possibly pathogenic</th>
<th>Frequency(^1) in references classifying as possibly pathogenic</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Familial cases</td>
<td>Controls</td>
<td></td>
</tr>
<tr>
<td>139C&gt;G</td>
<td>Pro47Ala</td>
<td>Cantor et al. 2001</td>
<td>2.9% (1/35)</td>
<td>0% (0/200)</td>
</tr>
<tr>
<td>415T&gt;G</td>
<td>Ser139Ala</td>
<td>Guenard et al. 2008</td>
<td>1.0% (1/96)</td>
<td>0% (0/71)</td>
</tr>
<tr>
<td>430G&gt;A</td>
<td>Ala144Thr</td>
<td>Lewis et al. 2005(^a)</td>
<td>1.3% (1/75)</td>
<td>0% (0/93)</td>
</tr>
<tr>
<td>577G&gt;A</td>
<td>Val193Ile</td>
<td>Rutter et al. 2003</td>
<td>0.9% (1/116)</td>
<td>0% (0/60)</td>
</tr>
<tr>
<td>584T&gt;C</td>
<td>Leu195Pro</td>
<td>Rutter et al. 2003</td>
<td>0.9% (1/116)</td>
<td>0% (0/60)</td>
</tr>
<tr>
<td>790C&gt;T</td>
<td>Arg264Trp</td>
<td>Seal et al. 2006(^c)</td>
<td>0.4% (5/1212)</td>
<td>0.29% (6/2081)</td>
</tr>
<tr>
<td>823A&gt;G</td>
<td>Ile275Val</td>
<td>Guenard et al. 2008</td>
<td>1.0% (1/96)</td>
<td>0% (0/66)</td>
</tr>
<tr>
<td>897G&gt;A</td>
<td>Met299Ile</td>
<td>Cantor et al. 2001</td>
<td>3.3% (1/30)(^d)</td>
<td>0% (0/200)</td>
</tr>
<tr>
<td>1619A&gt;T</td>
<td>Gln540Leu</td>
<td>Rutter et al. 2003</td>
<td>2.4% (1/42)</td>
<td>0% (0/60)</td>
</tr>
<tr>
<td>2468G&gt;T</td>
<td>Arg823Ser</td>
<td>Seal et al. 2006(^d)</td>
<td>0.1% (1/1212)</td>
<td>0% (0/2081)</td>
</tr>
<tr>
<td>2755C&gt;T</td>
<td>Pro919Ser</td>
<td>Sigurdson et al. 2004(^e)</td>
<td>61% (34/56)</td>
<td>18% (10/56) for Ser919 homozygotes</td>
</tr>
<tr>
<td>2804T&gt;G</td>
<td>Val835Gly</td>
<td>Rutter et al. 2003</td>
<td>0.9% (1/116)</td>
<td>0% (0/60)</td>
</tr>
<tr>
<td>2830G&gt;C</td>
<td>Gin944Glu</td>
<td>Cao et al. 2009b</td>
<td>0.6% (2/357)</td>
<td>0% (0/864)</td>
</tr>
<tr>
<td>3101C&gt;T</td>
<td>Pro1034Leu</td>
<td>present study (I)</td>
<td>0.7% (1/151)</td>
<td>0.33% (1/304)</td>
</tr>
<tr>
<td>3464G&gt;A</td>
<td>Gly1155Glu</td>
<td>Lewis et al. 2005(^b)</td>
<td>1.3% (1/75)</td>
<td>0% (0/93)</td>
</tr>
</tbody>
</table>

\(^1\) For heterozygotes. \(^a\) Predicted by ESE finder to affect exonic splicing enhancers, \(^b\) Based on prediction by mFOLD to affect mRNA folding, \(^c\) Based on PolyPhen and/or SIFT predictions, \(^d\) Unselected cases, \(^e\) Association with familial breast cancer risk in premenopausal women.
6.2 BARD1 alterations and breast cancer susceptibility (II, III)

