Department of Obstetrics and Gynecology
and
Department of Biosciences, Division of Genetics,
University of Helsinki

BREAST CANCER-PREDISPOSING GENES IN FINNISH BREAST AND OVARIAN CANCER FAMILIES

Paula Vehmanen

Academic Dissertation
To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in the auditorium 2041 of the Viikki Biocentre on June 15th, 2001 at 12 noon.

Helsinki 2001

Supervised by
Docent Heli Nevanlinna, Ph.D.
Department of Obstetrics and Gynecology
University of Helsinki

Reviewed by
Docent Jukka Partanen, Ph.D.
Department of Tissue Typing
Finnish Red Cross Blood Transfusion Service
Helsinki

Docent Minna Nyström-Lahti, Ph.D.
Department of Biosciences
Division of Genetics
University of Helsinki

Official opponent
Research Professor Lauri A. Aaltonen, MD, Ph.D.
Finnish Cancer Institute
University of Helsinki

ISBN 952-91-3495-9 (printed version)
ISBN 952-10-0004-X (pdf version)

Yliopistopaino
Helsinki 2001

TABLE OF CONTENTS

ORIGINAL PUBLICATIONS

This thesis is based on the following original publications that are referred to in the text by their Roman numerals.

I **Vehmanen P***, Friedman LS*, Eerola H, Sarantaus L, Pyrhönen S, Ponder B, Muhonen T, Nevanlinna H: A low proportion of *BRCA2* mutations in Finnish breast cancer families. *American Journal of Human Genetics 60: 1050-1058, 1997*

II **Vehmanen P***, Friedman LS*, Eerola H, McClure M, Ward B, Sarantaus L, Kainu T, Syrjäkoski K, Pyrhönen S, Kallioniemi OP, Muhonen T, Luce M and Nevanlinna H: A low proportion of *BRCA1* and *BRCA2* mutations in Finnish breast cancer families: evidence for additional susceptibility genes. *Human Molecular Genetics 6: 2309-2315, 1997*

III Kainu T*, Juo SH, Desper R, Schäffer AA, Gillanders E, Rozenblum E, Freas-Lutz D, Weaver D, Stephan D, Bailey-Wilson J, Kallioniemi OP, Tirkkonen M*, Syrjäkoski K, Kuukasjärvi T, Koivisto P, Karhu R, Holli K, Arason A*, Johannesdottir G, Bergthorsson JT, Johannsdottir H, Egilsson V, Barkardottir RB, Johannsson O*, Haraldsson K, Sandberg T, Holmberg E, Grönberg H, Olsson H, Borg Å, **Vehmanen P***, Eerola H, Heikkilä P, Pyrhönen S, Nevanlinna H: Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proceedings of the National Academy of Sciences of the United States of America 97: 9603-8, 2000*

* These authors contributed equally to the respective study

In addition, unpublished data is presented (study IV).

ABBREVIATIONS

ASO	allele-specific oligonucleotide
bp	base pairs
BRCA1	breast cancer susceptibility gene 1
BRCA2	breast cancer susceptibility gene 2
CI	confidence interval
CGH	comparative genomic hybridization
cM	centiMorgan
del	deletion
DNA	deoxyribonucleic acid
dsDNA	double-stranded deoxyribonucleic acid
HA	heteroduplex analysis
HLOD	heterogeneity LOD score
HRAS	Harvey rat sarcoma viral oncogene homolog
JUN	v-jun avian sarcoma virus 17 oncogene homolog
kb	kilo base pairs
kD	kilodalton
LOD	logarithm of the odds
LOH	loss of heterozygosity
Mb	megabase
mRNA	messenger ribonucleic acid
OMIM	Online Mendelian Inheritance in Man
p	chromosome short arm
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PTT	protein truncation test
q	chromosome long arm
RNA	ribonucleic acid
SSCP	single-strand conformation polymorphism

In addition, one-letter abbreviations of nucleotides and one-letter abbreviations of amino acids are used

REVIEW OF THE LITERATURE

1. FAMILIAL BREAST AND OVARIAN CANCER

1.1 Familial breast cancer

Hereditary (familial) cancer essentially means an inherited predisposition to cancer and clustering of disease within families is typical of such a dominantly inherited trait. The observation that breast cancer runs in families was already documented in the 18th century [1]. The highest risks of breast cancer are observed among first degree relatives of breast cancer patients [2]. Risk becomes higher when the age at diagnosis of the proband declines and when bilateral disease is involved [3]. The presence of male breast cancer is also considered as evidence of familial breast cancer [4]. This disease is not completely penetrant among susceptible persons; expression depends on gender, age and nongenetic risk factors. Overall, 4-10% of the patients with breast cancer show features of the hereditary breast cancer syndrome [5-7]. In Finland, there are over 3000 new cases of breast cancer diagnosed annually (during the years 1990-1997). Given 4-10% frequency of susceptibility allele, up to 300 cases a year may be explained by mutations in susceptibility genes (*1*).

1.1.1 Genes that predispose to breast cancer

Several studies have been carried out to identify breast cancer predisposing genes. Based on a segregation analysis in a large sample of families, Newman and colleagues concluded that disease clustering was due to an autosomal dominant allele. The highly penetrant susceptibility allele was estimated to be rare in the general population (frequency 0.0033) [7]. The cumulative lifetime risk of breast cancer for gene carriers is predicted to be high (82-92%) while for noncarriers risk is 8-10% [7, 8]. The major loci for breast cancer predisposition are *BRCA1* (p. 10) and *BRCA2* (p. 14) while other loci contribute to a lesser extent (Table 1) (*2*). Studies to identify deleted genomic regions in tumor tissue have been important in identifying breast cancer predisposing genes. According to the two-hit model of cancer development, hereditary cancers arise as a result of a germline mutation in a recessive tumor suppressor gene. Mutation is followed by the somatic deletion of the wildtype allele [9].

High frequency of LOH at the tumor suppressor gene *TP53* region (17p13.1) has been identified in breast cancer samples [10]. The gene product of the *TP53* gene, the p53 protein, prevents the cell from completing the cell cycle if DNA is not properly replicated. The p53 protein also triggers programmed cell death (apoptosis) if the damage to the cell is too great to be repaired. The overexpression of p53 in primary invasive human breast cancers is caused by a mutation in the *TP53* gene in a region that is highly conserved [11]. However, germline mutations in *TP53* account for only a small portion of familial breast cancer. The *TP53*-associated familial breast cancer is typically one feature of the Li-Fraumeni syndrome [12, 13].

The *PTEN* (a.k.a. *MMAC1*) locus mutated in Cowden syndrome displays frequent LOH in breast carcinomas and increased risk of breast cancer is one feature of this syndrome [14]. Somatic mutations are detected in both breast cancer cell lines and primary cancers. Germline mutations have not been identified in breast cancer patients [15] or in members of families with multiple breast cancer patients [16]. Moreover, linkage studies suggest that *PTEN* involvement in familial breast cancer is minor [17].

Breast cancer is also a feature of ataxia-telangiectasia, a recessive degenerative disease of childhood. The *ATM* (ataxia-telangiectasia mutated) protein is involved in detecting DNA damage and interrupting the cell cycle when damage is found. The *ATM* gene mutations are associated with increased risk of breast cancer [18] but these mutations do not contribute to a large extent to familial breast cancer [19-21].

The tumor susceptibility gene 101 (*TSG101*) involvement in human breast cancer has been implicated based on a high frequency of somatic mutations in primary breast carcinomas [22]. However, germline mutations have not been detected in this gene in breast cancer patients [23]. In addition, the Harvey rat sarcoma viral oncogene homolog (*HRAS*) locus has been implicated in breast cancer. Certain rare alleles at the *HRAS* locus may be associated with susceptibility to breast cancer but clear involvement has not been established [24]. Along with *TSG101* and *HRAS*, genes on 11q and/or 22q may be involved in breast cancer tumorigenesis. The observed number of breast cancer cases is significantly higher than expected among the carriers of the constitutional reciprocal translocation t(11; 22)(q23; q11) which is the most frequent translocation observed in man [25].

Table 1 *Genes predisposing to breast cancer*

Gene	**Location**	**OMIM title *(2)***	**OMIM entry *(2)***
BRCA1	17q21	BREAST CANCER, TYPE 1	113705
BRCA2	13q12	BREAST CANCER 2, EARLY-ONSET	600185
TP53	17p13	TUMOR PROTEIN p53	191170
ATM	11q22	ATAXIA-TELANGIECTASIA	208900
PTEN ***MMAC1***	10q22	PHOSPHATASE AND TENSIN HOMOLOG MUTATED IN MULTIPLE ADVANCED CANCERS 1	601728

1.2 Familial ovarian cancer

A certain number of ovarian cancer cases are familial. Approximately 10% of epithelial ovarian carcinoma cases are associated with inheritance of an autosomal dominant genetic trait [26]. Like breast cancer, ovarian cancer shows an earlier age of onset in hereditary settings [27]. Familial ovarian cancer is often observed in families with breast cancer, but also without any associated cancers as site-specific ovarian cancer [28]. Ovarian cancer is also a feature of the hereditary nonpolyposis colorectal cancer (HNPCC) syndrome [29, 30].

Table 2 *Genes predisposing to ovarian cancer*

Gene	Location	OMIM title *(2)*	OMIM entry*(2)*
BRCA1	17q21	BREAST CANCER, TYPE 1	113705
BRCA2	13q12.	BREAST CANCER 2, EARLY ONSET	600185
MSH2	2p22-21	COLON CANCER, FAMILIAL NONPOLYPOSIS, TYPE 1	120435
MLH1	3p21	COLON CANCER, FAMILIAL NONPOLYPOSIS, TYPE 2	120436

1.2.1 Genes that predispose to ovarian cancer

Ovarian cancer is a heterogeneous disease with environmental factors contributing to incidence. The gene frequency for the ovarian cancer predisposing gene in the general population is estimated to be 0.0028 and penetrance 50% by 70 years [31]. Presently, a few such predisposing genes are known. The hereditary breast and ovarian cancer syndrome is linked to the *BRCA1* (p. 10) locus and to a lesser extent, to the *BRCA2* (p. 14). locus The subtype of HNPCC syndrome with increased risk of ovarian cancer is due to mutations in two genes, *MSH2* or *MLH1*. These genes encode proteins involved in DNA mismatch repair [30]. However, a recent population-based study indicates that germline mutations in these four loci contribute to only a minority of cases of early-onset epithelial ovarian cancer [32]. Thus, the major ovarian cancer locus, if existing, has yet to be found.

2. THE BREAST CANCER SUSCEPTIBILITY GENE 1 (*BRCA1*)

2.1 The *BRCA1* gene is large with two distinct promoters

The *BRCA1* gene is located at 17q21 [33]. In 1994, a strong candidate for the *BRCA1* gene was identified [34] and predisposing mutations were detected in several breast cancer and ovarian cancer kindreds [35]. The 24 exons of the *BRCA1* span an 81-kb region with an unusually high density of Alu repetitive DNA (41.5%) [36]. *BRCA1* has two promoters generating two distinct transcripts: alpha and beta [37]. Neither of the promoters contains a TATA box [37]. Instead, a 36 bp positive regulatory region mediates the promoter activity [38]. Both promoters also contain several transcription factor-binding sites at their 5' flanking regions [37]. The promoter alpha is shared with the adjacent gene (called *NBR2*) and is bi-directional. The two transcripts differ having either copy of the duplicated exon 1: exon 1A is first exon for transcript alpha and exon 1B for transcript beta. Expression of both transcripts is detected in normal and cancer cell lines and in primary breast and ovary tumor tissues. The BRCA1a transcript is the form expressed in the mammary gland and the BRCA1b transcript in the placenta [37].

2.2 The BRCA1 protein is nuclear and has several functional domains

The *BRCA1* gene product is a nuclear phosphoprotein of 1863 amino acids [34] (Figure 1). This 220-kD protein is localized in the endoplasmic reticulum-Golgi complex and in tubes invaginating the nucleus [39]. During mitosis, it localizes with the centrosome [40]. The BRCA1 protein has two nuclear localization signals (NLS) in exon 11, of which only one (NLS1) is required for proper nuclear localization [41]. The RING-finger domain [34] situated at the amino-terminal portion of the protein is a member of the

extensive family of zinc-binding proteins implicated in cell growth, differentiation and development [42]. The C-terminal region of the BRCA1 contains a protein-binding site denoted as the BRCT module. The BRCT is a globular domain found in proteins involved in repair and cell cycle control [43].

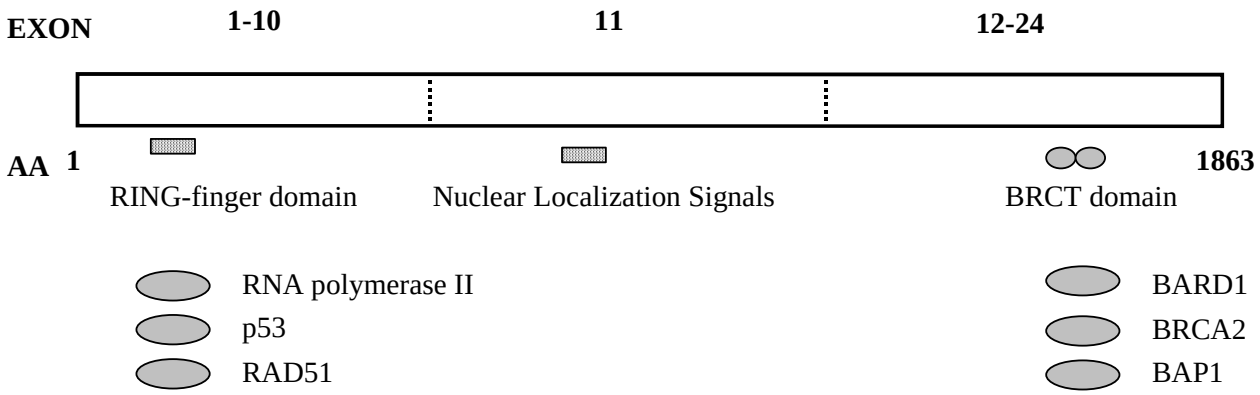


Figure 1 *A schematic illustration of the BRCA1 protein. The corresponding exons are shown above the figure. Functional domains and proteins that bind to these domains are shown below the figure. AA = amino acid*

The wildtype BRCA1 protein inhibits the growth of breast and ovarian cancer cell lines [44] and binds to a wide array of cellular proteins (Figure 1). The BRCT module and the RING-finger domain of the BRCA1 protein mediate most interactions with other cellular proteins. These interacting proteins include the DNA repair protein Rad51 [45] and the tumor suppressor protein p53 [46]. The BRCA1 protein is also assumed to be involved in transcription, because it binds the RNA polymerase II holoenzyme [47]. The BRCA1-associated RING domain protein, BARD1 [48], is thought to be a critical factor in BRCA1-mediated tumor suppression [49]. Moreover, BRCA1 binds BRCA2 (p. 14) [50] and the BAP1-proteins. The BAP1 protein enhances BRCA1-mediated inhibition of breast cancer cell growth [51]. The BRCA1 protein eventually induces the expression of the cyclin-dependent kinase inhibitor *p21* $^{Waf1/CiP1}$ which leads to cell cycle arrest [52]. The second target gene is the DNA-damage responsive gene *GADD45* (Growth Arrest- and DNA Damage-inducible) [53]. Induction of the *GADD45* by the BRCA1 protein then triggers apoptosis [54].

In breast and ovarian cancer cell lines and in breast cancer biopsies BRCA1 is localized mainly in the cytoplasm [34, 55, 56]. Location of the naturally occurring 110-kD splice variant BRCA1-delta11b (missing NLS1) is also cytoplasmic [41, 57]. The dislocation of the BRCA1 protein in breast cancer cells may be due to a defect in the NLS mediated pathway of nuclear import [58].

2.3 *BRCA1* expression

The greatest levels of *BRCA1* expression are observed during DNA synthesis and mitosis [34, 59]. A high level of expression is also detected in meiotic germ cells, while premeiotic germ cells express little or no *BRCA1* mRNA [60]. In the mammary gland, the expression is induced during puberty, pregnancy and following treatment with 17 beta-estradiol and progesterone [61]. *BRCA1* expression is initially upregulated followed by a reduction below basal levels in response to treatment with DNA damaging agents. This reduction of *BRCA1* expression is dependent on the presence of wildtype p53 [62].

