Search for new breast cancer susceptibility genes

Oldenburg, R.A.

Search for new breast cancer susceptibility genes

The Dutch Cancer Society provided financial support for the author (RUL1999-2021)
The printing of this thesis was sponsored by: Roche NimbleGen, Boehringer Ingelheim BV,
Michiel Oldenburg, Merck Sharp & Dohme BV, The Dutch Cancer Society,
my father Hans Oldenburg, Agnes Oldenburg-Herpers and my Mother Nibs Bloem.

Design: Pepijn Oldenburg, Oldenburg Visuele Communicatie, Leiden
Printed by CPI Wöhrmann Print Service

Voor Jet, Michelle en Cathelijne.

SEARCH FOR NEW BREAST CANCER SUSCEPTIBILITY GENES

PROEFSCHRIFT

Ter verkrijging van de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties te verdedigen
op donderdag 29 mei 2008 klokke 15:00 uur

door Rogier Abel Oldenburg
geboren te Lower Hutt, Nieuw Zeeland in 1968

PROMOTIECOMMISSIE

Promotores
Prof. Dr. P. Devilee
Prof. Dr. E.J. Meijers-Heijboer (VU Medisch Centrum, Amsterdam)
Prof. Dr. C.J. Cornelisse

Referent
Prof. Dr. M.R. Stratton (Institute of Cancer Research, London, UK)

Overige Leden
Prof. J.G.M. Klijn (Erasmus Medical Centre, Rotterdam)
Dr. P.M. Nederlof (Nederlands Kanker Instituut, Amsterdam)

CONTENTS

CHAPTER 1

AIMS AND OUTLINE OF THE THESIS

Worldwide, breast cancer is the most commonly occurring cancer among women. It accounts for 22% of all female cancers and the estimated annual incidence of breast cancer is about one million cases. Many risk factors have been identified but a positive family history remains among the most important ones established for breast cancer, with first-degree relatives of patients having an approximately two-fold elevated risk. It is currently estimated that approximately 20-25% of this risk is explained by known breast cancer susceptibility genes, mostly those conferring high risks, such as *BRCA1* and *BRCA2*.

However, these genes explain less than 5% of the total breast cancer incidence, even though several studies have suggested that the proportion of breast cancer that can be attributed to a genetic factor may be as high as 30%. It is thus likely that there are still breast cancer susceptibility genes to be found. It is presently not known how many such genes there still are, nor how many will fall into the class of rare high-risk (e.g. *BRCAX*) or of common low-risk susceptibility genes, nor if and how these factors interact with each other to cause susceptibility (a polygenic model). In general high-risk susceptibility genes will cause typical breast cancer families, which are characterized by breast cancer at an early age, bilateral breast cancer, the occurrence of other specific cancer types in the family (for example ovarian cancer or male breast cancer) and an autosomal dominant inheritance pattern.

On the other hand individual low-risk genes probably do not couse familial clustering of breast cancer. However it is possible that if there are many low-risk genes, different combinations of such genes could be involved in individual breast cancer susceptibility and familial clustering of breast cancer might occur. Early work of the Breast Cancer Linkage Consortium (BCLC) showed that respectively 52% and 32% of families with at least four cases of breast cancer diagnosed under 60 are caused by *BRCA1* and *BRCA2*. When selecting families with breast cancer and one or more cases with ovarian cancer 81% of the families is explained by *BRCA1* and 14% by *BRCA2*. However, when selecting families with four or more cases of breast cancer diagnosed under 60 and no cases of ovarian cancer or male breast cancer only 33% could be explained by *BRCA1* and *BRCA2* together. In some of these families the breast cancer will not be inherited, but on the whole this group is too big to be totally explained by coincidental clustering. More likely, most of these families are explained by mutations in other unknown genes.