As BARD1 is a major cellular binding partner of BRCA1 and many of its functions are mediated through association with BRCA1, it has been suggested that it may be involved in breast or ovarian cancer susceptibility. In this study, one possible disease-associated BARD1 alteration, Cys557Ser, was observed. Besides the present study, involvement of BARD1 alterations in breast and ovarian cancer development has been suggested in several other studies (Thai et al. 1998, Ghimenti et al. 2002, Ishitobi et al. 2003, Sauer & Andrulis 2005, Stacey et al. 2006, Gorringe et al. 2008) while others have failed to find significant association (Vahteristo et al. 2006a, Huo et al. 2007, Jakubowska et al. 2008, Johnatty et al. 2009). Altogether 13 potentially disease-associated alterations in BARD1 have been reported and are summarised in Table 10. Five of these alterations occur in the BRCT regions (Thr598Ile, Cys645Arg, Ile692Thr, Val695Leu and Ser761Asn), five in the ankyrin repeats (Asn470Ser, Val507Met, Cys557Ser, Gln564His and Arg565His) and three occur outside the known functional domains (Asn295Ser, Lys312Asn and del358-364). No alterations have been found in the RING finger of BARD1, unlike for BRCA1, suggesting that BARD1 may have a BRCA1-independent role in tumour progression.

The first association between breast cancer and alterations in the BARD1 gene has been observed in the screening of patients with breast, ovarian and endometrial cancers that identified one possible cancer associated germline (Gln564His) and two somatic alterations (Ser761Asn and Val695Leu) (Thai et al. 1998). A study of an Italian cohort of familial breast and ovarian cancers has reported five alterations in BARD1 of which four, Val507Met, Cys557Ser, Gln564His and Arg565His) and three occur outside the known functional domains (Asn295Ser, Lys312Asn and del358-364). No alterations have been found in the RING finger of BARD1, unlike for BRCA1, suggesting that BARD1 may have a BRCA1-independent role in tumour progression.

The first association between breast cancer and alterations in the BARD1 gene has been observed in the screening of patients with breast, ovarian and endometrial cancers that identified one possible cancer associated germline (Gln564His) and two somatic alterations (Ser761Asn and Val695Leu) (Thai et al. 1998). A study of an Italian cohort of familial breast and ovarian cancers has reported five alterations in BARD1 of which four, Val507Met, Cys557Ser, Gln564His and Arg565His) and three occur outside the known functional domains (Asn295Ser, Lys312Asn and del358-364). No alterations have been found in the RING finger of BARD1, unlike for BRCA1, suggesting that BARD1 may have a BRCA1-independent role in tumour progression.

Analysis of BARD1 in Japanese patients with familial breast cancer has observed that the Val507Met alteration associates with breast cancer in postmenopausal women (Ishitobi et al. 2003). The study has also revealed another alteration (Asn470Ser) likely to associate with breast cancer. A study of unselected Canadian breast and ovarian tumours and breast cancer cell lines revealed one novel putative disease-associated alteration Cys645Arg (Sauer & Andrulis 2005). Furthermore, an Australian study has identified three alterations, Arg565His, Thr598Ile and Ile692Thr, that may associate with breast cancer (Gorringe et al. 2008). Most of these cancer-associated alterations have only been observed in single studies, while Cys557Ser and Ser761Asn have been reported as disease-associated in several studies (Sauer & Andrulis 2005, Stacey et al.
In contrast, other studies have reported that Met507Val, 1144del21, Cys557Ser, Ser761Asn do not associate with breast or ovarian cancer (Thai et al. 1998, Ishitobi et al. 2003, Valteristo et al. 2006a, Huo et al. 2007, Jakubowska et al. 2008, Gorringe et al. 2008). The present study identified two BARD1 alterations, Val507Met and Cys557Ser, that have been reported to associate with breast cancer in other studies. The current results provide further evidence for the involvement of Cys557Ser in breast carcinogenesis. Additionally, the study is consistent with reports that classify the Val507Met variant as a neutral polymorphism rather than a pathogenic alteration.