In cancer cell lines, the amount of the *BRCA1* mRNA is relatively low [63]. Expression is reduced or undetectable in breast carcinomas [64, 65] and the mRNA level is markedly decreased during the transition from carcinoma in situ to invasive cancer [66]. The steady state level of the *BRCA1* mRNA is elevated by estrogen and blocked by antiestrogens [67] in breast cancer cell lines. This effect is mediated through the estrogen receptor [68].

2.4 Pathological variants of the *BRCA1* gene

2.4.1 There is a wide spectrum of BRCA1 germline mutations

Somatic mutations involving the *BRCA1* gene have been detected in primary breast and ovarian tumor samples at a low frequency [69]. Contrary to somatic mutations, there are multiple disease-causing mutations in the germline. The number of distinct mutations exceeds 900. Some 400 of these are found recurrently. All mutation types are present, but the majority of mutations are nucleotide substitutions followed by small deletions (*3,4*). Large germline deletions, which appear to result from homologous recombination between closely related Alu repeats [70-72], have also been detected. Carriers homozygous for *BRCA1* mutations have yet to be found, suggesting that homozygosity is lethal in humans. Segregation of two different frameshift mutations in a single family has been detected; however, no individual carried both mutations [73].

2.4.2 Germline mutations are found at high frequency in families with multiple affected members

Estimates obtained from population-based studies suggest that the frequency of the mutated allele in the general population is low (0. 0006) [74]. The proportion of breast cancer cases diagnosed in the general population below the age of 40 due to *BRCA1* mutation is approximately 5%. The corresponding estimate for ovarian cancer cases is 5.7% [74]. Of families with multiple breast cancer cases, 45-52% are thought to be *BRCA1*-associated [75-77]. Of the families featuring the hereditary breast and ovarian cancer syndrome, *BRCA1* is estimated to explain up to 76-90 % of families [75, 77, 78]. In certain ethnic groups mutations appear with a higher frequency, for instance, the 185delAG mutation is present at 1% frequency among Ashkenazim [79].

2.4.3 BRCA1 germline mutations predispose to several cancer types

Obtained estimates of the cumulative risk of breast cancer (54% by 60 years) and ovarian cancer (30% by age 70) are derived from families showing genetic linkage to *BRCA1* [80]. The risk of ovarian cancer is presumed to differ with specific cancer-disposing mutation. A significant correlation is observed between the location of the mutation and the ratio of breast to ovarian cancer incidence within the family. Mutations in the 3-prime third of the gene are associated with a lower proportion of ovarian

cancer [81, 82]. There is also experimental evidence for the relative susceptibility to ovarian versus breast cancer. Ovarian cancer cell growth is not affected by *BRCA1* mutations in the 5-prime portion of the gene but is inhibited by 3-prime *BRCA1* mutations [83]. *BRCA1* mutation also predisposes carriers to prostate and colon cancer and the cumulative risks by 70 years are 8% and 6%, respectively [84]. It has been noted that the age at breast cancer diagnosis is earlier among those mutation carriers who carry very long trinucleotide (GAG) repeats in the androgen receptor gene. This suggests that pathways involving androgen signaling may affect the risk of *BRCA1*-associated breast cancer [85]. The variable penetrance of *BRCA1* suggests that several genetic and nongenetic factors play a role in tumorigenesis. In addition, the *HRAS* gene minisatellite polymorphism has been proposed as one possible genetic modifier of penetrance. The risk of ovarian cancer is 2.11 times greater for those *BRCA1* mutation carriers harboring one or two rare HRAS alleles compared to carriers with only common alleles. Susceptibility to breast cancer does not appear to be affected by the presence of rare *HRAS* alleles [86].

2.4.4 The survival of mutation carriers does not seem to differ from sporadic cases

Most studies have not found a significant difference in the survival of breast cancer patients with *BRCA1* mutations when compared to controls [87, 88] but also conflicting reports have been published [89, 90]. Results are presumed to be conflicting at least in part because of methodological problems. Rubin and coworkers suggested a favorable clinical course in primary epithelial ovarian cancer patients with *BRCA1* mutation when compared to sporadic ovarian cancers [91]. However, other authors have found contradicting results suggesting that survival in familial ovarian cancer cases is worse than in sporadic cases, whether or not a *BRCA1* mutation is present [92]. Johannsson and colleagues have suggested no difference in the survival of ovarian cancer patients with *BRCA1* mutations when compared to controls [87].

2.4.5 Somatic deletions encompassing the BRCA1 locus are frequent in breast cancer tissue

Loss of heterozygosity on chromosome 17 is common genetic alteration in breast cancer tumors [93] and LOH is frequently observed at the *BRCA1* region [94]. In both tumors from families linked to the *BRCA1* gene [95] and individuals carrying *BRCA1* mutations [96], the observed allele loss involves the wildtype allele. Genome-wide CGH analysis of breast cancers from *BRCA1* mutation carriers shows that losses of 5q, 4q, 4p, 2q and 12q being significantly more common than in the control group [97]. The loss of the whole chromosome 17 is also a relatively frequent event in ovarian tumors. LOH of the *BRCA1* gene region is observed in 50% of tumors [98]. LOH of the *BRCA1* region is also detected in endometrial cancer [99] and gastric carcinomas [100]. The molecular pathogenesis and histological phenotype of tumors from *BRCA1* mutation carriers differ from mutation-negative cases as well. Breast cancers occurring in *BRCA1* mutation carriers have higher levels of p53 expression and a higher proliferation rate [101]. *BRCA1* mutation carrier cell line shows aneuploidy, an acquired *TP53* mutation with the loss of the wildtype allele and an acquired homozygous deletion of the *PTEN* gene [102].

2.5 Murine *Brca1*

The mouse homologue of *BRCA1* maps to chromosome 11, an area with extensive homology to the human chromosome 17 [103]. Null mice die early in embryonic development. The mutant embryos are poorly developed and show a variety of neuroepithelial defects [104] and reduced cell proliferation [105]. Heterozygous mice have defective cell cycle control and chromosomes contain multiple functional

centrosomes [106]. Excision of exon 11 of the *Brca1* causes abnormal ductal development and increased apoptosis in epithelial cells [106]. Disruption of the *Brca1* causes genetic instability and triggers further alterations, including inactivation of p53 [107]. After about 1 year of age, heterozygous *Brca1* female mice show no evidence of cancer [104, 105].

3. THE BREAST CANCER SUSCEPTIBILITY GENE 2 (*BRCA2*)

3.1 The *BRCA2* locus is situated at 13q12

Stratton and coworkers presented strong evidence against linkage to *BRCA1* in families with at least one case of male breast cancer [108]. Strong evidence of linkage to the *BRCA2* region was shown in families characterized by multiple cases of male breast cancer [109, 110]. The *BRCA2* gene was localized to region 13q12 [111]. *BRCA2* is composed of 27 exons (the first exon is noncoding) distributed over roughly 70 Kb of genomic DNA [112]. Simple sequence variation is frequent both in the coding and noncoding region [113].

3.2 The BRCA2 protein has no similarity to other known proteins

The gene product of the *BRCA2* gene is a 3418 amino acid (390-kD) nuclear protein and it bears no similarity to other known proteins [112] (Figure 2). The 5-prime portion of the exon 11 contains repeated sequences termed BRC motifs that are conserved in mammalian BRCA2 proteins [114]. The BRC module consists of four 30-residue repeats [115]. A sequence in exon 3 shows sequence similarity to the activation domain of the *JUN* proto-oncogene. The activation potential within exon 3 is under negative control by two inhibitory regions present immediately on either side of the exon 3 [116]. The human RAD51 binds directly to the BRC repeats [117]. Brca2 physically and functionally interacts with p53 [118] and it has been shown to acetylate the histones H3 and H4 when bound to the P/CAF transcriptional co-activator [119].

3.2.1 BRCA2 is expressed in several tissues

Expression *BRCA2* mRNA is observed in several tissues: breast epithelium, thymus, testis, placenta, lung, ovary and spleen [112, 120]. Expression is initiated before DNA synthesis [121] and it is associated with proliferation both in normal [122] and tumor-derived breast epithelial cells [123]. The steady state levels of *BRCA1* and *BRCA2* mRNAs are coordinately elevated by estrogen but not progesterone and blocked by antiestrogens [67]. *BRCA2* mRNA undergoes differential splicing giving rise to a variant protein BRCA2a. Both BRCA2 and BRCA2a are present at high levels in thymus and testis but moderate levels in the mammary gland and prostate suggesting that they have a role in the development and differentiation of these tissues [124].

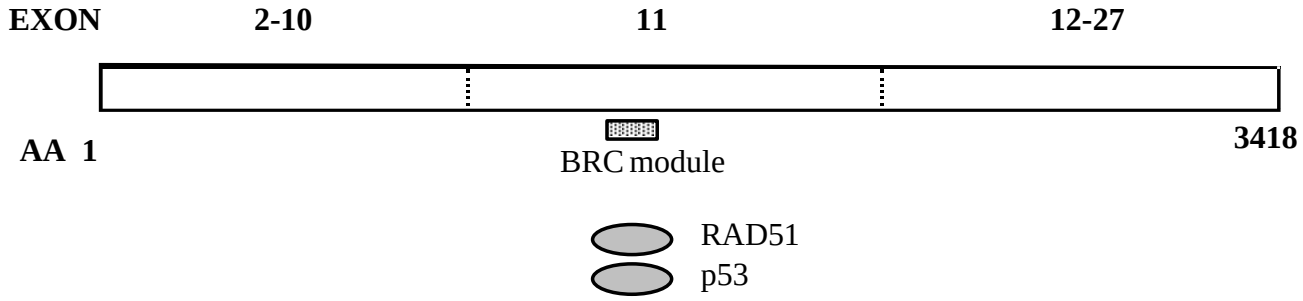


Figure 2 *Schematic representation of the BRCA2 protein. The functional domain, the BRC module and proteins binding to the module are presented below the figure. The corresponding exons are presented above the figure.AA=amino acid*

3.3 Pathological variants of the *BRCA2* gene

3.3.1 The germline mutation spectrum of BRCA2 resembles that of BRCA1

Wooster and colleagues reported first mutations in the *BRCA2* gene [120] followed by several other research groups [125, 126]. Somatic mutations in the *BRCA2* gene are rare [127-129]. On the contrary, the number of germline mutations is high. There are over 900 mutations, of which 620 have been found only once. The mutation spectrum closely resembles that of *BRCA1* (*3*,*4*). However, large deletions are more rare in the *BRCA2* gene than in *BRCA1*, with only one case reported so far [130]. No homozygotes for *BRCA2* mutations have been observed, although certain mutations have an exceptionally high carrier frequency [131].

3.3.2 BRCA2 is associated with fewer breast cancer cases than BRCA1

In families with multiple breast cancer cases, the disease is estimated to be due to mutations in the *BRCA2* gene in 32% of families [77]. Mutations in the *BRCA2* gene account for 4% of male breast cancer cases without family history [132] and for 14% of cases with family history [125]. Population-based estimates predict that of early-onset breast cancer cases 3.4% carry *BRCA2* mutations [133]. As in *BRCA1*, certain mutations are enriched: the 999delTCAAA mutation frequency is 0.6 % in the Icelandic population [131], 10.4% in female and 38% in male breast cancer cases [134]. Overall, *BRCA2* is associated with fewer breast cancer cases than *BRCA1* [135]. Recently, it was proposed that common variants might confer an increased risk of breast cancer also. The HH homozygotes for the polymorphism N372H (exon 10) have a 1.31-fold greater breast cancer risk than the NN homozygotes. Moreover, the HH variant appears to affect fetal survival in a sex-dependent manner [136].

3.3.3 BRCA2 mutations predispose to cancers in several organs

The estimated cumulative risk of breast cancer in families showing genetic linkage to *BRCA2* is 80% in females and 6% in males by the age of 70. These risks are similar to the ones estimated for *BRCA1* mutation carriers. However, there is some evidence of a lower risk of breast cancer in *BRCA2* carriers under 50 years of age [77, 137]. A significant excess of ovarian cancer is observed in gene carriers and the

ovarian cancer risk is 27% by the age of 70 [77]. Variation in phenotypic risk has been observed in mutation carriers. Truncating mutations in carriers with the highest risk of ovarian cancer relative to breast cancer are clustered in a region of approximately 3.3 Kb in exon 11 [138]. Laryngeal, prostate and pancreatic cancers are also frequent in individuals carrying *BRCA2* mutations [134, 139, 140]. Several studies have shown that despite their younger age at presentation, breast cancer patients harboring *BRCA2* mutations have a similar prognosis when compared with sporadic breast cancer patients [88, 141, 142].

3.3.4 Loss of heterozygosity involving the BRCA2 region is seen in various tumor types

In breast carcinomas, LOH involving the *BRCA2* region is found in 51% of studied samples [94]. In tumors from families showing genetic linkage to *BRCA2*, LOH at the *BRCA2* region is seen frequently. In most of the cases, there is a confirmed loss of the wildtype allele [143]. In *BRCA2* mutation carriers, LOH at the *BRCA2* region is observed in tumors of the prostate, ovary, cervix, colon, male breast, and ureter. In addition, breast tumors from female carriers demonstrate loss of the wildtype chromosome [144]. A CGH-based analysis revealed that tumors from *BRCA2* mutation carriers exhibited a significantly higher number of chromosomal aberrations than sporadic tumors. The most common genetic changes were gains of 8q, 20q and 17q and losses of 13q and 6q [145]. The histology of *BRCA2* mutation-positive cases differs from sporadic cases and from *BRCA1* mutation-positive ones. The findings suggest that breast cancers in these three groups have a different natural history [146].

3.4 Murine Brca2

The mouse *Brca2* maps to the distal end of chromosome 5 and it is expressed in thymus, testis and ovary [147]. In mammary epithelial cells *Brca2* expression is induced during puberty, pregnancy and in response to glucocorticoids. Both the *Brca1* and *Brca2* genes are highly expressed in rapidly proliferating cells, with the expression peaking before DNA replication begins [148]. Cells that harbor truncated *Brca2* spontaneously incur chromosomal rearrangements and DNA breaks during division [149]. These cells also exhibit hypersensitivity to DNA damage [150]. The majority of homozygous mutant mice die prenatally or perinatally. The mutants that survive to adulthood die with thymic lymphomas between 12 and 14 weeks of age. Heterozygote mice are free of tumors [151].

4. THE FINNISH POPULATION HAS A UNIQUE HISTORY AND GENE POOL

Finns are described as outliers among European populations with the gene pool showing a unique pattern. The genetic isolation of Finland is due to both geographical and linguistic reasons. The Finnish language is a member of the Finno-Ugric language family [152]. Genetically Finns are close to Indo-European speakers with an estimated 90% of genes being European, assumed to be the outcome of long close contact with Germanic populations [153].

Analysis based on mitochondrial mutations suggests that the founding population went through a bottleneck approximately 4000 years ago, presumably when populations using agriculture and animal husbandry arrived in Finland [154]. There is not a full agreement between historical data and genetic data on the origin of Finns. The single origin model proposes a single route of immigration over the Gulf of Finland [155, 156]. A dual-origins model based on HLA genes [157] and Y chromosome markers [158] proposes that two different groups settled in Finland: one from the east and the other from the south via the Gulf of Finland [159]. However, both models agree that the population has remained an isolate and that the population expansion started 2000-2500 years ago in southern and western parts of the country. The number of founders was small and this group founded the population known today as Finns (5.2 million). The northern and eastern parts of the country were settled mainly after the 16th century. This expansion was characterized by sub-isolates with a relatively small number of individuals. Genetic drift worked in these sub-isolates modifying the gene pool resulting in enrichment of mainly recessive diseases denoted as The Finnish Disease Heritage. Originally, the concept consisted of some 10 diseases [160] . Subsequently it was widened to comprise some 30 diseases [156]. The geographical distributions of rare alleles show a distinct pattern according to time the of origin [160].

Recently expanded populations such as Finns are suitable to map single disease genes affected by recent mutations. This characteristic has been utilized in several linkage disequilibrium studies. A simplified population genetics model was introduced by Hästbacka and co-workers [161] and since, several genes have been localized [156]. In complex diseases, the localization of disease genes has been aided by using families either from only the late settlement areas or from a sub-isolate [162].

5. METHODS

5.1 Methods for gene localization

Gene localization is the first step in positional cloning. Genetic linkage analysis is the major technique for the initial localization of genes and several statistical approaches exist. Novel gene localization methods are constantly being developed to accompany linkage analysis.