The objective of this thesis is to describe our endeavours to localize new high-risk breast cancer susceptibility genes by genome wide linkage analysis and to set the first steps in isolating these genes. For this purpose we selected families which had to satisfy the following criteria: (1) at least three women diagnosed with breast cancer below age 60 years, (2) no case of ovarian cancer or male breast cancer in a blood relative (since these phenotypes are strongly predictive of the presence of *BRCA1* or *BRCA2* mutation), and (3) DNA samples available for genotyping from at least three women affected with breast cancer. In addition, to minimize the probability that the family segregated a *BRCA1* or *BRCA2* mutation, DNA from at least one affected individual was screened for mutations across both genes. Whenever possible a second affected individual was screened. Subsequently, we collected genotype data on at least three microsatellite markers flanking the *BRCA1* and *BRCA2* loci. Families with insufficient mutation screening or linkage data were not included in further analyses. Due to the excellent structure of the eight different departments of clinical genetics in the Netherlands and the willingness of the pathological departments to cooperate, it was relatively easy to collect data and tumor material from sufficient families.

One of the families we selected for the genome-wide linkage analysis harbours an extraordinarily high number of tumours, comprising, breast, lung, colon cancers, malignant melanoma and oral squamous cell carcinomas (OSCC). In this family a *p16-Leiden* germline mutation was found. Other researchers suggested a relationship between *p16* germline mutations and breast cancer. Therefore we studied the possibility of *p16* acting as a breast cancer susceptibility gene. See chapter 3.1. In the meanwhile Meijers-Heijboer et al.[175] identified *CHEK2* as a low-risk breast cancer susceptibility allele and Kainu et al.[228] suggested the 13q21 region as a candidate breast cancer susceptibility locus. Chapter 3.2 describes the role of the *CHEK2*1100delC* mutation in causing breast cancer in our group of families. As described in chapter 3.3 we could not confirm the claim by Kainu.

One of the biggest problems one might encounter in linkage analysis is the extent of genetic heterogeneity in the selected families. Chapter 4 describes attempts to subclassify the heterogenic group of families in more homogeneous groups of families by determining tumor characteristics.

Chapter 5.1 describes the results of the international genome wide linkage analysis conducted by the BCLC. Chapter 5.2 presents the genomewide linkage analysis in the Dutch population and in which suggestive linkage for a new breast cancer susceptibility locus at 9q was identified.

CHAPTER 2

GENERAL INTRODUCTION

Based on the article:

GENETIC SUSCEPTIBILITY FOR BREAST CANCER: HOW MANY MORE GENES TO BE FOUND?

R.A. Oldenburg, H. Meijers-Heijboer, C.J. Cornelisse, P. Devilee
Critical Reviews in Oncology/Hematology. 2007 Aug; 63(2): 125-49

1. BACKGROUND

Breast tumors have been noted since antiquity and were probably first described in the Edwin Smith surgical papyrus originating from Egypt at around 2.500 BC.[1] In this document tumors were described as 'cold and hard to the touch' whereas abscesses were 'hot'.

Adenocarcinomas represent the vast majority of invasive malignant breast tumors and are believed to originate from the mammary parenchymal epithelium, particularly cells of the terminal duct lobular unit (TDLU). These tumors are characterized by invasion of adjacent tissues and a marked tendency to metastasize to distant sites. The most common being the bones, lungs and pleurae, liver, adrenals, ovaries, skin and brain.

In the clinical practice breast cancer patients are classified in four stages. This is based on the clinical and pathological extent of the disease according to the TNM system, where T refers to tumor size, N to the presence of metastases in the local regional lymph nodes, and M to distant metastases (beyond the ipsilateral supraclavicular lymph nodes).

Histologically invasive breast carcinomas (and all other invasive tumors) are routinely graded based on the assessment of tubule/gland formation, nuclear pleomorphism and mitotic counts. In addition they are classified as well differentiated (grade I), moderately differentiated (grade II), or poorly differentiated (grade III). Both the TNM classification and the histological grade are associated significantly with survival and are now recognized as powerful prognostic factors.

Breast abnormalities should always be evaluated by triple assessment including clinical examination, imaging (mammography and ultrasound) and tissue sampling by either fine needle aspiration cytology or needle core biopsy.