Seven of the disease-associated BARD1 alterations (Asn295Ser, Lys312Asn, Cys557Ser, Gln564His, Cys645Arg, Val695Leu and Ser761Asn) have been characterised in functional assays (Irminger-Finger et al. 2001, Sauer & Andrulis 2005), and appear to affect the stability of p53, as well as abrogate the growth-suppressive and apoptotic activities of BARD1. Gln564His results in reduced binding of CSTF1 and thereby abolishes the inhibition of mRNA processing in response to DNA damage (Kleiman & Manley 2001). Furthermore, a recent study provides mechanisms by which three of these alterations, Cys645Arg, Val695Leu and Ser761Asn, may adversely affect the structure and tumour suppressor function of BARD1 (Birrane et al. 2007): Cys645Arg destabilises the folding of the first BRCT domain, Val695Leu has a destabilising effect on the folding of the second BRCT domain, and Ser761Asn reduces the affinity of BARD1 for BRCA1. None of the alterations studied appears to affect the cellular localisation of BRCA1, and only Ser761Asn may disrupt the interaction between BARD1 and BRCA1. These results indicate further that BARD1 has a BRCA1-independent role in tumourigenesis and that the disease-associated BARD1 variants affect BRCA1-independent functions of BARD1. The biological significance of the rest of the disease-associated BARD1 alterations has not been resolved in any study.

Many of the cancer-associated BARD1 alterations are individually rare, while others, like the Cys557Ser variant, are frequent. In general, it seems that BARD1 mutations associated with inherited and spontaneous cases of breast and ovarian cancer are quite infrequent and the involvement of many alterations seems controversial. Based on the present data, some BARD1 alterations may be associated with an increased risk of breast cancer. Therefore, more extensive studies will still be needed to address the role of BARD1 in breast cancer predisposition. However, it seems that the contribution of the BARD1 germline variants to breast cancer predisposition may be small.
Table 10. Summary of the BARD1 alterations reported as possibly disease-associated.