5.1.1 Genetic markers and maps

A genetic linkage map shows the relative locations of specific DNA markers along the chromosome. Polymorphic DNA sequence differences between individuals are used as such markers. Restriction fragment length polymorphisms were the first markers used in genetic mapping. These are sequence variations in DNA at sites that can be cleaved by restriction enzymes [163]. The next generation of markers was based on short repeated sequences that vary in the number of repeated units. These minisatellite markers were detected by electrophoresis [164]. Weber and May introduced microsatellites in which the repeated sequences are 2-4 bases long. These short tandem repeats are at present the standard method for genetic mapping. PCR is performed with a fluorescent primer and products are directly analyzed on an automated instrument [165]. The latest generation of markers, single nucleotide polymorphisms, was recently introduced. Densely spaced biallelic markers are expected to advance linkage studies [166].

The human genetic linkage map is constructed by observing how frequently two markers are inherited together. Recombination frequency provides an estimate of the distance between two markers. Distances between markers are measured in terms of centiMorgans (cM). Distance over which probability of recombination is 1% is known as 1 cM. A genetic distance of 1 cM is roughly equal to a physical distance of 1 million bp (1 Mb). The probability that a recombination occurs between two loci is the recombination fraction, theta. Thus, theta is a measure of genetic distance between two loci. Inherited disease is located on the genetic map by following the inheritance of a DNA marker present in affected individuals [167].

5.1.2 Statistical methods used for gene localization

Statistical methods applied for linkage analysis are often divided into parametric and nonparametric methods. The parametric methods involve testing whether the inheritance pattern fits a specific model for a trait-causing gene. Parametric methods require the inheritance model to be defined and are sensitive to misspecification of the model [168, 169]. Nonparametric methods do not require the inheritance model to be known. In addition, nonparametric methods allow analysis in less optimal situations [170]. Before actual genotyping is performed, the power to detect linkage in a given set of families can be estimated by simulation [171]. In addition to genetic linkage, one can study allelic association.

5.1.2.1 The LOD-score method is the standard parametric method of linkage analysis
The LOD score method is the standard parametric method when analyzing human pedigrees. The LOD scores are obtained by calculating the overall likelihood of the data under two alternative assumptions: that the two loci are linked with the given recombination fraction value (theta) and that they are unlinked. Likelihood is similar to probability and it is dependent on the distance between the two loci. The ratio of those two likelihoods is called a likelihood ratio and it indicates the odds for or against linkage. The logarithm to the base 10 of this ratio is the LOD score. This procedure is repeated for a selection of theta values. The theta value that gives the highest LOD score represents the best estimate of the distance between the two loci. The power to detect linkage depends on the number of informative meioses, almost regardless of family structure. An LOD score of three or over is considered as conclusive [172].

If linkage cannot be established based on the given set of pedigrees, it is studied whether a subset of the pedigrees shows collectively evidence of linkage. One can explicitly allow genetic heterogeneity by inclusion of an admixture parameter (alpha) specifying the proportion of linked families. The LOD score can be constructed under heterogeneity depending on both alpha and theta. Maximizing the LOD score provides estimates of both alpha and theta [173]. The probability of obtaining false evidence of linkage depends on several parameters. Such parameters are the correctness of the marker allele frequencies, pedigree structure, number of the families in the sample and the threshold of accepting linkage. These effects are most pronounced in small pedigrees when essential individuals are unavailable for typing. Difficulties in linkage analysis include incomplete information (unfavorable structure of the pedigrees and uninformativeness of the genetic markers) as well as incomplete penetrance of the susceptibility gene and presence of sporadic cases in the families [170].

Multipoint linkage is used for more precise gene localization by typing a series of markers. Multipoint offers added power, especially with loci with poor information content or with small samples. LOD scores are computed for different locations and results are presented in the form of a graph. The LOD score is plotted against the position along the chromosome. The evidence in favor of particular location is expressed in terms of “relative odds” of a particular order. When odds of 1000:1 or greater are obtained, it is conventionally regarded as conclusive [167]. Multipoint includes some computational difficulties when the number of loci increases. In addition to localization, multipoint linkage is used for constructing maps of genetic markers by marker typings on large reference families [174].

5.1.2.2 Nonparametric methods of linkage analysis
Nonparametric methods test whether the inheritance pattern deviates from expectation under independent assortment. These methods do not depend on a model but on counting marker alleles or haplotypes between affected relatives. Linkage is measured by allele sharing. Typically affected individuals are sibling pairs but analyses can be extended to other types of relative pairs [175]. The APM (Affected Pedigree Member) approach enables the use of pedigrees with more than two affected individuals [176].

5.1.2.3 Allelic association
Allele-sharing methods involve direct testing of candidates by looking for associations between particular alleles and disease by comparing allele frequencies in affected individuals and matched

controls. The choice of appropriate control group is essential. The observed high-risk allele does not need to be the disease causing allele itself. A more likely explanation for association is linkage disequilibrium (LD). LD occurs when a particular marker allele lies so close to the disease susceptibility allele that these alleles co-segregate over many generations. The same allele will be detected in affected individuals in multiple apparently unrelated families [177]. Linkage disequilibrium can arise in a number of different ways. The most common situations are after recent admixture of two previously isolated populations [178] or, when a large population has expanded rapidly from a relatively small founder population [161]. When population association has been established, analysis may be extended to families. The most widely used test, TDT (transmission disequilibrium test) evaluates the transmission of associated marker allele from a heterozygous parent to the affected offspring [179].

5.1.3 Novel methods for gene localization

Novel methods test either the differences in genomic DNA or variations in gene expression. Comparative genomic hybridization produces a map of DNA sequence copy number as a function of chromosomal location throughout the entire genome. Regions of gains or losses of DNA sequences are seen as changes in the ratio of the intensities of the used fluorochromes along the target chromosome [180]. Genome-scale LOH analysis is labor-intensive. Contrary to LOH, CGH allows the identification of chromosomal imbalances in a comprehensive manner, providing genome-scale information in a single experiment. With CGH, it is possible to identify recurrent chromosomal gains and losses as starting points for the characterization and isolation of pathogenetically relevant cancer genes. The power of CGH in identifying putative tumor suppressor loci has been demonstrated in Peutz-Jeghers syndrome where CGH in hamartomas pinpointed a recurrently deleted region. Subsequent targeted linkage analysis revealed a high-penetrance locus in 19p13.3 [181]. This novel locus encodes the serine/threonine kinase STK11 [182]. However, DNA gains and losses in tumor cells do not occur entirely at random. Models that relate tumor progression to the occurrence of DNA gains and losses have been developed to facilitate the distinction of primary events involved in cancer initiation from the numerous events that arise during cancer progression [183].

The multiple genetic alterations in cancer result in complex changes in the expression of many genes. Analysis of gene expression profiles of tumors using a DNA microarray enables the identification of up-regulated or down-regulated genes. The modification cellular phenotype by successive genetic alterations and the examination the resulting expression profiles may facilitate the identification of tumor suppressor genes [184]. The application of the DNA microarray technology may also facilitate gene identification when mutations result in changes in the mRNA level of the disease gene [185].

5.2 Methods for mutation detection

Changes in the genome of an individual can lead to a disease phenotype. Such changes can occur in a number of different ways, such as deletion, insertion, fusion, point mutation, or amplification of unstable sequences. Changes in DNA structure may occur both in germ cells and somatically. Methodologies used in screening for mutations differ in sensitivity and thus biases the types of mutations found. No currently available single technique can guarantee the identification of all types of mutations [186]. In addition, mutations of uncertain clinical significance may be identified. The clinical significance of a mutation can be determined by studying whether the mutation segregates with the

disease in a given family and whether the allele has a higher frequency in patients than in the general population. Ultimately the significance of the mutation can be revealed by direct protein function assays [187].

5.2.1 Screening methods for finding unknown mutations

The PCR method is often used for generation of the DNA fragments or is in itself used as part of the detection method. One set of methods relies on differences in electrophoretic mobilities between wildtype and mutant nucleic acids. The second group of methods is based on the cleavage of mismatches in heteroduplexes. The third group uses mismatch-binding proteins in heteroduplexes. Direct sequencing is considered as the gold standard method of mutation detection [188]. The dideoxy (Sanger) method for sequencing is the most widely used variant [189].

SSCP (single strand conformation polymorphism) is based on the electrophoretic mobility of single-stranded nucleic acids. Mobility depends not only on size but also on sequence and most single base changes can be detected as mobility shifts [190]. Using different conditions sensitivity can be increased [191]. In DGGE (denaturing gradient gel electrophoresis), dsDNA is electrophoresed through a gradient of increasingly denaturing conditions. Sequence differences result as differences in partial melting behavior of the DNA fragments [192]. In heteroduplex analysis (HA), the heteroduplex molecules that form between a mutant and a wildtype DNA strand are distinguished from homoduplex molecules upon gel electrophoresis [193]. Modifications and enhancements of the HA method have been developed that increase the sensitivity of detection of single-base pair alterations [194]. Chemical or enzymatic cleavage of mismatch is based on the formation of heteroduplexes between wildtype and mutant sequences in the sample. Strands are cleaved at the site of mismatch chemically [195] or enzymatically [196]. The protein truncation test (PTT) involves in vitro-coupled transcription and translation of PCR products, resolved by SDS-PAGE analysis. The size of the translation product pinpoints the position of the stop codon in the corresponding PCR product fragment. It should be noted that mutations located very close to the primer binding sites may escape detection. Designing overlapping sets of primers can alleviate this problem [197]. Mutations involving large genomic rearrangements can be detected by Southern blotting [198].

5.2.2 Diagnostic methods for finding known mutations

The simplest ways to screen for known mutations is to use naturally occurring restriction enzyme cleavage sites [163]. Artificial restriction sites may be introduced using mismatched primers [199]. The allele-specific amplification (ASA) method relies on two parallel reactions where only one allele of either mutant or wildtype is amplified, and several variants of this technique are available [188]. In allele-specific oligonucleotide (ASO) hybridization one uses probes which can detect single base substitutions in PCR-amplified DNA. An allele-specific oligonucleotide will only anneal to sequences that match it perfectly, a single mismatch being sufficient to prevent hybridization under appropriate conditions [200]. Oligonucleotide-ligation assay (OLA) is based on the covalent attachment of labeled reporter oligonucleotide to an immobilized target oligonucleotide. The sample DNA serves as the template. In the case of a mismatch no ligation will occur. The assay products are resolved electrophoretically from one another and from unligated probes under denaturing conditions [201]. Minisequencing has several variants, which are all based on the extension of the 3’ end of the primer by a single, complementary nucleotide [202, 203].

AIMS OF THE PRESENT STUDY

1. To determine the prevalence and spectrum of the *BRCA1* and *BRCA2* mutations in Finnish breast and ovarian cancer families (studies I and II).

2. To further characterize the most common *BRCA2* mutation in Finland, 999delTCAAA, using genealogical records and haplotype data (study IV, unpublished).

3. To find evidence of a third breast cancer-predisposing gene in Finnish breast cancer families (study III).

MATERIAL AND METHODS

1. ETHICAL IMPLICATIONS

Appropriate permission was obtained from the Ministry of Social Affairs and Health. Written informed consent was obtained from all patients and relatives donating blood samples for the genetic analysis. The DNA samples were coded and no data from the research work was provided to the patients or family members.

2. KINDREDS AND TUMOR SAMPLES

2.1 Kindreds included in the *BRCA1* and *BRCA2* mutation analyses (I,II)

One hundred patients with a positive family history were included in the study (Table 3). The patients were originally selected from the following three cohorts:
A. Breast cancer patients diagnosed below 40 years of age who were treated at the Department of Oncology of the Helsinki University Central Hospital (HUCH) during the years 1985–1993.
B. Breast cancer patients diagnosed with bilateral breast cancer who were treated at the Department of Oncology (HUCH) during the years 1985–1993.
C. Unselected breast cancer patients who were treated at the Department of Oncology (HUCH) between November 10th 1993 and April 30th 1995.

Patients were interviewed for family history of breast cancer and the patients with positive family history were included in the study. The genealogy of these patients was confirmed using church-parish registries. The diagnoses of the patients and relatives were confirmed using hospital records of the HUCH Department of Oncology and the Finnish Cancer Registry. In addition, DNA samples from 93 unrelated healthy individuals were used as controls. The study population and the control group were of Finnish (Caucasian) origin.

Table 3 *Breast and ovarian cancer cases in families included in mutation analyses*

Number and type of cancer cases in the family	**Number of families**
Four or more breast cancer cases	23
Four or more breast cancer cases and ovarian cancer	14
Three breast cancers	50
Three cancers, breast and ovarian	11
Two cancers, breast and ovarian	2
Total	*100*

2.2 Kindreds included in the haplotype analysis of the *BRCA1* and *BRCA2* regions (II)

Haplotype analysis was carried out in 12 families in which *BRCA1* or *BRCA2* mutations had not been detected. Six of the families had four or more cases of breast cancer and six had four or more cases of breast and ovarian cancer. All available individuals were genotyped.

2.3 Kindreds included in the haplotype analysis of the 999delTCAAA mutation families (IV, unpublished)

Ten Finnish kindreds carrying the 999delTCAAA mutation ascertained at four research centers were included in the study. Two of these families were from the cohort of 100 families described in the Material and Methods section 2.1. A total of 54 individuals were genotyped, of these, 28 were mutation carriers and 26 noncarriers. Mutation carrier status had been determined in all individuals by ASO hybridization. The birthplaces of grandparents to the proband were traced using church parish records.

2.4 Tumor samples used in the CGH analysis (III)

A total of 61 tumor specimens from 37 non-*BRCA1/BRCA2* breast cancer families were selected for the CGH analysis from the Helsinki University Central Hospital (Finland, 24 tumors), Lund University Hospital (Sweden, 23 tumors) and The University Hospital of Iceland (14 tumors). All tumors originated from families with three or more cases of breast cancer.

2.5 Kindreds included in the linkage analysis (III)

The 77 families for linkage analysis were identified through the Oncology clinics of the Helsinki and Tampere University Hospitals, Lund University Hospital in Sweden and the Department of Pathology at the University Hospital of Iceland. Twenty-three of these were families included in the CGH analysis and from which blood samples from multiple affected were available. Paraffin tissue samples were obtained from deceased patients when available. The initial set of families consisted of 57 Finnish families and the second set of 12 Icelandic and 8 Swedish families.

3. METHODS

3.1 Methods used in the *BRCA1* and *BRCA2* mutation analyses (I-II)

All 100 families were analyzed for the coding regions (including splice boundaries) of both the *BRCA1* and *BRCA2* genes using the combined HA/SSCP method. Detection of the HA/SSCP bands was performed by silverstaining as described in reference Gayther et al 1995 [81]. The large exons (exon 11 of *BRCA1*, 10 and 11 of *BRCA2*) were analyzed using PTT. Primer sequences used in these analyses are described in reference Friedman et al 1997 [132]. When aberrant mobility was detected on HA/SSCP or PTT gels, the variants were re-amplified from genomic DNA and directly sequenced. ASO hybridization was used for analyzing polymorphic status of the changes and for confirming the segregation of the putative mutation in other family members. Oligonucleotides used in ASO hybridization assays were designed using genomic sequences of the *BRCA1* and *BRCA2* genes (*5*). Seventy index cases from these families were also sequenced for the entire coding region of the *BRCA1* gene (Myriad Genetics) and three index cases for the *BRCA2* gene. The principal laboratory methods used are listed in Table 4.

Table 4 *Laboratory methods used in studies I-IV*

Laboratory method	**Used in study**	**Reference**
ASO	I,II, IV	Saiki et al 1986 [200]
PTT	I,II	Roest et al 1993 [197]
HA	I,II	Nagamine et al 1989 [193]
SSCP	I,II	Orita et al 1989 [190]
Sequencing	I,II	Sanger et al 1977 [189]
Genotyping	III, IV	Smith et al 1995 [204]
CGH	III	Kallioniemi et al 1992 [180]

3.2 Polymorphic markers used for the haplotype analysis of the *BRCA1* and *BRCA2* regions (II)

Polymorphic microsatellite repeat markers for the 17q21 region (*BRCA1*) haplotype analysis were D17S1185, D17S855 and D17S579 and markers for the 13q12 region (*BRCA2*) D13S260, D13S1701 and D13S267.