There is a slightly higher frequency of invasive breast cancer in the left breast, with a left to right ratio of 1.07:1. Between 40 and 50% of the tumors occur in the upper outer quadrant of the breast. There is a decreasing order of frequency in the other quadrants from the central, upper inner, lower outer to the lower inner quadrant.[2]
Today, breast cancer is the most common occurring cancer amongst women. It accounts for 22% of all female cancers. The estimated annual incidence of breast cancer worldwide is about one million cases. A significant difference in the incidence rates of breast cancer has been observed between so-called low risk areas such as the Far East, Africa and South America, and the high-risk areas North America and Northern Europe. Together, the USA and Europe roughly account for 16% of the world population and 60% of the worldwide incidence of breast cancer.[3,4] Studies on migrants have demonstrated that breast cancer incidence increases in people who move from a region with a low incidence to a region with higher breast cancer incidence. This effect is then passed on to the next generation until, within one or two generations the migrant's descendents acquire the same breast cancer risk as the native population.[5,6] This underlines the crucial contribution of environmental factors to breast cancer risk. To date many other risk factors have been identified. See also Table 1 for presently known risks and protective factors for breast cancer.

2. BREAST CANCER RISK FACTORS

2.1. Ethnicity, gender and age

Incidence rates correlate with gender, ethnic origin and show age specific patterns. Compared to the female breast cancer incidence rate the incidence rate of male breast cancer is far less. Approximately one out of every 150 breast cancer cases occurs in a male.[7] Breast cancer incidence is less than 10 cases per 100.000 women aged 25 or younger and increases up to 10-fold by the age of 40.[8] In the United States, the incidence rates are 20-40% higher in white women than in African American women,[9] except in younger age groups where rates are higher in African-American than in white women.[10] The age- and geographic-specific differences become even more profound after menopause. In the USA and Sweden the age-specific risk continues to rise up to 75 years, while in Colombia, the age specific risk increase is considerably less after the age of 45. In contrast, in Japan breast cancer incidence after the age of 45 exhibits a plateau followed by a slow decrease.[8]

2.2. Hormonal factors

The extent and duration of exposure to sex hormones has been consistently identi-

TABLE 1

Summary of protective factors and factors that increase breast cancer risk

Genetic constitution	Positive family history of breast cancer (any first or second degree family member with breast cancer) Carrier of a know breast cancer susceptibility gene (see also table 3)
Demographic factors	Geographical region (Western Countries) Female sex Increasing age Low socio-economical status
Endogenous factors	Older age at menopause (>54 years) Early age of menarge (<12 years) Nulliparity and older age at first born No breastfeeding Low physical activity
Exogenous factors	Usage of oral contraceptives Usage of hormone replacement therapy Exposure to ionizing radiation at young adolescent age
Physical characteristics	Obesity in postmenopausal women Tall stature High insulin-like growth factor I (IGF-I) levels History of atypical proliferative benign breast disease History of breast cancer Dense tissue at mammography High bone density in postmenopausal women
Dietary factors	Alcohol use Low folate intake High intake of unsaturated fat and well-done meat
Protective factors	Geographical region (Asia, Africa) Early age of first full term pregnancy High parity Breast feeding Early age at menopause Obesitas in premenopausal women Fruit and vegetables consumption Physical activity Usage of non-steroidal anti inflammatory drugs Chemopreventive agents

fied as a risk factor in many epidemiological studies. This includes endogenous sex hormones related to the menstrual cycle, as well as exogenous hormones derived from contraceptives, hormonal replacement therapy (HRT) and diet.[11] The specific hormone or hormone combination responsible for breast cancer initiation has not been identified. However, estrogen is believed to be a major factor in modifying breast cancer risk. Two mechanisms have been proposed to explain the carcinogenicity of estrogens. Firstly, the receptor-mediated hormonal activity, which is generally related to stimulation of cellular proliferation result in more opportunities for the accumulation of genetic damage leading to carcinogenesis.[12] Secondly, the potential genotoxic activity of estrogen metabolites, in particular the hydroxylated (catechol) estrogens may lead to an increase of breast cancer risk.[13] Accordingly, longer periods of exposure are expected to increase breast cancer risk.