<table>
<thead>
<tr>
<th>Nucleotide change</th>
<th>Effect on protein</th>
<th>Cancer type</th>
<th>References classifying as possibly pathogenic</th>
<th>Frequency(^1) in references classifying as possibly pathogenic</th>
<th>References classifying as nonpathogenic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Familial cases</td>
<td>Unselected cases</td>
<td>Controls</td>
</tr>
<tr>
<td>957A&gt;G</td>
<td>Asn295Ser</td>
<td>BC</td>
<td>Ghimenti \textit{et al.} 2002</td>
<td>2.5% (1/40)</td>
<td>0% (0/20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sauer &amp; Andrulis 2005(^5)</td>
<td>NA NA NA</td>
<td></td>
</tr>
<tr>
<td>1009A&gt;T</td>
<td>Lys312Asn</td>
<td>BC</td>
<td>Ghimenti \textit{et al.} 2002</td>
<td>NA 0% (0/20)</td>
<td>0% (0/60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sauer &amp; Andrulis 2005(^5)</td>
<td>2.5% (1/40) NA NA</td>
<td></td>
</tr>
<tr>
<td>1145del21bp</td>
<td>del358-364</td>
<td>BC</td>
<td>Ghimenti \textit{et al.} 2002</td>
<td>2.5% (1/40) 5.0% (1/20)</td>
<td>0% (0/60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1743G&gt;C</td>
<td>Cys557Ser</td>
<td>BC, Ov</td>
<td>Ghimenti \textit{et al.} 2002</td>
<td>2.5% (1/40)</td>
<td>0% (0/20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sauer &amp; Andrulis 2005(^5)</td>
<td>1.7% (1/60) NA NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stacey \textit{et al.} 2006</td>
<td>3.7% (7/190) 2.6% (21/802)</td>
<td>1.6% (11/703)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>present study (II)</td>
<td>5.2% (39/757) 4.4% (88/1984)</td>
<td>2.7% (88/3591)</td>
</tr>
<tr>
<td>1765G&gt;C</td>
<td>Glu564His</td>
<td>Ov</td>
<td>Thai \textit{et al.} 1998</td>
<td>NA 1.7% (1/58)</td>
<td>0% (0/300)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sauer &amp; Andrulis 2005(^5)</td>
<td>NA NA NA</td>
<td></td>
</tr>
<tr>
<td>1767G&gt;A</td>
<td>Arg565His</td>
<td>BC</td>
<td>Gorringe \textit{et al.} 2007</td>
<td>0% (0/210)</td>
<td>0.3% (1/354)</td>
</tr>
<tr>
<td>1866C&gt;T</td>
<td>Thr598Ile</td>
<td>BC</td>
<td>Gorringe \textit{et al.} 2007</td>
<td>0.5% (1/210)</td>
<td>0% (0/354)</td>
</tr>
<tr>
<td>2006T&gt;C</td>
<td>Cys645Arg</td>
<td>BC, Ov</td>
<td>Sauer &amp; Andrulis 2005</td>
<td>3.3% (260)</td>
<td>NA 0% (0/45)</td>
</tr>
<tr>
<td>2148A&gt;C</td>
<td>Ile692Thr</td>
<td>BC</td>
<td>Gorringe \textit{et al.} 2007</td>
<td>0.5% (1/210)</td>
<td>0% (0/354)</td>
</tr>
<tr>
<td>2156G&gt;C</td>
<td>Val695Leu</td>
<td>BC</td>
<td>Thai \textit{et al.} 1998</td>
<td>NA 2.0% (1/50) (^a)</td>
<td>0% (0/300)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sauer &amp; Andrulis 2005(^5)</td>
<td>NA NA NA</td>
<td></td>
</tr>
<tr>
<td>2355G&gt;A</td>
<td>Ser761Asn</td>
<td>BC, Ut</td>
<td>Thai \textit{et al.} 1998</td>
<td>NA 1.7% (1/60) (^a)</td>
<td>0% (0/300)</td>
</tr>
</tbody>
</table>

\(^1\) For heterozygotes, \(^a\) Somatic change, \(^b\) Based on functional study, BC breast cancer, NA not available, Ov ovarian cancer, Ut uterine cancer.
6.2.1 The BARD1 Cys557Ser variant seems to be associated with genetic predisposition to breast cancer (II, III)

Since the first observation that BARD1 Cys557Ser is associated with an increased risk of breast cancer, a number of studies have examined its role among various populations. The results of the present study provides additional evidence that Cys557Ser is a breast cancer susceptibility allele that confers an about two-fold increased risk of breast cancer, even though other studies have not found a statistically significant association between the allele and breast cancer (Table 10).

BARD1 Cys557Ser was initially seen in a sporadic Caucasian tumour material consisting of breast, ovarian and uterine cancers with a frequency of 2%, and was classified as a rare polymorphism (Thai et al. 1998). Later, in an Italian study of 40 breast and/or ovarian cancer families, the alteration was observed in four women belonging to one family (Ishitobi et al. 2003). Three of the women had breast or breast and/or ovarian cancer, and the fourth was unaffected by the age of 39. As Cys557Ser was not present among any of the 60 controls or 20 sporadic cases studied, it was suggested to be a possible cancer susceptibility allele. In an Icelandic study of 1,090 breast cancer cases and 703 controls, the Cys557Ser variant was associated with a 1.8-fold increased risk of breast cancer (Stacey et al. 2006). The frequency was particularly high in women with breast cancer carrying the Icelandic founder mutation BRCA2 995del5. In contrast, BARD1 Cys557Ser has been reported not to associate with breast cancer in Southern Finland, Australia and Poland (Vahteristo et al. 2006a, Jakubowska et al. 2008, Gorringe et al. 2008, Johnatty et al. 2009). In Southern Finnish cohort, the Cys557Ser allele was observed in 1.4% of familial and 2.2% of unselected breast cancer patients and in 2.5% of healthy controls (Vahteristo et al. 2006a). The prevalence of the BARD1 Cys557Ser allele in a Polish study of 3,188 unselected women with breast cancer and 1,038 healthy controls was slightly greater in cases than in controls (4.7% vs 3.8%), but the difference was not statistically significant (Jakubowska et al. 2008). Two Australian studies found the allele in 4.8% and 4.4% of breast cancer cases and 4.8% and 5.0% of controls, respectively (Gorringe et al. 2008, Johnatty et al. 2009).