3.3 Polymorphic markers used for the haplotype analysis of the 999delTCAAA mutation families (IV, unpublished)

Haplotype analysis of the ten families included 24 markers spanning a genomic region of 29 cM [205], [206] (Figure 3). The map distances for microsatellite markers were obtained from Généthon (*6*). Haplotypes were constructed with Genehunter and manually assuming minimal parsimony.

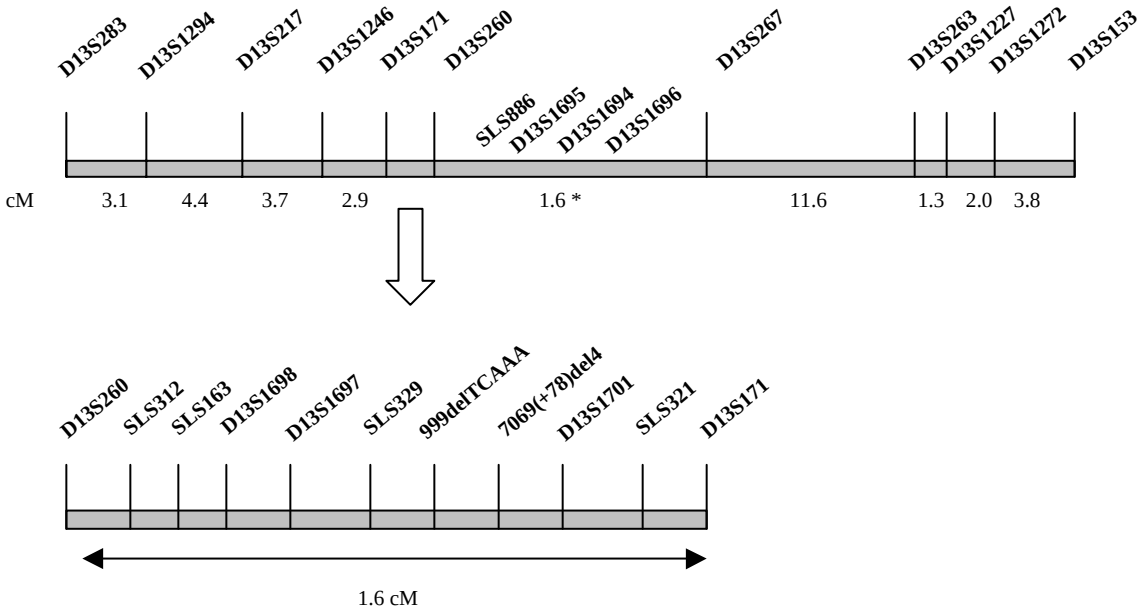


Figure 3 *The polymorphic markers studied in the 999delTCAAA families. The region between markers D13S 260 and D13S171 harboring the BRCA2 gene is shown in greater detail. * This 1.6cM region is shown out of proportion for clarity.*

3.4 CGH and related data analysis (III)

CGH was performed essentially as described in reference Kallioniemi et al. 1992 [180]. To select which recurrent events are most likely to be nonrandom (truly related to the cause of tumor development and progression), we used a well-established method of Brodeur et al. 1982 [207].

3.5 Genotyping of the families and related data analysis (III)

Genotyping was performed using 23 genetic markers from the Genome Database, distances as given by Généthon (*6,7*). The software used in the analysis is listed in Table 5. The population frequency of the disease allele was set to 0.0033 [8]. The autosomal dominant model was used. The age-dependent penetrances for mutation carriers and phenocopies are shown in Table 6 [76]. Females diagnosed with breast cancer were coded as affected and females with ovarian cancer as unknown. The statistical significance of the peak 2-point LOD score with marker D13S1308 was assessed by simulation, using the principles outlined in Ott 1999 [208].

Table 5 *Software used in study III*

Software	**Analysis**	**References**
Pedcheck	Genotype data inconsistency	209
FASTLINK software	Model dependent LOD score	210, 211, 212, 213
HOMOG2	Locus heterogeneity [1]	208
HOMOG3R	Locus heterogeneity [2]	208
SIBPAL (S.A.G.E.)	Model-independent affected sib-pair analysis	(8)
SIB-PAIR	Population-specific allele frequencies	214
FASTLINK	Multipoint linkage analyses	211, 212, 213
HOMOG	Multipoint heterogeneity LOD scores	208
Genehunter	Multipoint linkage analyses	170

1 *BRCA2* and the novel locus
2 *BRCA1* and the novel locus

Table 6 *The age-adjusted penetrances for mutation carriers and the phenocopy rates used in the linkage analysis.*

	< 30 y.	**30-39 y.**	**40-49 y.**	**50-59 y.**	**60-69 y.**	**70-79 y.**	**> 80 y.**
Penetrance	0.00167	0.01276	0.02305	0.01711	0.0126	0.00908	0.00654
Phenocopy Rate	0.00002	0.00026	0.00112	0.00137	0.00226	0.00218	0.00213

RESULTS AND DISCUSSION

1. *BRCA1* AND *BRCA2* MUTATIONS ARE FOUND AT A LOW FREQUENCY (I-II)

In the *BRCA1* gene, 10 different cancer-predisposing mutations were found, one in each family (Table 7 and Figure 4). The mutations were unevenly distributed over the *BRCA1* gene, with five mutations found in exon 11 and five mutations in exons 12-20. Nine of these mutations are truncating mutations. The splice site mutation 4216(-2)A>G destroys the invariant AG at the splicing acceptor site [215]. This splice substitution was not seen in 186 normal control chromosomes. The mutations 1924delA, 3904C>A, 3744delT, 4216(-2)A>G and 5145-5155del are so far unique to Finland. Of these five Finnish mutations, the 3744delT and 4216(-2)A>G have subsequently been found in Finland by another research group as well [216]. The 2803delAA mutation in exon 11 is also found in Holland, where it is present as a founder mutation. The 3604delA is present in Belgium and Holland [217]. The 4446C>T has been reported several times, in both Europe and Northern America and haplotype studies suggest multiple origins for this mutation [82]. The 5370C>T mutation is also found also in Austria [218]. The frameshift mutation 5382insC is found recurrently in Europe and is a Jewish founder mutation [219, 220]. The cancers occurring in the families carrying mutations in the *BRCA1* gene are summarized in Table 8.

Table 7 *BRCA1 mutations*

Exon	Mutation	Type	Found in
11	1924delA	frameshift	Finland
11	2803delAA	frameshift	Europe
11	3604delA	frameshift	Europe
11	3744delT	frameshift	Finland
11	3904C>A	nonsense	Finland
11*	4216(-2)A>G	splice site	Finland
13	4446C>T	nonsense	Europe, North America
17	5145-5155del	frameshift	Finland
20	5370C>T	nonsense	Europe
20	5382insC	frameshift	Europe

* Intron 10

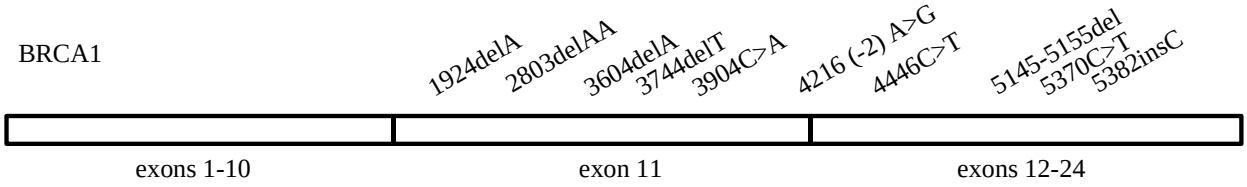


Figure 3 *Schematic representation of the distribution of the BRCA1 mutations detected in the mutation screening.*

Seven of the 10 *BRCA1* families included ovarian cancer, and three of these families had multiple cases. The median age of ovarian cancer onset was 49 years. A statistically significant association (P <0.005, Fisher's exact test) between the location of the mutation and prevalence of breast or ovarian cancer in the *BRCA1* families was discovered. In the five families with exon 11 mutations, a total of 9 breast cancers and 10 ovarian cancers were seen. On the contrary, in families with mutations 3' of the exon 11, a total of 19 breast cancers but only two ovarian cancers were seen. The proposed variation in breast and ovarian cancer risk of mutations in different halves of the gene [81] is supported by these data. Mutations in the *BRCA1* gene also predisposes carriers to prostate and colon cancer [84], however no such cases were observed in this set of mutation-positive families.

Table 8 *Cancer cases in the families with BRCA1 mutations*

***BRCA1* mutation**	**No. of BC cases**	**No. of OC cases**	**Other cancers in family**
1924delA	2	1	skin
2803delAA	2	2	unknown
3604delA	1	3	lung
3744delT	2[1]	1	stomach, skin
3904C>A	3	3	skin, corpus, pancreas, lung, salivary gland, peritoneal metastasis[2]
4216(-2)A>G	4[1]	0	-
4446C>T	9	0	lung
5145-5155del	3	0	gall bladder, peritoneal metastasis[2]
5370C>T	3	1	peritoneal metastasis[2]
5382insC	1	1	skin, lymphoma, multiple myeloma

BC Breast cancer OC Ovarian cancer

1 bilateral cases are summarized as two separate cases

2 primary unknown

In the *BRCA2* gene, six different mutations were discovered (Table 9 and Figure 5). Four of these are recurrent found in more than one family. Four of the mutations lead to premature termination of translation. One of these truncating mutations is a combination of a substitution and a deletion in positions 6495 and 6496, respectively. Both nucleotide changes reside on the same chromosome, which was determined using ASO hybridization with four different oligos detecting different allelic combinations. The 9346-2A>G is predicted to destroy the invariant splice acceptor site sequence AG. The 9730-12T>G variant is positioned at the conserved run of pyrimidines [215]. Neither of these splice-site substitutions were seen in 186 normal chromosomes and both segregated with the disease in families.

Table 9 *BRCA2 mutations*

Exon	Mutation	Type	Found in
9	999delTCAAA	frameshift	Iceland^{R}
11	6495 G>C 6496delCA	frameshift	Finland
15	7708 C>T*	nonsense	Finland^{R}
18	8555 T>G	nonsense	Finland^{R}
23**	9346(-2)A>G	splice site	Finland^{R}
25**	9730(-12)T>G	splice site	Finland

* Detected by sequencing

** Intron

R Recurrent

Five of these *BRCA2* mutations are novel Finnish mutations. Two of these five novel mutations (999delTCAAA and 9346(-2)A>G) have subsequently been found in other Finnish breast cancer families as well [216]. The 999delTCAAA has been reported previously as a strong founder mutation in Iceland [221]. It has not been observed in other Scandinavian countries [222-224]. There are two other mutations reported in this region, the 1000delAA (reported as 258delAA) and the 995delCAAAT, suggesting that this area is a hotspot for mutations [225, 226]. Mavraki and coworkers proposed that this codon could be associated with increased susceptibility to male breast cancer [225]. The absence of male breast cancer patients among these Finnish gene carriers does not support this hypothesis. However, it should be noted that the number of families is very small. It has been noted that *BRCA2* carriers are at an increased risk of prostate cancer, and indeed, four cases were observed in these ten families [221]. Four of these *BRCA2* mutations described here were found recurrently. More than one frequent mutation may be explained by multiple founder effect with genetic drift in smaller communities, which is thereafter mixed to form the larger population [227]. In fact, it has been recently shown that carriers of these recurrent Finnish *BRCA1* and *BRCA2* mutations have common ancestors. The variation in the length of the shared haplotype indicates that distinct mutations have started to spread at different times [228].

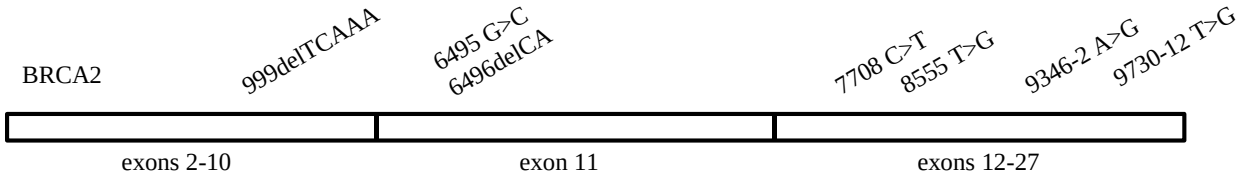


Figure 5 *Schematic representation of the distribution of the BRCA2 mutations detected in the mutation screening.*

Table 10 shows the cancer cases occurring in the families carrying *BRCA2* mutations. A total of 46 breast cancer cases were found in the families. Ovarian cancers were detected in seven of these families. In addition, other types of malignancies were present in all families.

Table 10 *Cancer cases in the families with BRCA2 mutations*

Mutation	**No. of BC cases**	**No. of OC cases**	**Other cancers in family**
999delTCAAA	5	0	melanoma, skin[1]
999delTCAAA	6*	0	rectum, lung, cervix, corpus uteri, liver, tonsil, prostate
6496delCA	3	0	stomach, lung, skin[1]
7708C>T	6*	1	-
7708C>T	6	1	prostate, cervix, bladder
7708C>T	3	1	-
8555T>G	3*	1	tongue
8555T>G	5	0	prostate
9346(-2)A>G	3	1	liver
9346(-2)A>G	5	2	melanoma, prostate, pancreas
9730(-12)T>G	1	1	skin[1] , lymphoma, multiple myeloma

BC Breast cancer OC Ovarian cancer
* Bilateral cases are summarized as two separate cases
1 unclassified

Altogether, sixteen different mutations were found in this set of 100 high-risk families. *BRCA1* mutations account for 10% (95% CI = 5-18) of these 100 families and *BRCA2* mutations for an additional 11% (95% CI = 6-19). The *BRCA1* mutations were found more frequently in families with both breast and ovarian cancer than in families with breast cancer only. *BRCA2* mutations were most

common in the largest families with four or more cases of breast cancer or both breast and ovarian cancer. Together, mutations in the two genes were found in 12/27 (44%; 95% CI = 25-65) of the breast-ovarian cancer families and in 4/23 (17%; 95% CI = 5-39) of site specific breast cancer families with four or more cases of breast cancer. Only 4/50 (8%, 95% CI = 2-19) of breast cancer families with three affected members were mutation-positive. The mutation spectrum is wide in these Finnish families, which is consistent with mutation studies in other populations. In summary, a total of 21% (95% CI = 13-30) of the large breast cancer pedigrees carry mutation in either gene. The proportion of the high-risk breast cancer families attributable to *BRCA1* and *BRCA2* has been shown to vary considerably between countries. The observed 21 % frequency is lower than in any other country surveyed so far [229].

An early age of breast cancer onset is a clear indicator of *BRCA1* as well as *BRCA2* mutation carrier status in these Finnish breast cancer families. The age of onset in the *BRCA1* (median 43.9 years) (P <0.0001, unpaired t-test) and in the *BRCA2* families (median 49.2 years) (P = 0.0001) was significantly lower than in the mutation negative families (median 57.4 years). An older age of onset in the *BRCA2* families compared with *BRCA1* families has also been found in other studies [226] and is reflected as a smaller contribution of *BRCA2* to early onset breast cancer [135].

It is possible that a few mutations have been missed due to limitations of the screening methods used. The 7708 C>T mutation in *BRCA2* was detected by sequencing previously missed by the HA/SSCPmethod. Although the combined use of SSCP with HA offers higher sensitivity, clearly some mutations will be missed [188]. The methods used in this study lack the capacity to detect gross genomic rearrangements comprising several exons. In addition, small in-frame deletions that result in the loss of a few residues in the resulting protein would not be detected by the PTT approach [197]. This is due to the resolution of the SDS-PAGE method used for separating the produced protein fragments. A recent study revealed that 12% of the mutation-negative families had actually large deletions in the *BRCA1* gene [230]. In order to obtain a more accurate frequency of the *BRCA1* and *BRCA2* mutations in this cohort, screening for such deletions is necessary. Mutations involving the promoter region of the *BRCA1* gene have been reported twice [71, 231], while none have been reported for *BRCA2*. The possibility that promoter-involving mutations are present in the study sample can not be excluded, and additional mutation screening is needed to clarify the issue.

In addition, a few families may be explained by the presence of mutations in the *TP53* gene, which has been shown to account for a small fraction of the high-risk families in Finland. These *BRCA1* and *BRCA2* mutation-negative families fulfilled the criteria of either the Li-Fraumeni syndrome or the Li-Fraumeni-like syndrome [232]. Despite the involvement of other known genes, the results shown here strongly suggest that additional yet unidentified breast cancer predisposition genes are likely to be important in Finland.