Early menarche (younger than 12 years of age compared to older than 14 years) increases the risk by 10-20%.[14,15] Delayed menopause increases it by approximately 3% for every one year increase in age of menopause.[16] Usage of exogenous hormones, such as hormone replacement therapy (especially a combination of progestin and estrogen) and oral contraceptives increases breast cancer risk as well. There is a small transient increase in the relative risk of breast cancer among users of oral contraceptives but, since use typically occurs at young age when breast cancer is relatively rare, such an increase has little effect on overall incidence rates.[16]

Surgically induced menopause (ovariectomy or hysterectomy) before the age of 35 decreases breast cancer risk by about 60% relative to women experiencing natural menopause.[17]

Epidemiological studies suggest that diets (particularly soy and unrefined grain products) rich in phytoestrogens, which embody several groups of nonsteroidal estrogens that are widely distributed within the plant kingdom, including isoflavones and lignans, may be associated with lower risk of breast cancer. However, much controversy exists regarding this subject, and there seems to be no clear evidence that phytoestrogen intake influences the risk of developing breast cancer.[18]

Obesity among postmenopausal women increases breast cancer risk. For every 5kg of weight gain above the lowest adult weight, breast cancer risk increases by 8%.[19-21] One plausible mechanism by which postmenopausal obesity increases the risk of breast cancer is through higher levels of endogenous estrogen present in obese women, as adipose tissue is an important source of estrogens.[22]

Studies in postmenopausal women have found a positive correlation between increased bone density and high breast cancer risk with the relative risk varying from 2.0

to 3.5.[23] Since estrogens help to maintain the bone mass, this correlation may again be explained by an increased total amount of estrogen.
Physical activity in adolescence and young adulthood decreases breast cancer risk with 20%. This effect maybe a result of delaying the onset of menarche and modifying the bioavailable hormone levels.[24,25] The use of antiestrogens (e.g. tamoxifen), early pregnancy, breastfeeding and higher parity also has a protective effect against breast cancer.

2.3. Other risk factors

2.3.1. Breast density

Women with a more than 75% increased breast density on mammography have an approximately five-fold increase in the risk of developing breast carcinoma over a woman with less than 5% increased breast density.[26,27] Null parity and high breast density seem to act synergistically since the risk increases sevenfold when they are both present in a person compared to parous women with low breast density.[28] Twin studies have shown that the population variation in the percentage of dense and non-dense tissue on mammography at a given age has a high heredity. Thus genetic factors probably play a large role in explaining the observed variation and finding the genes responsible for this phenotype could be important for understanding the causes of breast cancer.[27,29]

2.3.2. Benign breast disease

Some benign lesions are acknowledged risk factors for subsequent invasive breast cancer in the same area in the breast and are therefore considered precursor lesions. Severe atypical epithelial hyperplasia for example increases the risk of developing breast cancer four to five fold compared with women who do not have any proliferative changes in their breast. Women with this change and a family history of breast cancer (first degree relative) have a nine-fold increase in risk. Women with palpable cysts, complex fibro adenomas, duct papillomas, sclerosis adenosis, and moderate or florid epithelial hyperplasia have a slightly higher risk for breast cancer (1.5-3 times) than women without these changes.[17]

2.3.3. Radiation

Exposure of the mammary gland to high-dose ionizing radiation has been demonstrated to increase the risk of breast cancer. For example, long-term follow-up of women exposed to the Hiroshima or Nagasaki nuclear explosions indicates an incre-

ased risk of breast cancer, in particular for women exposed around puberty.[30] In addition, repeated fluoroscopies for treatment of tuberculosis, and more recently, treatment of women for Hodgkin's disease have been demonstrated to increase the risk of breast carcinoma also. The risk is dose-dependent and decreases gradually over time.[8,11,23]

2.4. MMTV

Another intriguing possibility, which potentially could explain a significant part of the breast cancer occurrence, was raised by the discovery of mouse mammary tumor virus (MMTV) in 1942. It has been postulated that a similar, or related, virus could be involved in the etiology of human breast cancer, which could potentially be of considerable clinical significance because this would permit the development of new preventive measures and treatment modalities and also raise the possibility of prophylactic and therapeutic vaccines. Today, viruses are believed to cause about 15% of all human cancers.[31,32,33,34,35]