The cysteine 557 of BARD1 locates on the recently identified fourth ankyrin repeat of the protein (Fig. 5) (Fox et al. 2008), in a region that has been shown to be essential for the binding of p53 as well as for the induction of p53-dependent apoptosis (Jefford et al. 2004, Feki et al. 2005). The region is also required for the
interaction of BARD1 with proto-oncogene BCL3 (Dechend et al. 1999). In fact, according to a functional study, Cys557Ser has a harmful effect on protein function, as it was shown to abrogate the growth-suppressive and apoptotic functions of BARD1 (Sauer & Andrulis 2005). The induction of apoptosis is an important mechanism for tumour suppression, and the decreased apoptotic activity of Cys557Ser matches its association with cancer. Furthermore, Cys557 is conserved between several species (H. Sapiens, M. musculus, R. norvegicus, X. laevis).

Based on the data observed in both studies II and III, the segregation of the Cys557Ser allele with breast cancer appeared to be incomplete. However, the limited evidence of segregation in the pedigrees is the typical, and expected, pattern of low- and moderate-penetrance susceptibility alleles (Meijers-Heijboer et al. 2002, Renwick et al. 2006). The results from study III also provided further evidence for the indication in study II that Cys557Ser appeared to be enriched among families displaying breast but not ovarian cancer (Table 6). Moreover, patients from families without BRCA1/BRCA2 mutations were more likely to carry Cys557Ser than those from BRCA1/BRCA2 mutation-positive families (p=0.01), among whom no association was observed when compared to controls (p=0.7). This trend was seen within different Nordic populations. The frequency of Cys557Ser among BRCA1/BRCA2 mutation-negative families was elevated particularly in Finland (6.0% vs 1.6%, p=0.005) and Sweden (10.1% vs 3.7%, p=0.04) when compared to healthy controls. In Iceland the frequency was over two-fold higher than in controls, but the difference did not quite reach statistical significance (6.7% vs 3.1%, p=0.08). To obtain a more critical analysis of the combined data, Mantel-Haenszel pooled risk estimates were calculated for the comparisons. The Mantel-Haenszel combined OR values were quite similar to the initially obtained $\chi^2$-values, showing significant association, thereby confirming the results.

The present study is consistent with the reports classifying BARD1 Cys557Ser as a breast cancer-associated variant rather than a polymorphism. However, based on the current data the role of Cys557Ser in cancer predisposition seems controversial. There might be several reasons for the differences between the results of different studies reporting the role of the Cys557Ser allele among patients with breast cancer. Some of the observed variation in frequencies between different populations may be due to different patient ascertainment criteria as well as variability in methods. There may be differences between populations or in the tumour characteristics. For example, in the Icelandic study
the greatest association of Cys557Ser was seen for breast cancers of lobular or medullar histology, with a relative risk of 2.5 and 6.9, respectively (Stacey et al. 2006). The genetic background may also differ in a way that is relevant to susceptibility to breast cancer (e.g. the allele might predispose to breast cancer only in the presence of an additional allele of another gene). Differences in modifier loci or environmental differences may influence the penetrance of different mutations between different populations. In addition, chance may play a part, or there may be a modestly elevated risk of breast cancer associated with the allele, but only a sufficiently large study would have detected it. Furthermore, in the present study, no significant association of the BARD1 Cys557Ser variant has been observed with either male breast, ovarian, colorectal or prostate cancer. The possible association of BARD1 alterations with other than breast and ovarian cancer has not been investigated in other studies. Therefore, these observations need to be confirmed with larger patient series.