2. BASED ON THE HAPLOTYPE ANALYSIS THERE ARE SEVERAL FAMILIES NOT LINKED TO THE *BRCA1* OR *BRCA2* GENES (II)

In order to confirm the low involvement of these two genes, we also performed haplotype analysis in the *BRCA1* and *BRCA2* regions of chromosome 17q21 and 13q12 in 12 of the larger families where mutations had not been found. Not enough samples were available for formal linkage analysis in the families, thus, affected family members were studied for sharing alleles, or haplotypes where possible, of three microsatellite markers on each chromosome. In seven of these families, linkage to both

chromosome 17q21 (*BRCA1*) and 13q12 (*BRCA2*) was definitely excluded. Analysis of one breast-ovarian cancer family suggested linkage to *BRCA2*. Subsequently the 7708C>T nonsense mutation was found in this family by direct sequencing. In four families, analysis was not informative for either the *BRCA1* or *BRCA2* gene. However, no mutations were found by direct sequencing or by PTT and HA/SSCP analysis. There was no evidence of a common haplotype between families for which exclusion was not obtained. Mutation analysis combined with haplotype analysis suggests that additional susceptibility genes may account for a large proportion of breast and breast-ovarian cancer families in Finland.

3. THE 999delTCAAA MUTATION IN THE *BRCA2* GENE IS FOUND IN TWO DIFFERENT HAPLOTYPE BACKGROUNDS (IV, UNPUBLISHED)

Haplotype analysis with 24 markers (29 cM) revealed two distinct haplotypes that are associated with this mutation in Finland (Table 11, Figure 6). The more common haplotype (denoted as haplotype A) was present in 7 families and the rarer haplotype (denoted as haplotype B) in 3 families. The families carrying haplotype A all share a common region of approximately 6 cM. The families segregating haplotype A originate from two geographical regions: in the old agricultural area Pirkanmaa, and in the new settlement region, in the most Eastern part of the country. In haplotype B the core haplotype D13S260 to D13S267 (3 cM) is present in all three families with the exception that family 158 carries allele 1 (instead of the allele 2) at marker D13S1698. This could be due to mutation [233] or a null allele [234]. All families with haplotype B are located in Pirkanmaa only (Figure 6). The families carrying haplotype B have their origins in Pirkanmaa, in the very same small region as the families segregating haplotype A.

Table 11 *The haplotypes in the 999delTCAAA families. The boundaries of the shared region in haplotypes between families is indicated by the arrows.*

	Haplotype A								Haplotype B			
Family	*130*	*397*	*63*	*11*	*113*	*97*	*42*	Marker	*102*	*4*	*158*	
	2	7	2	5	1	2	2	D13S283	7	1	2	
	2	3	3	6	6	2	2	D13S1294	2	2	2	
	5	3	3	3	3	3	3	D13S217	2	5	1	
→	**6**	**6**	**6**	**6**	**6**	**6**	**6**	D13S1246	3	5	6	←
	4	**4**	**4**	**4**	**4**	**4**	**4**	D13S260	**7**	**7**	**7**	
	3	**3**	**3**	**3**	**3**	**3**	**3**	SLS312	**1**	**1**	**1**	
	1	**1**	**1**	**1**	**1**	**1**	**1**	D13S1699	**2**	**2**	**2**	
	2	**2**	**2**	**2**	**2**	**2**	**2**	D13S1698	**2**	**2**	**1**	
	1	**1**	**1**	**1**	**1**	**1**	**1**	SLS163	**1**	**1**	**1**	
	2	**2**	**2**	**2**	**2**	**2**	**2**	D13S1697	**1**	**1**	**1**	
	2	**2**	**2**	**2**	**2**	**2**	**2**	SLS329	**1**	**1**	**1**	
	2	**2**	**2**	**2**	**2**	**2**	**2**	7069(+78)del4	**2**	**2**	**2**	
	5	**5**	**5**	**5**	**5**	**5**	**5**	D13S1701	**2**	**2**	**2**	
	1	**1**	**1**	**1**	**1**	**1**	**1**	SLS321	**1**	**1**	**1**	
	2	**2**	**2**	**2**	**2**	**2**	**2**	D13S171	**2**	**2**	**2**	
	2	**2**	**2**	**2**	**2**	**2**	**2**	SLS886	**1**	**1**	**1**	
	2	**2**	**2**	**2**	**2**	**2**	**2**	D13S1695	**3**	**3**	**3**	
	1	**1**	**1**	**1**	**1**	**1**	**1**	D13S1694	**1**	**1**	**1**	
→	**2**	**2**	**2**	**2**	**2**	**2**	**2**	D13S1696	**2**	**2**	**2**	
	6	6	6	1	6	6	6	D13S267	**4**	**4**	**4**	←
	1	5	1	7	1	1	1	D13S263	4	8	8	
	8	5	5	7	7	5	5	D13S1227	2	6	6	
	1	1	3	1	1	1	1	D13S1272	2	4	4	
	5	6	2	3	8	4	4	D13S153	1	4	4	

The presence of two distinct haplotypes supports the hypothesis that this region is in fact is a hotspot. In eukaryotic genomes, short tandem repeats are often hotspots for deletions and insertions [235]. It has been proposed that short symmetric elements predispose DNA sequences to meiotic microdeletion [236]and indeed, the 999delTCAAA-mutation region is both preceded and followed by such symmetric elements (Figure 7). Recurrent mutations with distinct origins have been reported in the *BRCA1* gene [237] and a similar scenario may be true in the *BRCA2* gene.

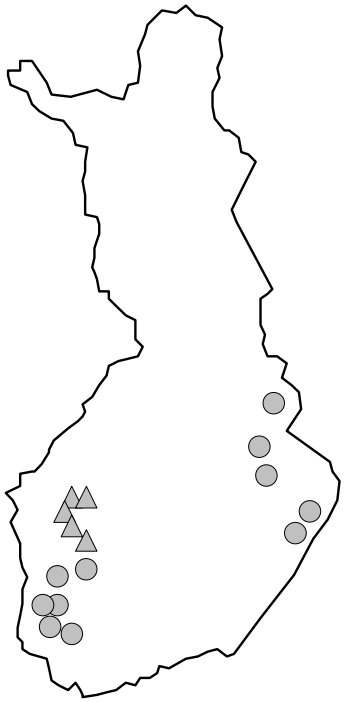


Figure 6 *Map of Finland showing the ancestries of the 999delTCAAA families. The circles indicate origins of the families carrying the haplotype A. The triangles indicate origins of the families carrying haplotype B.*

AGTGAAAACACAAA**TCAAA**GAGAAGCTGCAAGTCATG

Figure 7 *Sequence of the region encompassing the 999delTCAAA mutation. The deleted 5-bp sequence is shown in bold. Symmetric elements on either side of the deleted region are underlined.*

However, an alternative explanation for two haplotypes in Finland is the occurrence of gene conversion. Both haplotypes A and B are present in a very restricted geographical region, in adjacent communities in the old agricultural region. Gene conversion becomes possible when a certain allele reaches a high frequency in a population and is subsequently mixed with the surrounding one with a different haplotype composition. Haplotype B was possibly originally present in a small, geographically restricted area that may well have been an isolated sub-population. The gene conversion may have occurred when the gene pool of this local isolate was mixed with a larger one due to population growth following migration. A situation very similar to the 999delTCAAA has previously been reported in Finland in the transglutaminase gene. One mutation was observed in two distinct haplotype backgrounds. Both observed haplotypes were present in a restricted geographical region and also distinct mutations involving the same genomic region had been described in other populations [238].

4. THE NOVEL BREAST CANCER-PREDISPOSING GENE, *BRCA4**, IS LOCATED AT 13q22 (III)

Losses in *BRCAX** tumors were most often seen at 13q and gains at 1q (Figure 8). The method of Brodeur et al. [207]. selected the following events as nonrandom in *BRCAX* cases: +1q, +8q, +16p, +17q, +19p, +19q, +20q, -6q, -8p, -9p, -11q, and -13q. The loss of 13q was selected as being an early event in the progression of *BRCAX* tumors with both branching and phylogenetic (distance matrix) trees. In contrast, 6q loss was located far from the root of both tree types, suggesting that it may not be an early event in the progression of these tumors.

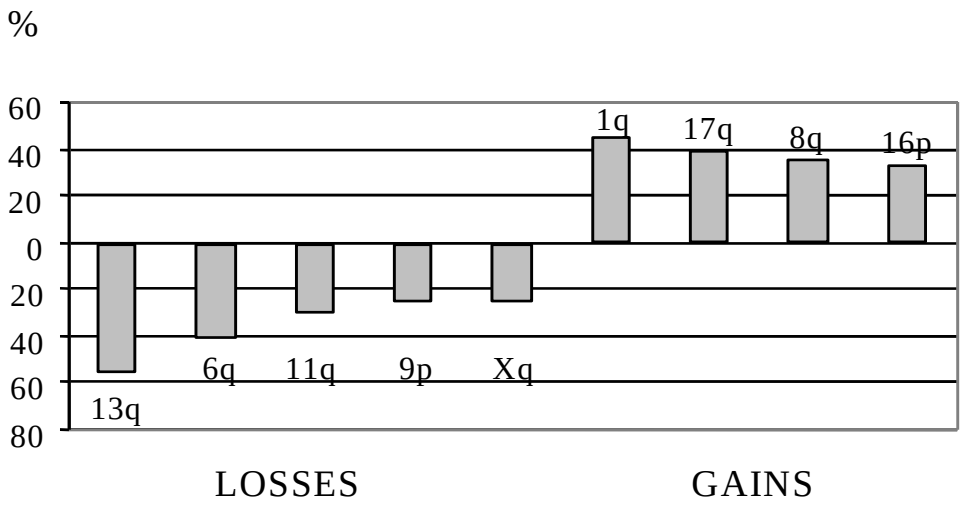


Figure 8 *Gains and losses in the BRCAX tumors*

* The novel, yet unidentified, gene was named *BRCAX* by the authors at the time these analyses were performed. Presently, the putative gene is listed in OMIM (*2*) as *BRCA4* (BREAST CANCER, TYPE 4). The designation *BRCAX* is used instead of BRCA4 throughout the text.

Compelling evidence implicating 13q loss as an early event in the *BRCAX* tumors came from one large family (L5), where all five samples analyzed from different patients showed a deletion at 13q, with the consensus region of deletion at 13q21-q22. Loss of 13q was the only genetic alteration common to all five tumors from this family. Several other families showed individual tumors with the same narrow region of loss. Loss at 13q was common in samples from each country (Finland 13/24, Sweden 13/23, and Iceland 8/14). To test whether losses of the 13q21-q22 region were selected for, because this region harbors a germline alteration predisposing to breast cancer, we genotyped genetic markers along the 13q region in the germline DNA of the L5 family members. A 7.3 cM haplotype was shared between all five affected cases in this family with markers along 13q. The region of sharing was apparent between markers D13S1317 and D13S166. Markers at the *BRCA2* region were not shared. A metaphase fluorescence *in situ* hybridization experiment localized a yeast artificial chromosome probe for a shared genetic marker (D13S1257) to 13q22. This haplotype-sharing region was therefore identical to the region of loss in the tumors.

In the targeted linkage analysis 3 markers (D13S1308, D13S1257 and D13S791) had two-point LOD scores > 1 in the Finnish population. The peak two-point LOD score (2.89) was observed at D13S1308 with $\theta = 0.04$. After combination with the Swedish and the Icelandic data, six markers within a 10 cM region (D13S1313 to D13S162) had two-point LOD scores greater than 1 (Table 12). The peak LOD score was 2.76, $\theta = 0.10$ with marker D13S1308. There was no evidence for sex dependence of the recombination fraction.

HOMOG3R analyses showed no evidence of heterogeneity, indicating a low fraction of *BRCA1*-linked families in this data set. The HOMOG2 analyses did not result in significant evidence of heterogeneity ($p>0.3$). This suggests that while there may be a few *BRCA2*-linked families that were not discovered in the extensive mutation detection, the proportion of such families is likely to be small.

Multipoint LOD scores were positive (2.60) between markers D13S1296 and D13S1308. The highest HLOD score was 3.46 with the estimated proportion of linked families (α) of 0.65 in the region between D13S1296 and D13S1308. The HLODs were positive from D13S1262 to D13S160.

Nonparametric affected sib-pair analysis also supported evidence of linkage at 13q21-q22. The number of informative affected sib-pairs ranged from 33 to 65 in the Finnish (mean 50), from two to 12 in the Icelandic (mean 7) and from 10 to 23 in the Swedish (mean 20) families. The most significant results were obtained in the Finnish families at marker D13S1257 ($p=0.003$) with D13S1308 showing borderline significance ($p=0.05$). In all 77 families, three markers were slightly positive: D13S1257 ($p=0.015$), D13S791 ($p=0.05$) and D13S269 ($p=0.04$).

Table 12 *Combined LOD scores from the 77 families for markers on 13q. LOD scores greater than 1.0 are shown in bold.*

		LOD score at θ					
Marker	Genetic Distance cM	0.00	0.05	0.10	0.20	0.30	0.40
D13S260*		-8.24	-4.81	-2.92	-1.01	-0.27	-0.04
D13S1701*	1.6	-4.36	-2.48	-1.43	-0.39	-0.05	0.03
D13S267*	0	-2.35	-1.16	-0.56	-0.04	0.08	0.06
D13S1301	20.3	-1.75	-0.69	-0.16	0.23	0.21	0.08
D13S1313	0.8	0.94	**1.13**	**1.18**	0.98	0.56	0.17
D13S1317	1.9	-1.87	-0.18	0.58	0.96	0.67	0.26
D13S1262	0	-4.01	-2.03	-1.09	-0.21	0.06	0.07
D13S275	0.8	-0.24	0.7	0.98	0.87	0.48	0.11
D13S1296	0.8	-0.27	0.79	**1.13**	**1.05**	0.59	0.15
D13S1308	0	**1.61**	**2.56**	**2.76**	**2.19**	**1.24**	0.4
D13S1291	0.5	-2.64	-0.85	-0.03	0.52	0.41	0.13
D13S152	2.6	-5.58	-2.12	-0.67	0.41	0.44	0.15
D13S1257	0	-0.4	0.64	**1.14**	**1.26**	0.84	0.3
D13S745	0	0.63	0.86	0.89	0.68	0.39	0.11
D13S791	0	-1.55	0.36	0.99	**1.09**	0.65	0.18
D13S1326	0	-3.34	-1.2	-0.28	0.32	0.27	0.07
D13S1249	0.7	-0.45	0.16	0.47	0.59	0.39	0.13
D13S166	0	-4.93	-2.03	-0.7	0.39	0.45	0.19
D13S269	1.0	-1.66	-0.2	0.5	0.85	0.56	0.16
D13S162	1.4	-1.13	0.68	**1.39**	**1.55**	0.98	0.33
D13S1306	1.2	-3.01	-0.98	-0.07	0.51	0.41	0.15
D13S160	1.8	-2.26	-0.5	0.18	0.47	0.27	0.04
D13S1255	1.6	-2.87	-1.08	-0.16	0.48	0.4	0.12

**BRCA2* region

Tumor tissue for loss of heterozygosity analysis was available from 16 individuals from six families with evidence of linkage to 13q22. Allelic imbalance with multiple markers at 13q22 was observed in 11/16 cases (69%), whereas the frequency observed in sporadic tumors in this region is 25% [239]. In eight of the 11 cases, the putative wildtype allele demonstrated loss. LOH provides molecular evidence supporting the hypothesis that the gene in this region functions as a tumor suppressor-gene. LOH is frequent in both *BRCA1* [96, 240] and *BRCA2* [241] with demonstrated loss of the wildtype. It is most tempting to speculate that the observed somatic loss of the wildtype allele in these *BRCAX* tumors implies that the new locus at 13q is a classical tumor suppressor gene.

A simulation experiment was carried out to validate the significance of the two point linkage score empirically. The simulation also addressed the concern that undetected *BRCA2* mutation positive families had influenced the observed LOD scores at 13q21-q22. We conservatively estimated that the upper limit of *BRCA2*-linked families would be 17/57 Finnish families, 1/12 Icelandic families, and 1/8 Swedish families. In all possible configurations allowing this number of families to have undetected *BRCA2* mutations, the simulated LOD score at D13S1308 exceeded the threshold of 2.76 (corresponding to the peak true LOD score) in at most one of the 3000 replicates. This corresponds to significance level of $p<0.0017$. For a linkage analysis that targets a chromosome no larger than X due to substantial prior evidence for linkage, $p<0.01$ is considered sufficient evidence for linkage [208].