Early studies were able to demonstrate MMTV-like virus particles in human breast cancer biopsies[36], cell-lines[37] and breast milk.[38] Wang et al.[39] found a 660-bp sequence of the *env* gene with 90-98% homology to MMTV, which could be detected in 38% of 314 unselected human breast carcinomas from the USA, but only in 1% in normal breast specimens. Similar findings have been reported by others.[40,41] Interestingly, a recently conducted gene expression analysis[42] identified a very similar percentage (40%) of cases with an interferon-inducible gene (IIG) signature, which may be a reflection of an immune response to viral infection. However, this is not the only reasonable explanation. The up regulation of IIG's may reflect the response of the cancer cells to interferon secreted by host immune cells.[43]

Despite the initial molecular findings, more recent observations have cast doubt on a role for MMTV-like viruses in the etiology of human breast cancer. The predominant fact is an inability of independent researchers to confirm an association between an MMTV-like virus and human breast cancer.[44,45] Others were able to detect PCR amplicons of the expected size, using the same PCR-condition described by Wang et al., but upon DNA-sequencing, all PCR-products turned out to be false-positive, comprising host genomic DNA.[46]

Besides these findings there are several other fundamental arguments against MMTV-like viruses playing a role in the etiology of breast cancer. For example, there is no evidence of transmission of human MMTV-like viruses via breast milk[47], as is the case for MMTV. Traces of MMTV are detected in normal mouse breast tissues. To date this

is not the case for human MMTV-like viruses. Pregnancy has a well-established protective effect against the risk of developing breast cancer in humans. The opposite is true for MMTV. In contrast to all established human oncogenic viruses, chronic immunosuppression does not predispose to breast cancer in humans[48,49] and, finally, human cells lack the receptor necessary for the viral entry of MMTV.[50] Thus, although the debate remains unsettled, it appears unlikely that an MMTV-like agent is a causal agent for breast cancer.

2.5.Family history

The Ancient Romans already noted the occurrence of familial clustering, but formal documentation began in the mid-nineteenth century.[51] Probably the oldest report of familial occurrence of breast cancer was written in 1757 by a French surgeon, Le Dran who had diagnosed a 19-year old nun with breast cancer and documented her family history of breast cancer.[52] Another French surgeon Broca, who in 1866 had observed an association between breast cancer and heredity in his wife's family, wrote the second oldest report of hereditary breast cancer. To date, a positive family history for breast cancer is a well established risk factor for breast cancer, with first-degree relatives of patients having an approximately two-fold elevated risk.[53] This risk increases with the number of affected relatives and is greater for women with relatives affected at a young age, bilateral disease or a history of benign breast disease.[17,54] About 13% of all patients have a first-degree relative with breast cancer. In Western countries, the overall lifetime risk for women who have no affected relative is 7.8%, for those who have one, the risk is 13.3%, and for those who have two, the risk is 21.1%.[53] The estimated probability for a woman aged 20 to develop breast cancer by age 50 is 1.7%, 3.7%, and 8.0%, respectively, for women with zero, one, and two affected first-degree relatives. Even in third - to fifth - degree relatives a significant increase in breast cancer risk has been observed.[55] Table 2 provides lifetime cumulative breast cancer risk estimates for women having a positive family history, which is widely used in the Dutch clinical genetic practice (based on Claus et al.[56]).

3. KNOWN BREAST CANCER SUSCEPTIBILITY GENES

To date up to 5-10% of all breast cancers are caused by germ-line mutations in well-identified breast cancer susceptibility genes. These genes can be roughly divided into 'high-risk' and 'low to moderate risk' breast cancer susceptibility genes. The high-risk breast cancer susceptibility genes include *BRCA1*, *BRCA2*, *PTEN*, *TP53*, *LKB1/STK11* and *CDH1,* with relative lifetime risks higher than 4 (but generally much

TABLE 2

Cumulative risk for breast cancer when having a positive family history (based on Claus et al.[56])