6.3 **TOPBP1 germline alterations do not seem to play a significant role in breast cancer in Finland (IV)**

TOPBP1 encodes a checkpoint protein that is required for DNA damage and replication processes and associates with BRCA1 in response to DNA damage. Therefore, to uncover additional predisposing genes we have performed a comprehensive screening for germline alterations in the TOPBP1 gene. The recent finding that TOPBP1 serves as an activator of the ATR-ATRIP complex and is needed for the activation of the checkpoint kinase CHK1 has further emphasised its importance in maintaining genomic integrity.

The present study identified one possible breast cancer associated alteration, Arg309Cys, in TOPBP1. This alteration is caused by a C>T transition at codon 1010 and is located within a consensus splicing sequence and predicted to affect the consensus acceptor site. The characterisation of TOPBP1 protein levels in wild-type and Arg309Cys heterozygous LCLs showed that TOPBP1 expression was not reduced in LCLs derived from Arg309Cys carriers, indicating that Arg309Cys has no effect on TOPBP1 protein levels. The Western blot analysis did not show any aberrant-sized protein products, either. These observations indicated that the variant appears not to have an effect at the mRNA level. Sequence comparison between several species showed that the Arg309 residue is conserved in both dog and frog, and falls within a region that shows conservation in TOPBP1 homologues from yeast to human. In fact, the replacement by a
cysteine residue could have a harmful impact on polypeptide structure and function, since arginine and cysteine are biochemically profoundly different. Both amino acids are hydrophilic, but while arginine has a guanidinium group as its side chain and is positively charged, cysteine is uncharged and carries a very reactive sulfhydryl group.

The significantly higher frequency among familial breast cancer patients compared to controls \((p=0.002)\) prompted us to investigate whether Arg309Cys has an effect on protein function. Since TOPBP1 is obviously involved in initiation of DNA replication and involved in the ATR signalling pathway, we hypothesised that Arg309Cys could affect the cellular responses to DNA replication-blocking lesions. Therefore, we investigated the phosphorylation of two key target proteins, p53 and CHK1, involved in DNA damage response pathways and activated through phosphorylation after genotoxic stress. We did not observe any difference in the phosphorylation levels of either protein between wild-type and heterozygous LCLs, indicating that the checkpoint response in heterozygous cell lines was triggered normally after HU-induced replication stalling. Furthermore, the activation of both CHK1 and p53 occurred at the same time points after HU treatment. The present study did not find any evidence that TOPBP1 Arg309Cys had a harmful impact on protein function. However, the possible influence on other cellular functions not considered here cannot be excluded. For instance, arginine at position 309 locates to the N-terminal part of the protein, adjacent to the second BRCT domain (Fig. 7), and some lines of evidence suggest that the amino-terminal region of TOPBP1 is responsible for transcriptional co-activation functions. The activity of human papilloma virus (HPV) transcription/replication factor E2 was shown to be enhanced by TOPBP1, and removal of the amino-terminal portion of TOPBP1 abolishes this function (Boner & Morgan 2002). Previously the amino terminus of TOPBP1 was shown to activate transcription in yeast (Mäkiniemi et al. 2001).

In the present study, genomic and functional approaches were integrated in order to analyse the possible role of TOPBP1 in breast and/or ovarian cancer predisposition. The results suggest that alterations in the TOPBP1 gene do not play a considerable role in breast cancer development in Finland. The novel Arg309Cys substitution was the only putative pathogenic alteration found in TOPBP1. It was observed at significantly higher frequency in familial breast and/or ovarian cancer patients compared to healthy controls, suggesting an about 2.4-fold increased disease risk for carriers. Segregation analysis with a limited number of samples from additional family members revealed incomplete
segregation of the alteration with the disease phenotype, as both affected non-carriers and healthy carriers were identified. This observation together with the high prevalence of the allele also in healthy controls and breast cancer patients without known family history of the disease indicates that TOPBP1 Arg309Cys is a commonly occurring allele that may have a subtle impact on cancer risk. To our knowledge, this is the first report of a mutation screening of the TOPBP1 gene, and further evaluation of the current findings will require additional studies.
7  Summary and concluding remarks

Since the currently known susceptibility genes explain only part of the hereditary predisposition to breast and/or ovarian cancer, additional predisposing genes remain to be identified. Based on observations from the known breast cancer associated genes, suitable candidates are likely to be found among genes whose protein products function in DNA damage response pathways or interact with the known susceptibility genes.