In the 77 families, a maximum two-point LOD score of 2.76 and a peak multipoint HLOD of 3.46 (alpha = 0.65) were observed. These LOD scores were obtained from a linkage analysis targeting a single candidate site in the genome, based on prior suggestions from CGH. This situation is in sharp contrast to genome-wide linkage scans where hundreds of comparisons are performed. Therefore, much more stringent thresholds for statistical significance need to be applied in genome-wide approaches. The two-point linkage result was further evaluated by means of a simulation analysis. Based on haplotype sharing at the *BRCA2* locus, we modeled that up to 17 Finnish families, one Icelandic family, and one Swedish family to be putative *BRCA2* families, in which *BRCA2* mutations may have been missed. This was done to evaluate the influence of putative "contaminating" *BRCA2* families on the LOD scores observed in the 13q21-q22 region. The simulation results suggested that the observed linkage to D13S1308 was significant at a level of $P < 0.0017$.

In our analyses of *BRCA1* and *BRCA2* mutation-negative breast cancer families, four intersecting lines of evidence suggest that the 13q losses observed in these tumors might represent early, initiating genetic alterations. First, the overall frequency of 13q loss was about two times higher in *BRCAX* tumors than in control breast cancers. Second, exploring the CGH data with two different tree models placed the loss of 13q at or near the root of the hypothetical breast cancer progression pathways. Third, tumor tissues from patients belonging to the same cancer family showed the same specific deletion at 13q21-q22 and shared a germline haplotype at this locus. Finally, and most importantly, unselected *BRCAX* breast cancer families showed significant evidence of linkage to the 13q21-q22 region. It will be critically important to independently validate this genetic locus on larger set of families, especially in populations outside of the Nordic countries. Further studies will allow assessment of allelic heterogeneity and penetrance estimates. To determine whether any other cancers besides breast cancer are present at an elevated level, further study based on this same set of families is necessary.

SUMMARY

Identification of individuals at a high-risk of developing breast cancer is a major task. Hereditary breast cancer already occurs at a young age and mutation carriers are not detected through general screening programs. Determining the incidence of cancer predisposing mutations facilitates further analyses of cancer risks and phenotypes associated with specific mutations. These are important parameters when establishing genetic testing and counseling. High-risk families also serve as models for understanding breast cancer in general, the full extent of associated heterogeneity, and hopefully also interactions of environmental and cultural risks. Hereditary breast cancer susceptibility is a complex phenomenon, in which multiple genes may play a role. The genes known to be important to the development of cancer regulate diverse cellular pathways. The two known breast cancer predisposing genes, *BRCA1* and *BRCA2,* are tumor suppressor genes with high penetrance. Both have a function in DNA repair.

The proportion of families in which breast cancer is due to germline *BRCA1* and *BRCA2* mutations has been studied in a cohort of 100 Finnish families with 3 or more affected individuals. In this set of high-risk breast cancer families a total of 16 different mutations in the *BRCA1* and *BRCA2* genes were found and accounted for 21% of the families analyzed. In the *BRCA1* gene, 10 different cancer predisposing mutations were identified each in one family. In the *BRCA2* gene, six different mutations were identified in 11 families, with four recurrent mutations found each in more than one family. The most common *BRCA2* mutation in Finland, 999delTCAAA, is present in two different haplotype backgrounds. Either a hotspot or a gene conversion may explain this finding. The two major breast cancer predisposing genes, *BRCA1* and *BRCA2* are involved in hereditary breast cancer families in Finland. Despite the fact that some mutations may have been missed due to the limitations of the screening methods used, the majority of the high-risk families are not explained by mutations in these two genes. Therefore, the presence of an additional, yet unidentified gene was studied.

The identification of a candidate susceptibility locus at 13q21 was based on a novel strategy, which assumed that somatic genetic changes in cancer tissues might give insights into the nature of the germline predisposing loci. Detection of somatic deletions in the wildtype gene by comparative genomic hybridization in 61 tumors was followed by targeted linkage analysis out in a set of 77 Finnish, Icelandic, and Swedish breast cancer families. Distinction of early genetic events was facilitated by the application of 2 complementary mathematical tree models for analysis of the CGH data. The studies predicted that the loss of 13q was one of the earliest genetic events in hereditary cancer. In a family with 5 breast cancer cases, all analyzed tumors showed distinct 13q deletions, with the minimal region of loss at 13q21-q22. Genotyping revealed segregation of a shared 13q21 germline haplotype in the family. In the targeted linkage analysis, a maximum parametric 2-point LOD score of 2.76 was obtained for a marker at 13q21 (D13S1308, theta = 0.10). The multipoint LOD score under heterogeneity was 3.46. The findings provide preliminary evidence for the location of a novel predisposition locus for breast cancer at 13q21-q22.

ACKNOWLEDGMENTS

This work was carried out during 1995-1997 at the Research Laboratory of the Department of Obstetrics and Gynecology, Helsinki University Central Hospital and during 1997-1998 at The National Human Genome Research Institute, Cancer Genetics Branch, USA. I wish to thank the professors Olavi Ylikorkala and Markku Seppälä, heads of the Department of Obstetrics and Gynecology and Jeffrey Trent, chief of the Cancer Genetics Branch at NHGRI. Both institutes are acknowledged for placing excellent research equipment at my disposal.

First of all I would like to thank the families that have participated and thus made these studies possible. I also wish to thank the great number of people who have contributed to this thesis if not specifically mentioned.

My supervisor, Docent Heli Nevanlinna is thanked for introducing me to the subject in the first place and supervising this thesis. I also wish to acknowledge the skillful people involved in these BRCA studies: Professor Seppo Pyrhönen at the Department of Oncology and Radiotherapy in Turku University Hospital, Hannaleena Eerola and Timo Muhonen at the Department of Oncology and Laura Sarantaus at the Department of Obstetrics and Gynecology in the Helsinki University Central Hospital. I am grateful to Lori Friedman, Professor Olli-Pekka Kallioniemi and Assistant Professor Juha Kere, who are acknowledged for additional scientific guidance in these studies. I am indebted to the skillful members of the collaborative groups, especially Tommi Kainu, Elizabeth Gillanders, Diana Freas-Lutz and Ester Rozenblum at NHGRI and Kirsi Syrjäkoski at the Tampere University.

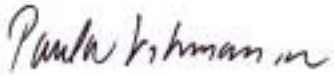
I also wish to thank Docent Minna Nyström-Lahti and Docent Jukka Partanen who did an excellent job with the critical review of this thesis. James Woolley is thanked for the language revision.

I want to express my gratitude to the head of the Department of Biosciences, Kielo Haahtela for giving me good advice. The Professor of the Division of Genetics, Pekka Heino is thanked for collaboration with finishing my studies. My thanks are also due to Professor Jim Schröder for encouraging me to finish this thesis. All the teachers at the Division of Genetics are thanked for teaching me the various aspects of genetics.

The Academy of Finland, The Cancer Foundation of Finland, The Clinical Research Fund of Helsinki University Hospital, The Nordic Cancer Union and The Sigrid Juselius Foundation have made this work financially possible.

Previous colleagues at the Department of Obstetrics and Gynecology and at The National Human Genome Research Institute are warmly remembered. Present colleagues at the Finnish Red Cross Blood Transfusion Service are thanked for their kind interest toward my thesis. Good friends have been valuable during this work.

Helsinki, June 2001



REFERENCES

www information

1 The Cancer Registry
http://www.cancerregistry.fi/v97/contents.html

2 Online Mendelian Inheritance in Man
http://www.ncbi.nlm.nih.gov/Omim/

3 Breast Cancer Information Core
http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic/

4 Human Gene Mutation Database
http://www.uwcm.ac.uk/uwcm/mg/hgmd0.html

5 GenBank
http://www.ncbi.nlm.nih.gov/Genbank/index.html

6 Genethon
ftp://ftp.genethon.fr/pub/Gmap/Nature-1995/

7 Genome Database
http://gdbwww.gdb.org/

8 Statistical Analysis for Genetic Epidemiology
http://darwin.cwru.edu/pub/sage.html

Literature cited

1. Eisinger, F., Sobol, H., Serin, D. & Whorton, J.C. Hereditary breast cancer, circa 1750. *Lancet* 351, 1366 (1998).
2. Sattin, R.W. *et al.* Family history and the risk of breast cancer. *Jama* 253, 1908-1913 (1985).
3. Anderson, D.E. Genetic study of breast cancer: identification of a high risk group. *Cancer* 34, 1090-1097 (1974).
4. Rosenblatt, K.A. *et al.* Breast cancer in men: aspects of familial aggregation. *J Natl Cancer Inst* 83, 849-854 (1991).
5. Eeles, R., Ponder, B., Easton, D. & Horwich, A. *Genetic Predisposition to Cancer*. Chapman & Hall Medical, London (1996).
6. Lynch, H.T. *et al.* Genetic predisposition to breast cancer. *Cancer* 53, 612-622 (1984).
7. Newman, B., Austin, M.A., Lee, M. & King, M.C. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci U S A* 85, 3044-3048 (1988).
8. Claus, E.B., Risch, N. & Thompson, W.D. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet* 48, 232-242 (1991).
9. Knudson, A.G., Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 68, 820-823 (1971).
10. Coles, C. *et al.* Evidence implicating at least two genes on chromosome 17p in breast carcinogenesis. *Lancet* 336, 761-763 (1990).
11. Davidoff, A.M., Humphrey, P.A., Iglehart, J.D. & Marks, J.R. Genetic basis for p53 overexpression in human breast cancer. *Proc Natl Acad Sci U S A* 88, 5006-5010 (1991).

12. Borresen, A.L. *et al.* Screening for germ line TP53 mutations in breast cancer patients. *Cancer Res* 52, 3234-3236 (1992).
13. Sidransky, D. *et al.* Inherited p53 gene mutations in breast cancer. *Cancer Res* 52, 2984-2986 (1992).
14. Brownstein, M.H., Wolf, M. & Bikowski, J.B. Cowden's disease: a cutaneous marker of breast cancer. *Cancer* 41, 2393-2398 (1978).
15. Tsou, H.C. *et al.* The role of MMAC1 mutations in early-onset breast cancer: causative in association with Cowden syndrome and excluded in BRCA1-negative cases. *Am J Hum Genet* 61, 1036-1043 (1997).
16. Freihoff, D. *et al.* Exclusion of a major role for the PTEN tumour-suppressor gene in breast carcinomas. *Br J Cancer* 79, 754-758 (1999).
17. Shugart, Y.Y. *et al.* Linkage analysis of 56 multiplex families excludes the Cowden disease gene PTEN as a major contributor to familial breast cancer. *J Med Genet* 36, 720-721 (1999).
18. Swift, M., Reitnauer, P.J., Morrell, D. & Chase, C.L. Breast and other cancers in families with ataxia-telangiectasia. *N Engl J Med* 316, 1289-1294 (1987).
19. Stankovic, T. *et al.* ATM mutations and phenotypes in ataxia-telangiectasia families in the British Isles: expression of mutant ATM and the risk of leukemia, lymphoma, and breast cancer. *Am J Hum Genet* 62, 334-345 (1998).
20. FitzGerald, M.G. *et al.* Heterozygous ATM mutations do not contribute to early onset of breast cancer. *Nat Genet* 15, 307-310 (1997).
21. Chen, J., Birkholtz, G.G., Lindblom, P., Rubio, C. & Lindblom, A. The role of ataxia-telangiectasia heterozygotes in familial breast cancer. *Cancer Res* 58, 1376-1379 (1998).
22. Li, L., Li, X., Francke, U. & Cohen, S.N. The TSG101 tumor susceptibility gene is located in chromosome 11 band p15 and is mutated in human breast cancer. *Cell* 88, 143-154 (1997).
23. Hofferbert, S., Brohm, M. & Weber, B.H. Search for TSG101 germ-line mutations in BRCA1/BRCA2-negative breast/ovarian cancer families. *Cancer Genet Cytogenet* 102, 86-87 (1998).
24. Hall, J.M. *et al.* Rare HRAS alleles and susceptibility to human breast cancer. *Genomics* 6, 188-191 (1990).
25. Lindblom, A. *et al.* Predisposition for breast cancer in carriers of constitutional translocation 11q;22q. *Am J Hum Genet* 54, 871-876 (1994).
26. Piver, M.S. *et al.* Familial ovarian cancer. A report of 658 families from the Gilda Radner Familial Ovarian Cancer Registry 1981-1991. *Cancer* 71, 582-588 (1993).
27. Lynch, H.T., Watson, P., Lynch, J.F., Conway, T.A. & Fili, M. Hereditary ovarian cancer. Heterogeneity in age at onset. *Cancer* 71, 573-581 (1993).
28. Anderson, D.E. & Badzioch, M.D. Familial breast cancer risks. Effects of prostate and other cancers. *Cancer* 72, 114-119 (1993).
29. Aarnio, M. *et al.* Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int J Cancer* 81, 214-218 (1999).
30. Watson, P. & Lynch, H.T. Extracolonic cancer in hereditary nonpolyposis colorectal cancer. *Cancer* 71, 677-685 (1993).
31. Eccles, D.M. *et al.* Segregation analysis of ovarian cancer using diathesis to include other cancers. *Ann Hum Genet* 61, 243-252 (1997).
32. Stratton, J.F. *et al.* The genetic epidemiology of early-onset epithelial ovarian cancer: a population-based study. *Am J Hum Genet* 65, 1725-1732 (1999).
33. Hall, J.M. *et al.* Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250, 1684-1689 (1990).

34. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266, 66-71 (1994).
35. Friedman, L.S. *et al.* Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nat Genet* 8, 399-404 (1994).
36. Smith, T.M. *et al.* Complete genomic sequence and analysis of 117 kb of human DNA containing the gene BRCA1. *Genome Res* 6, 1029-1049 (1996).
37. Xu, C.F. *et al.* Distinct transcription start sites generate two forms of BRCA1 mRNA. *Hum Mol Genet* 4, 2259-2264 (1995).
38. Thakur, S. & Croce, C.M. Positive regulation of the BRCA1 promoter. *J Biol Chem* 274, 8837-8843 (1999).
39. Coene, E., Van Oostveldt, P., Willems, K., van Emmelo, J. & De Potter, C.R. BRCA1 is localized in cytoplasmic tube-like invaginations in the nucleus. *Nat Genet* 16, 122-124 (1997).
40. Hsu, L.C. & White, R.L. BRCA1 is associated with the centrosome during mitosis. *Proc Natl Acad Sci U S A* 95, 12983-12988 (1998).
41. Thakur, S. *et al.* Localization of BRCA1 and a splice variant identifies the nuclear localization signal. *Mol Cell Biol* 17, 444-452 (1997).
42. Boddy, M.N., Freemont, P.S. & Borden, K.L. The p53-associated protein MDM2 contains a newly characterized zinc-binding domain called the RING finger. *Trends Biochem Sci* 19, 198-199 (1994).
43. Bork, P. *et al.* A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *Faseb J* 11, 68-76 (1997).
44. Rao, V.N., Shao, N., Ahmad, M. & Reddy, E.S. Antisense RNA to the putative tumor suppressor gene BRCA1 transforms mouse fibroblasts. *Oncogene* 12, 523-528 (1996).
45. Scully, R. *et al.* Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* 88, 265-275 (1997).
46. Zhang, H. *et al.* BRCA1 physically associates with p53 and stimulates its transcriptional activity. *Oncogene* 16, 1713-1721 (1998).
47. Scully, R. *et al.* BRCA1 is a component of the RNA polymerase II holoenzyme. *Proc Natl Acad Sci U S A* 94, 5605-5610 (1997).
48. Wu, L.C. *et al.* Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nat Genet* 14, 430-440 (1996).
49. Thai, T.H. *et al.* Mutations in the BRCA1-associated RING domain (BARD1) gene in primary breast, ovarian and uterine cancers. *Hum Mol Genet* 7, 195-202 (1998).
50. Yarden, R.I. & Brody, L.C. BRCA1 interacts with components of the histone deacetylase complex. *Proc Natl Acad Sci U S A* 96, 4983-4988 (1999).
51. Jensen, D.E. *et al.* BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene* 16, 1097-1112 (1998).
52. Somasundaram, K. *et al.* Arrest of the cell cycle by the tumour-suppressor BRCA1 requires the CDK-inhibitor p21WAF1/CiP1. *Nature* 389, 187-190 (1997).
53. Harkin, D.P. *et al.* Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1. *Cell* 97, 575-586 (1999).
54. Takekawa, M. & Saito, H. A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK. *Cell* 95, 521-530 (1998).
55. Chen, Y. *et al.* Aberrant subcellular localization of BRCA1 in breast cancer. *Science* 270, 789-791 (1995).
56. Thomas, J.E. *et al.* Subcellular localization and analysis of apparent 180-kDa and 220-kDa proteins of the breast cancer susceptibility gene, BRCA1. *J Biol Chem* 271, 28630-28635 (1996).