In this case-control study, three genes (BACH1, BARD1 and TOPBP1) whose protein products associate with BRCA1 were studied for mutations in breast and/or ovarian cancer families originating from Northern Finland. The prevalence of the BARD1 Cys557ser allele was also studied in families originating from the other Scandinavian countries, and in other types of cancer.

Subsequent to our investigations, BACH1 has recently been identified as a moderate breast cancer susceptibility gene. Certain rare missense variants as well as protein truncating mutations in BACH1 have been reported to associate with breast cancer. To investigate whether BACH1 mutations are found in Finnish families with breast and ovarian cancer, we screened altogether 214 affected women for possible germline alterations. One rare, putative disease-associated alteration, Pro1034Leu, was observed with a two-fold higher frequency in cancer cases compared to healthy controls (0.7% vs 0.3%), but the difference did not reach statistical significance. The results suggest that BACH1 alterations are rare in the Finnish population, and are unlikely to explain a significant proportion of familial breast and ovarian cancer susceptibility. Overall, the results are in line with other reports that indicate that mutations in BACH1 are rare and their contribution to familial breast cancer seems therefore marginal. Large-scale studies of breast and ovarian cancer cases and healthy controls would be needed to assess further the role of the rare Pro1034Leu variant. Additionally, further studies would also be required to investigate whether truncating BACH1 mutations that would not have been observed by the methods used in this study could be found among Finnish patients having a family history of breast cancer. Since breast cancer-associated truncating mutations in BACH1 cause Fanconi anaemia in biallelic carriers, the role of BACH1 in the Fanconi anaemia pathway and how this involvement relates to its tumour suppression function is also an interesting question.

As the constitutive nuclear partner of BRCA1, BARD1 has been considered as a candidate for breast and ovarian cancer susceptibility and has been under
intense investigation ever since its identification. There is accumulating evidence that certain missense mutations of BARD1 might contribute to the development of breast and ovarian cancers. Current mutation screening of the BARD1 gene in 126 Finnish breast and ovarian cancer patients revealed one putative disease-associated alteration, Cys557Ser. The alteration was found in 5.6% (7/126) of the cancer cases and in 1.4% (14/1018) of controls ($p=0.005$, OR 4.2, 95% CI 1.7–10.7). To evaluate further the possible effect of Cys557Ser on the risk of different types of cancer, its prevalence was analysed in altogether 3,956 cancer patients and 3,591 controls from the Nordic countries. The patients studied included 2,906 cases of breast and/or ovarian cancer, 734 cases of prostate cancer, 188 cases of colorectal cancer and 128 cases of male breast cancer. The results provided further evidence of the involvement of the Cys557Ser alteration in breast carcinogenesis, conferring an about two-fold increased risk. Cys557Ser was identified in 5.2% (39/757) of the index cases in families with breast and/or ovarian cancer from Finland, Iceland, Denmark and Sweden and in 2.7% (98/3,591) of the control individuals ($p=0.001$, OR 1.9, 95% CI 1.3–2.8). The highest frequency, 6.8% of the patients, was found both among patients from BRCA1/BRCA2 mutation-negative families ($p=0.000009$, OR 2.6, 95% CI 1.7–4.0) and among patients having no family history of ovarian cancer ($p=0.000009$, OR 2.6, 95% CI 1.7–4.0). The present study is consistent with the reports classifying BARD1 Cys557Ser as a breast cancer-associated variant rather than a harmless polymorphism. However, since the role of many BARD1 alterations, including Cys557Ser, in breast cancer predisposition is controversial, more extensive studies will still be needed to address the role of BARD1 in breast tumourigenesis. Additionally, although the current study did not indicate a role in the development of the any other studied malignancies, the results need to be confirmed with larger patient series. Overall it seems that the contribution of the BARD1 germline variants to breast cancer predisposition is marginal.