57. Wilson, C.A. *et al.* Differential subcellular localization, expression and biological toxicity of BRCA1 and the splice variant BRCA1-delta11b. *Oncogene* 14, 1-16 (1997).
58. Chen, C.F. *et al.* The nuclear localization sequences of the BRCA1 protein interact with the importin-alpha subunit of the nuclear transport signal receptor. *J Biol Chem* 271, 32863-32868 (1996).
59. Chen, Y. *et al.* BRCA1 is a 220-kDa nuclear phosphoprotein that is expressed and phosphorylated in a cell cycle-dependent manner. *Cancer Res* 56, 3168-3172 (1996).
60. Zabludoff, S.D., Wright, W.W., Harshman, K. & Wold, B.J. BRCA1 mRNA is expressed highly during meiosis and spermiogenesis but not during mitosis of male germ cells. *Oncogene* 13, 649-653 (1996).
61. Marquis, S.T. *et al.* The developmental pattern of Brca1 expression implies a role in differentiation of the breast and other tissues. *Nat Genet* 11, 17-26 (1995).
62. Andres, J.L. *et al.* Regulation of BRCA1 and BRCA2 expression in human breast cancer cells by DNA-damaging agents. *Oncogene* 16, 2229-2241 (1998).
63. Ribieras, S. *et al.* Abundance of BRCA1 transcripts in human cancer and lymphoblastoid cell lines carrying BRCA1 germ-line alterations. *Int J Cancer* 73, 715-718 (1997).
64. Wilson, C.A. *et al.* Localization of human BRCA1 and its loss in high-grade, non-inherited breast carcinomas. *Nat Genet* 21, 236-240 (1999).
65. Magdinier, F., Ribieras, S., Lenoir, G.M., Frappart, L. & Dante, R. Down-regulation of BRCA1 in human sporadic breast cancer; analysis of DNA methylation patterns of the putative promoter region. *Oncogene* 17, 3169-3176 (1998).
66. Thompson, M.E., Jensen, R.A., Obermiller, P.S., Page, D.L. & Holt, J.T. Decreased expression of BRCA1 accelerates growth and is often present during sporadic breast cancer progression. *Nat Genet* 9, 444-450 (1995).
67. Spillman, M.A. & Bowcock, A.M. BRCA1 and BRCA2 mRNA levels are coordinately elevated in human breast cancer cells in response to estrogen. *Oncogene* 13, 1639-1645 (1996).
68. Marks, J.R. *et al.* BRCA1 expression is not directly responsive to estrogen. *Oncogene* 14, 115-121 (1997).
69. Futreal, P.A. *et al.* BRCA1 mutations in primary breast and ovarian carcinomas. *Science* 266, 120-122 (1994).
70. Puget, N. *et al.* A 1-kb Alu-mediated germ-line deletion removing BRCA1 exon 17. *Cancer Res* 57, 828-831 (1997).
71. Swensen, J., Hoffman, M., Skolnick, M.H. & Neuhausen, S.L. Identification of a 14 kb deletion involving the promoter region of BRCA1 in a breast cancer family. *Hum Mol Genet* 6, 1513-1517 (1997).
72. Puget, N. *et al.* Screening for germ-line rearrangements and regulatory mutations in BRCA1 led to the identification of four new deletions. *Cancer Res* 59, 455-461 (1999).
73. Stoppa-Lyonnet, D. *et al.* Segregation of two BRCA1 mutations in a single family. *Am J Hum Genet* 59, 479-481 (1996).
74. Ford, D., Easton, D.F. & Peto, J. Estimates of the gene frequency of BRCA1 and its contribution to breast and ovarian cancer incidence. *Am J Hum Genet* 57, 1457-1462 (1995).
75. Roa, B.B., Boyd, A.A., Volcik, K. & Richards, C.S. Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2. *Nat Genet* 14, 185-187 (1996).
76. Easton, D.F., Bishop, D.T., Ford, D. & Crockford, G.P. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* 52, 678-701 (1993).

77. Ford, D. *et al.* Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* 62, 676-689 (1998).
78. Narod, S.A. *et al.* An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. Breast Cancer Linkage Consortium. *Am J Hum Genet* 56, 254-264 (1995).
79. Struewing, J.P. *et al.* The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nat Genet* 11, 198-200 (1995).
80. Easton, D.F., Ford, D. & Bishop, D.T. Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Am J Hum Genet* 56, 265-271 (1995).
81. Gayther, S.A. *et al.* Germline mutations of the BRCA1 gene in breast and ovarian cancer families provide evidence for a genotype-phenotype correlation. *Nat Genet* 11, 428-433 (1995).
82. Neuhausen, S.L. *et al.* Haplotype and phenotype analysis of six recurrent BRCA1 mutations in 61 families: results of an international study. *Am J Hum Genet* 58, 271-280 (1996).
83. Holt, J.T. *et al.* Growth retardation and tumour inhibition by BRCA1. *Nat Genet* 12, 298-302 (1996).
84. Ford, D., Easton, D.F., Bishop, D.T., Narod, S.A. & Goldgar, D.E. Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Lancet* 343, 692-695 (1994).
85. Rebbeck, T.R. *et al.* Modification of BRCA1-Associated Breast Cancer Risk by the Polymorphic Androgen-Receptor CAG Repeat. *Am J Hum Genet* 64, 1371-1377 (1999).
86. Phelan, C.M. *et al.* Ovarian cancer risk in BRCA1 carriers is modified by the HRAS1 variable number of tandem repeat (VNTR) locus. *Nat Genet* 12, 309-311 (1996).
87. Johannsson, O.T., Ranstam, J., Borg, A. & Olsson, H. Survival of BRCA1 breast and ovarian cancer patients: a population- based study from southern Sweden. *J Clin Oncol* 16, 397-404 (1998).
88. Lee, J.S. *et al.* Survival after breast cancer in Ashkenazi Jewish BRCA1 and BRCA2 mutation carriers. *J Natl Cancer Inst* 91, 259-263 (1999).
89. Porter, D.E., Dixon, M., Smyth, E. & Steel, C.M. Breast cancer survival in BRCA1 carriers. *Lancet* 341, 184-185 (1993).
90. Foulkes, W.D. *et al.* Germ-line BRCA1 mutation is an adverse prognostic factor in Ashkenazi Jewish women with breast cancer. *Clin Cancer Res* 3, 2465-2469 (1997).
91. Rubin, S.C. *et al.* Clinical and pathological features of ovarian cancer in women with germ- line mutations of BRCA1. *N Engl J Med* 335, 1413-1416 (1996).
92. Pharoah, P.D., Easton, D.F., Stockton, D.L., Gayther, S. & Ponder, B.A. Survival in familial, BRCA1-associated, and BRCA2-associated epithelial ovarian cancer. United Kingdom Coordinating Committee for Cancer Research (UKCCCR) Familial Ovarian Cancer Study Group. *Cancer Res* 59, 868-871 (1999).
93. Niederacher, D. *et al.* Patterns of allelic loss on chromosome 17 in sporadic breast carcinomas detected by fluorescent-labeled microsatellite analysis. *Genes Chromosomes Cancer* 18, 181-192 (1997).
94. Silva, J.M. *et al.* Loss of heterozygosity in BRCA1 and BRCA2 markers and high-grade malignancy in breast cancer. *Breast Cancer Res Treat* 53, 9-17 (1999).
95. Cornelis, R.S. *et al.* High allele loss rates at 17q12-q21 in breast and ovarian tumors from BRCAl-linked families. The Breast Cancer Linkage Consortium. *Genes Chromosomes Cancer* 13, 203-210 (1995).
96. Neuhausen, S.L. & Marshall, C.J. Loss of heterozygosity in familial tumors from three BRCA1-linked kindreds. *Cancer Res* 54, 6069-6072 (1994).
97. Tirkkonen, M. *et al.* Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. *Cancer Res* 57, 1222-1227 (1997).

98. Papp, J. *et al.* Allele loss from large regions of chromosome 17 is common only in certain histological subtypes of ovarian carcinomas. *Br J Cancer* 74, 1592-1597 (1996).
99. Niederacher, D. *et al.* Loss of heterozygosity of BRCA1, TP53 and TCRD markers analysed in sporadic endometrial cancer. *Eur J Cancer* 34, 1770-1776 (1998).
100. Semba, S., Yokozaki, H., Yasui, W. & Tahara, E. Frequent microsatellite instability and loss of heterozygosity in the region including BRCA1 (17q21) in young patients with gastric cancer. *Int J Oncol* 12, 1245-1251 (1998).
101. Armes, J.E. *et al.* Distinct molecular pathogeneses of early-onset breast cancers in BRCA1 and BRCA2 mutation carriers: a population-based study. *Cancer Res* 59, 2011-2017 (1999).
102. Tomlinson, G.E. *et al.* Characterization of a breast cancer cell line derived from a germ-line BRCA1 mutation carrier. *Cancer Res* 58, 3237-3242 (1998).
103. Bennett, L.M. *et al.* Isolation of the mouse homologue of BRCA1 and genetic mapping to mouse chromosome 11. *Genomics* 29, 576-581 (1995).
104. Gowen, L.C., Johnson, B.L., Latour, A.M., Sulik, K.K. & Koller, B.H. Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nat Genet* 12, 191-194 (1996).
105. Hakem, R. *et al.* The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse. *Cell* 85, 1009-1023 (1996).
106. Xu, X. *et al.* Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. *Mol Cell* 3, 389-395 (1999).
107. Xu, X. *et al.* Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. *Nat Genet* 22, 37-43 (1999).
108. Stratton, M.R. *et al.* Familial male breast cancer is not linked to the BRCA1 locus on chromosome 17q. *Nat Genet* 7, 103-107 (1994).
109. Gudmundsson, J. *et al.* Frequent occurrence of BRCA2 linkage in Icelandic breast cancer families and segregation of a common BRCA2 haplotype. *Am J Hum Genet* 58, 749-756 (1996).
110. Thorlacius, S. *et al.* Linkage to BRCA2 region in hereditary male breast cancer. *Lancet* 346, 544-545 (1995).
111. Schutte, M. *et al.* An integrated high-resolution physical map of the DPC/BRCA2 region at chromosome 13q12. *Cancer Res* 55, 4570-4574 (1995).
112. Tavtigian, S.V. *et al.* The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat Genet* 12, 333-337 (1996).
113. Wagner, T.M.U. *et al.* Global sequence diversity of BRCA2: analysis of 71 breast cancer families and 95 control individuals of worldwide populations. *Hum Mol Genet* 8, 413-423 (1999).
114. Bignell, G., Micklem, G., Stratton, M.R., Ashworth, A. & Wooster, R. The BRC repeats are conserved in mammalian BRCA2 proteins. *Hum Mol Genet* 6, 53-58 (1997).
115. Bork, P., Blomberg, N. & Nilges, M. Internal repeats in the BRCA2 protein sequence. *Nat Genet* 13, 22-23 (1996).
116. Milner, J., Ponder, B., Hughes-Davies, L., Seltmann, M. & Kouzarides, T. Transcriptional activation functions in BRCA2. *Nature* 386, 772-773 (1997).
117. Wong, A.K.C., Pero, R., Ormonde, P.A., Tavtigian, S.V. & Bartel, P.L. RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2. *J Biol Chem* 272, 31941-31944 (1997).
118. Marmorstein, L.Y., Ouchi, T. & Aaronson, S.A. The BRCA2 gene product functionally interacts with p53 and RAD51. *Proc Natl Acad Sci U S A* 95, 13869-13874 (1998).
119. Fuks, F., Milner, J. & Kouzarides, T. BRCA2 associates with acetyltransferase activity when bound to P/CAF. *Oncogene* 17, 2531-2534 (1998).

120. Wooster, R. *et al.* Identification of the breast cancer susceptibility gene BRCA2. *Nature* 378, 789-792 (1995).
121. Bertwistle, D. *et al.* Nuclear location and cell cycle regulation of the BRCA2 protein. *Cancer Res* 57, 5485-5488 (1997).
122. Su, L.K. *et al.* Characterization of BRCA2: temperature sensitivity of detection and cell-cycle regulated expression. *Oncogene* 17, 2377-2381 (1998).
123. Vaughn, J.P. *et al.* Cell cycle control of BRCA2. *Cancer Res* 56, 4590-4594 (1996).
124. Zou, J.P., Hirose, Y., Siddique, H., Rao, V.N. & Reddy, E.S. Structure and expression of variant BRCA2a lacking the transactivation domain. *Oncol Rep* 6, 437-440 (1999).
125. Couch, F.J. *et al.* BRCA2 germline mutations in male breast cancer cases and breast cancer families. *Nat Genet* 13, 123-125 (1996).
126. Miki, Y., Katagiri, T., Kasumi, F., Yoshimoto, T. & Nakamura, Y. Mutation analysis in the BRCA2 gene in primary breast cancers. *Nat Genet* 13, 245-247 (1996).
127. Weber, B.H., Brohm, M., Stec, I., Backe, J. & Caffier, H. A somatic truncating mutation in BRCA2 in a sporadic breast tumor. *Am J Hum Genet* 59, 962-964 (1996).
128. Lancaster, J.M. *et al.* BRCA2 mutations in primary breast and ovarian cancers. *Nat Genet* 13, 238-240 (1996).
129. Takahashi, H. *et al.* Mutations of the BRCA2 gene in ovarian carcinomas. *Cancer Res* 56, 2738-2741 (1996).
130. Nordling, M. *et al.* A large deletion disrupts the exon 3 transcription activation domain of the BRCA2 gene in a breast/ovarian cancer family. *Cancer Res* 58, 1372-1375 (1998).
131. Thorlacius, S. *et al.* Study of a single BRCA2 mutation with high carrier frequency in a small population. *Am J Hum Genet* 60, 1079-1084 (1997).
132. Friedman, L.S. *et al.* Mutation analysis of BRCA1 and BRCA2 in a male breast cancer population. *Am J Hum Genet* 60, 313-319 (1997).
133. Malone, K.E. *et al.* Frequency of BRCA1/BRCA2 mutations in a population-based sample of young breast carcinoma cases. *Cancer* 88, 1393-1402 (2000).
134. Thorlacius, S. *et al.* Population-based study of risk of breast cancer in carriers of BRCA2 mutation. *Lancet* 352, 1337-1339 (1998).
135. Krainer, M. *et al.* Differential contributions of BRCA1 and BRCA2 to early-onset breast cancer. *N Engl J Med* 336, 1416-1421 (1997).
136. Healey, C.S. *et al.* A common variant in BRCA2 is associated with both breast cancer risk and prenatal viability. *Nat Genet* 26, 362-364 (2000).
137. Easton, D.F. *et al.* Cancer risks in two large breast cancer families linked to BRCA2 on chromosome 13q12-13. *Am J Hum Genet* 61, 120-128 (1997).
138. Gayther, S.A. *et al.* Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. *Nat Genet* 15, 103-105 (1997).
139. Berman, D.B. *et al.* A common mutation in BRCA2 that predisposes to a variety of cancers is found in both Jewish Ashkenazi and non-Jewish individuals. *Cancer Res* 56, 3409-3414 (1996).
140. Goggins, M. *et al.* Germline BRCA2 gene mutations in patients with apparently sporadic pancreatic carcinomas. *Cancer Res* 56, 5360-5364 (1996).
141. Gaffney, D.K. *et al.* Response to radiation therapy and prognosis in breast cancer patients with BRCA1 and BRCA2 mutations. *Radiother Oncol* 47, 129-136 (1998).
142. Verhoog, L.C. *et al.* Survival in hereditary breast cancer associated with germline mutations of BRCA2. *J Clin Oncol* 17, 3396-3402 (1999).