This is the first study evaluating the role of TOPBP1 as a candidate gene for familial breast cancer predisposition. In the analysis of 125 Finnish patients, one putative disease-associated alteration, Arg309Cys, was identified and found to be associated with an approximately two-fold increased risk of breast cancer. The Arg309Cys allele was observed in 15.2% of the familial cancer cases and in 7.0% of healthy controls ($p=0.002$, OR 2.4, 95% CI 1.3–4.2). Only one individual homozygote for the variant allele was observed among familial breast cancer cases, and none were seen among the 697 healthy controls. Protein analyses showed that Arg309Cys had no effect on TOPBP1 protein levels and no aberrant-
sized protein products were observed in heterozygous cell lines either. Furthermore, this study did not find any evidence that Arg309Cys would have harmful impact on protein function. In conclusion, the results obtained suggest that germline alterations in the TOPBP1 gene are unlikely to play a considerable role in breast cancer development in Finland. The Arg309Cys alteration is a relatively common allele, so it is likely to have only a minor contribution to breast cancer predisposition.

The results of this study fit well in the hypothesis of a polygenic model of breast cancer susceptibility, according to which a considerable proportion of the remaining familial breast and/or ovarian cancer cases could be due to the multiplicative effect of various low- to moderate-penetrance genes, perhaps in combination with environmental factors. In addition to rare breast cancer susceptibility alleles, common polymorphisms may also be associated with an increased risk of breast cancer. Genetic variants that are more common in the population than high-penetrance gene mutations may in aggregate make a substantially greater contribution to cancer. Since the genetic contribution to breast cancer development remains largely unknown, extensive investigation will continue in this field. In the future new genome-scale technologies (i.e. deep sequencing and SNP chips) will be utilised in the search for unidentified variants that are involved in the development of familial breast cancer. The identification of new predisposing factors is challenging, but characterisation of the known risk factors will also be important. Therefore, more populations should be analysed, genomic rearrangements evaluated and a larger subset of candidate variants validated by functional assays. Many questions are to be answered: What genetic models and mechanisms explain the remaining familial risk of breast cancer? What environmental and genetic factors modify breast cancer risk? The fact that several genes that are known to associate with familial breast cancer susceptibility act in DNA damage response pathways suggests that the disease arises, at least in part, from a breakdown of genome stability control. This raises the question of whether breast is a particularly susceptible organ to carcinogenesis following a breakdown of the DNA damage response. Better understanding of pathways that lead to breast cancer formation and defining the molecular defects will bring benefits in the fight against breast cancer in terms of further progress in the prevention and treatment of the disease.
References


111


113


114


117


124


132


Original publications


Original publications have been reprinted with the permission of the copyright holders, Elsevier (I, IV) and BMJ Publishing Group (II, III).

The original publications are not included in the electronic version of this dissertation.
1005. Ilmastoiminen, Sari (2009) Fall accidents and exercise among a very old home-dwelling population
1006. Westerlund, Tarja (2009) Thermal, circulatory, and neuromuscular responses to whole-body cryotherapy
1008. Kuisma, Mari (2009) Magnetic resonance imaging of lumbar degenerative bone marrow (Modic) changes. Determinants, natural course and association with low back pain
1010. Löfgren, Johan (2009) Genetic polymorphisms in collectins and Toll-like receptor 4 as factors influencing susceptibility to severe RSV infections and otitis media
1013. Töri, Sami (2009) Factors affecting outcome after primary intracerebral hemorrhage
Sanna-Maria Karppinen

THE ROLE OF BACH1, BARD1 AND TOPBP1 GENES IN FAMILIAL BREAST CANCER