143. Collins, N. *et al.* Consistent loss of the wild type allele in breast cancers from a family linked to the BRCA2 gene on chromosome 13q12-13. *Oncogene* 10, 1673-1675 (1995).
144. Gudmundsson, J. *et al.* Different tumor types from BRCA2 carriers show wild-type chromosome deletions on 13q12-q13. *Cancer Res* 55, 4830-4832 (1995).
145. Tirkkonen, M. *et al.* Somatic genetic alterations in BRCA2-associated and sporadic male breast cancer. *Genes Chromosomes Cancer* 24, 56-61 (1999).
146. BCLC. Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Breast Cancer Linkage Consortium. *Lancet* 349, 1505-1510 (1997).
147. Connor, F. *et al.* Cloning, chromosomal mapping and expression pattern of the mouse Brca2 gene. *Hum Mol Genet* 6, 291-300 (1997).
148. Rajan, J.V., Wang, M., Marquis, S.T. & Chodosh, L.A. Brca2 is coordinately regulated with Brca1 during proliferation and differentiation in mammary epithelial cells. *Proc Natl Acad Sci U S A* 93, 13078-13083 (1996).
149. Patel, K.J. *et al.* Involvement of Brca2 in DNA repair. *Mol Cell* 1, 347-357 (1998).
150. Yu, V.P. *et al.* Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation. *Genes Dev* 14, 1400-1406 (2000).
151. Friedman, L.S. *et al.* Thymic lymphomas in mice with a truncating mutation in Brca2. *Cancer Res* 58, 1338-1343 (1998).
152. Cavalli-Sforza, L.L. & Piazza, A. Human genomic diversity in Europe: a summary of recent research and prospects for the future. *Eur J Hum Genet* 1, 3-18 (1993).
153. Guglielmino, C.R., Piazza, A., Menozzi, P. & Cavalli-Sforza, L.L. Uralic genes in Europe. *Am J Phys Anthropol* 83, 57-68 (1990).
154. Sajantila, A. *et al.* Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. *Proc Natl Acad Sci U S A* 93, 12035-12039 (1996).
155. Nevanlinna, H.R. The Finnish population structure. A genetic and genealogical study. *Hereditas* 71, 195-236 (1972).
156. de la Chapelle, A. Disease gene mapping in isolated human populations: the example of Finland. *J Med Genet* 30, 857-865 (1993).
157. Ikaheimo, I., Silvennoinen-Kassinen, S. & Tiilikainen, A. HLA five-locus haplotypes in Finns. *Eur J Immunogenet* 23, 321-328 (1996).
158. Kittles, R.A. *et al.* Dual origins of Finns revealed by Y chromosome haplotype variation. *Am J Hum Genet* 62, 1171-1179 (1998).
159. Eriksson, A.W. Genetic polymorphisms in Finno-Ugrian populations. Finns, Lapps and Maris. *Isr J Med Sci* 9, 1156-1170 (1973).
160. Norio, R., Nevanlinna, H.R. & Perheentupa, J. Hereditary diseases in Finland; rare flora in rare soil. *Ann Clin Res* 5, 109-141 (1973).
161. Hastbacka, J. *et al.* Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nat Genet* 2, 204-211 (1992).
162. Ekelund, J. *et al.* Genome-wide scan for schizophrenia in the Finnish population: evidence for a locus on chromosome 7q22. *Hum Mol Genet* 9, 1049-1057 (2000).
163. Botstein, D., White, R.L., Skolnick, M. & Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32, 314-331 (1980).
164. Jeffreys, A.J., Wilson, V. & Thein, S.L. Hypervariable 'minisatellite' regions in human DNA. *Nature* 314, 67-73 (1985).

165. Weber, J.L. & May, P.E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44, 388-396 (1989).
166. Gray, I.C., Campbell, D.A. & Spurr, N.K. Single nucleotide polymorphisms as tools in human genetics. *Hum Mol Genet* 9, 2403-2408 (2000).
167. Terwilliger, J.D. & Ott, J. *Handbook of genetic linkage* (The Johns Hopkins University Press, Baltimore (1994).
168. Clerget-Darpoux, F., Bonaiti-Pellie, C. & Hochez, J. Effects of misspecifying genetic parameters in lod score analysis. *Biometrics* 42, 393-399 (1986).
169. Risch, N. & Giuffra, L. Model misspecification and multipoint linkage analysis. *Hum Hered* 42, 77-92 (1992).
170. Kruglyak, L., Daly, M.J., Reeve-Daly, M.P. & Lander, E.S. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58, 1347-1363 (1996).
171. Ploughman, L.M. & Boehnke, M. Estimating the power of a proposed linkage study for a complex genetic trait. *Am J Hum Genet* 44, 543-551 (1989).
172. Lander, E.S. & Schork, N.J. Genetic dissection of complex traits. *Science* 265, 2037-2048 (1994).
173. Smith, C. Testing for heterogeneity of recombination fraction values in human genetics. *Ann Hum Genet* 27, 175-182 (1963).
174. Murray, J.C. *et al.* A comprehensive human linkage map with centimorgan density. Cooperative Human Linkage Center (CHLC). *Science* 265, 2049-2054 (1994).
175. Risch, N. Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am J Hum Genet* 46, 229-241 (1990).
176. Weeks, D.E. & Lange, K. The affected-pedigree-member method of linkage analysis. *Am J Hum Genet* 42, 315-326 (1988).
177. Haines, J.L. & Pericak-Vance, M.A. *Approaches to Gene Mapping in Complex Human Diseases*. Wiley-Liss, New York (1998).
178. Stephens, J.C., Briscoe, D. & O'Brien, S.J. Mapping by admixture linkage disequilibrium in human populations: limits and guidelines. *Am J Hum Genet* 55, 809-824 (1994).
179. Spielman, R.S., McGinnis, R.E. & Ewens, W.J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52, 506-516 (1993).
180. Kallioniemi, A. *et al.* Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258, 818-821 (1992).
181. Hemminki, A. *et al.* Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. *Nat Genet* 15, 87-90 (1997).
182. Jenne, D.E. *et al.* Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nat Genet* 18, 38-43 (1998).
183. Desper, R. *et al.* Inferring tree models for oncogenesis from comparative genome hybridization data. *J Comput Biol* 6, 37-51 (1999).
184. Su, Y.A. *et al.* Identification of tumor-suppressor genes using human melanoma cell lines UACC903, UACC903(+6), and SRS3 by comparison of expression profiles. *Mol Carcinog* 28, 119-127 (2000).
185. Jiang, H., Kang, D.C., Alexandre, D. & Fisher, P.B. RaSH, a rapid subtraction hybridization approach for identifying and cloning differentially expressed genes. *Proc Natl Acad Sci U S A* 97, 12684-12689. (2000).
186. Dowton, S.B. & Slaugh, R.A. Diagnosis of human heritable diseases--laboratory approaches and outcomes. *Clin Chem* 41, 785-794; discussion 813-786 (1995).

187. Petrij-Bosch, A. *et al.* BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nat Genet* 17, 341-345 (1997).
188. Nollau, P. & Wagener, C. Methods for detection of point mutations: performance and quality assessment. IFCC Scientific Division, Committee on Molecular Biology Techniques. *Clin Chem* 43, 1114-1128 (1997).
189. Sanger, F., Nicklen, S. & Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74, 5463-5467 (1977).
190. Orita, M., Suzuki, Y., Sekiya, T. & Hayashi, K. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 5, 874-879 (1989).
191. Ravnik-Glavac, M., Glavac, D. & Dean, M. Sensitivity of single-strand conformation polymorphism and heteroduplex method for mutation detection in the cystic fibrosis gene. *Hum Mol Genet* 3, 801-807 (1994).
192. Cotton, R.G. Mutation detection and mutation databases. *Clin Chem Lab Med* 36, 519-522 (1998).
193. Nagamine, C.M., Chan, K. & Lau, Y.F. A PCR artifact: generation of heteroduplexes. *Am J Hum Genet* 45, 337-339 (1989).
194. Glavac, D. & Dean, M. Applications of heteroduplex analysis for mutation detection in disease genes. *Hum Mutat* 6, 281-287 (1995).
195. Lambrinakos, A., Humphrey, K.E., Babon, J.J., Ellis, T.P. & Cotton, R.G. Reactivity of potassium permanganate and tetraethylammonium chloride with mismatched bases and a simple mutation detection protocol. *Nucleic Acids Res* 27, 1866-1874 (1999).
196. Youil, R., Kemper, B. & Cotton, R.G. Detection of 81 of 81 known mouse beta-globin promoter mutations with T4 endonuclease VII--the EMC method. *Genomics* 32, 431-435 (1996).
197. Roest, P.A., Roberts, R.G., Sugino, S., van Ommen, G.J. & den Dunnen, J.T. Protein truncation test (PTT) for rapid detection of translation-terminating mutations. *Hum Mol Genet* 2, 1719-1721 (1993).
198. Southern, E.M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98, 503-517 (1975).
199. Friedman, K.J., Highsmith, W.E., Jr., Prior, T.W., Perry, T.R. & Silverman, L.M. Cystic fibrosis deletion mutation detected by PCR-mediated site-directed mutagenesis. *Clin Chem* 36, 695-696 (1990).
200. Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B. & Erlich, H.A. Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. *Nature* 324, 163-166 (1986).
201. Landegren, U., Kaiser, R., Sanders, J. & Hood, L. A ligase-mediated gene detection technique. *Science* 241, 1077-1080 (1988).
202. Sokolov, B.P. Primer extension technique for the detection of single nucleotide in genomic DNA. *Nucleic Acids Res* 18, 3671 (1990).
203. Pastinen, T., Kurg, A., Metspalu, A., Peltonen, L. & Syvanen, A.C. Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res* 7, 606-614 (1997).
204. Smith, J.R. *et al.* Approach to genotyping errors caused by nontemplated nucleotide addition by Taq DNA polymerase. *Genome Res* 5, 312-317 (1995).
205. Dib, C. *et al.* A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380, 152-154 (1996).
206. Couch, F.J. *et al.* Generation of an integrated transcription map of the BRCA2 region on chromosome 13q12-q13. *Genomics* 36, 86-99 (1996).
207. Brodeur, G.M., Tsiatis, A.A., Williams, D.L., Luthardt, F.W. & Green, A.A. Statistical analysis of cytogenetic abnormalities in human cancer cells. *Cancer Genet Cytogenet* 7, 137-152 (1982).

208. Ott, J. *Analysis of Human Genetic Linkage*. Johns Hopkins University Press, Baltimore (1999).
209. O'Connell, J.R. & Weeks, D.E. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63, 259-266 (1998).
210. Ott, J. Computer-simulation methods in human linkage analysis. *Proc Natl Acad Sci U S A* 86, 4175-4178 (1989).
211. Lathrop, G.M., Lalouel, J.M., Julier, C. & Ott, J. Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci U S A* 81, 3443-3446 (1984).
212. Cottingham, R.W., Jr., Idury, R.M. & Schaffer, A.A. Faster sequential genetic linkage computations. *Am J Hum Genet* 53, 252-263 (1993).
213. Schaffer, A.A., Gupta, S.K., Shriram, K. & Cottingham, R.W., Jr. Avoiding recomputation in linkage analysis. *Hum Hered* 44, 225-237 (1994).
214. Duffy, D. *Am.J. Hum. Genet. suppl. 61* , A197 (1997).
215. Shapiro, M.B. & Senapathy, P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* 15, 7155-7174 (1987).
216. Huusko, P. *et al.* Evidence of Founder Mutations in Finnish BRCA1 and BRCA2 Families. *Am J Hum Genet* 62, 1544-1548 (1998).
217. Peelen, T. *et al.* A high proportion of novel mutations in BRCA1 with strong founder effects among Dutch and Belgian hereditary breast and ovarian cancer families. *Am J Hum Genet* 60, 1041-1049 (1997).
218. Wagner, T.M. *et al.* BRCA1-related breast cancer in Austrian breast and ovarian cancer families: specific BRCA1 mutations and pathological characteristics. *Int J Cancer* 77, 354-360 (1998).
219. Backe, J. *et al.* Frequency of BRCA1 mutation 5382insC in German breast cancer patients. *Gynecol Oncol* 72, 402-406 (1999).
220. Abeliovich, D. *et al.* The founder mutations 185delAG and 5382insC in BRCA1 and 6174delT in BRCA2 appear in 60% of ovarian cancer and 30% of early-onset breast cancer patients among Ashkenazi women. *Am J Hum Genet* 60, 505-514 (1997).
221. Thorlacius, S. *et al.* A single BRCA2 mutation in male and female breast cancer families from Iceland with varied cancer phenotypes. *Nat Genet* 13, 117-119 (1996).
222. Haraldsson, K. *et al.* BRCA2 germ-line mutations are frequent in male breast cancer patients without a family history of the disease. *Cancer Res* 58, 1367-1371 (1998).
223. Hakansson, S. *et al.* Moderate frequency of BRCA1 and BRCA2 germ-line mutations in Scandinavian familial breast cancer. *Am J Hum Genet* 60, 1068-1078 (1997).
224. Chen, J. *et al.* BRCA2 germline mutations in Swedish breast cancer families. *Eur J Hum Genet* 6, 134-139 (1998).
225. Mavraki, E., Gray, I.C., Bishop, D.T. & Spurr, N.K. Germline BRCA2 mutations in men with breast cancer. *Br J Cancer* 76, 1428-1431 (1997).
226. Schubert, E.L. *et al.* BRCA2 in American families with four or more cases of breast or ovarian cancer: recurrent and novel mutations, variable expression, penetrance, and the possibility of families whose cancer is not attributable to BRCA1 or BRCA2. *Am J Hum Genet* 60, 1031-1040 (1997).
227. Zlotogora, J. High frequencies of human genetic diseases: founder effect with genetic drift or selection? *Am J Med Genet* 49, 10-13 (1994).
228. Sarantaus, L. *et al.* Multiple founder effects and geographical clustering of BRCA1 and BRCA2 families in Finland. *Eur J Hum Genet* 8, 757-763 (2000).
229. Szabo, C.I. & King, M.C. Population genetics of BRCA1 and BRCA2. *Am J Hum Genet* 60, 1013-1020 (1997).

230. Unger, M.A. *et al.* Screening for genomic rearrangements in families with breast and ovarian cancer identifies BRCA1 mutations previously missed by conformation-sensitive gel electrophoresis or sequencing. *Am J Hum Genet* 67, 841-850 (2000).
231. Catteau, A. *et al.* Identification of a C/G polymorphism in the promoter region of the BRCA1 gene and its use as a marker for rapid detection of promoter deletions. *Br J Cancer* 79, 759-763 (1999).
232. Huusko, P. *et al.* Germ-line TP53 mutations in Finnish cancer families exhibiting features of the Li-Fraumeni syndrome and negative for BRCA1 and BRCA2. *Cancer Genet Cytogenet* 112, 9-14 (1999).
233. Weber, J.L. & Wong, C. Mutation of human short tandem repeats. *Hum Mol Genet* 2, 1123-1128 (1993).
234. Callen, D.F. *et al.* Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *Am J Hum Genet* 52, 922-927 (1993).
235. Li, W.-H. & Graur, D. *Fundamentals of Molecular Evolution*. Sinauer Associates Inc., Sunderland (1991).
236. Schmucker, B. & Krawczak, M. Meiotic microdeletion breakpoints in the BRCA1 gene are significantly associated with symmetric DNA-sequence elements. *Am J Hum Genet* 61, 1454-1456 (1997).
237. Xu, C.F. *et al.* Mutations and alternative splicing of the BRCA1 gene in UK breast/ovarian cancer families. *Genes Chromosomes Cancer* 18, 102-110 (1997).
238. Laiho, E. *et al.* Transglutaminase 1 mutations in autosomal recessive congenital ichthyosis: private and recurrent mutations in an isolated population. *Am J Hum Genet* 61, 529-538 (1997).
239. Eiriksdottir, G. *et al.* Mapping loss of heterozygosity at chromosome 13q: loss at 13q12-q13 is associated with breast tumour progression and poor prognosis. *Eur J Cancer* 34, 2076-2081 (1998).
240. Smith, S.A., Easton, D.F., Evans, D.G. & Ponder, B.A. Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. *Nat Genet* 2, 128-131 (1992).
241. Arason, A. *et al.* A population study of mutations and LOH at breast cancer gene loci in tumours from sister pairs: two recurrent mutations seem to account for all BRCA1/BRCA2 linked breast cancer in Iceland. *J Med Genet* 35, 446-449 (1998).