Age at diagnose family member	number of first degree family members with breast cancer						
	one first degree family member	Two first degree family members					
		Age at diagnose second first degree family member					
		20-29	30-39	40-49	50-59	60-69	70-79
20-29	21%	48%	46%	43%	40%	35%	31%
30-39	16%		44%	40%	35%	30%	25%
40-49	13%			35%	30%	25%	20%
50-59	11%				24%	19%	16%
60-69	10%					16%	13%
70-79	9%						11%

higher at young ages). The *CHEK2*, *TGFβ1*, *CASP8*, *BARD1*, *BRIP1*, *PALB2* and *ATM* genes belong to the 'low to moderate-risk' breast cancer susceptibility genes (see Table 3). The high-risk genes are the main cause for strong familial aggregation of breast cancer, and were mostly detected through linkage analysis (section 3.1). The low risk genes cannot be detected in this way because the relationship between genotype and phenotype is much weaker (section 3.2). The most widely used approach has been the association study, in which the allele frequencies of common variants within candidate genes are compared between a population of breast cancer cases and controls (Chapter 6). This research area has been problematic, however, because of the many associations that have been published to date, few have been established beyond reasonable doubt.[57,58] For example, one systematic meta-analysis examined 46 reports on 18 different genes.[57] Of the 12 significant associations reported, none were replicated by any of the other studies, and only four remained significant. For this reason, we will limit ourselves to those genes for which positive associations were replicated in independent studies.

3.1. High-risk breast cancer susceptibility genes

3.1.1. BRCA1 and BRCA2

The *BRCA1* gene is located on chromosome 17q21 and the *BRCA2* gene is located on chromosome 13q12.

TABLE 3

List of known high- and moderate to low risk breast cancer susceptibility genes

Gene	location	Gene Variant	Carrier status	Frequency	Breast Cancer Risk
BRCA1	17q21	Multiple	Heterozygous	Rare*	46-85% lifetime risk
BRCA2	13q12	Multiple	Heterozygous	Rare*	43-84% lifetime risk
TP53	17p13.1	Multiple	Heterozygous	Rare	28-56% by age 45
PTEN	10q23.3	Multiple	Heterozygous	Rare	25-50% lifetime risk
LKB1/STK11	19p13.3	Multiple	Heterozygous	Rare	29-54% lifetime risk
CDH1	16q22.1	Multiple	Heterozygous	Rare	20-40% lifetime risk
ATM	11q22-23	Multiple	Heterozygous	Moderate	RR: 2.2
TGFβ1	19q13.1	C-509T (promoter SNP)	Homozygous T	Frequent	OR: 1.25 (P=0.009)
		T-29C (L10P)	Homozygous C	Frequent	OR: 1.21 (P=0.01)
CASP8	2q33-34	G-1192C (D302H)	Heterozygous	Frequent	OR: 0.83
		G-1192C (D302H)	Homozygous H	Rare	OR: 0.58 (Ptrend=0.0002)
CASP10	2q33-34	G-1228A (V410I)	Heterozygous	Frequent	OR: 0.62 (P=0.0076)
CASP8/CASP10		410VI/II & 302DH/HH	Combination**	Moderate	OR: 0.37 (P=0.013)
BRIP1	17q22-24	Multiple	Heterozygous	Rare	RR: 2.0
PALB2	16p12	Multiple	Heterozygous	Rare	RR: 2.2
BARD1	2q34-35	Several (incl Cys557Ser)	Heterozygous	Moderate	OR: 2.6 (p=0.000003)
CHEK2	22q12.1	1100delC	Heterozygous	Moderate	RR: 2

* In, for example the Ashkenazi Jewish population some mutations have a moderate population frequency.

** Combination of the four different genotypes bearing the protective alleles of both CASP10 and CASP8 (i.e. 410VI-302DH, 410VI-302HH, 410II-302DH and 410II-302HH) compared with the most common genotype (410VV-302DD).

Rare: < 1% population frequency, Moderate 1-5%, Frequent >5%. OR = odds ratio, RR = relative risk

Although *BRCA1* and *BRCA2* do not share any obvious sequence homology, the parallels between the two genes are interesting. Both genes are reasonably large genes: *BRCA1* has 22 exons, spans approximately 100kb of genomic DNA, and encodes a 1863 amino acid protein, while *BRCA2* has 27 exons, spans around 70kb, and encodes a protein of 3418 amino acids.[59] They are both characterized by the presence of an extremely large exon 11. Both genes are ubiquitously expressed in humans with the highest levels in testis, ovaries and thymus. In contrast to most other known tu-

mor suppressor genes, they are relatively poorly conserved between other species, with the exception of a few small domains.

Both genes are generally considered to be 'caretaker' genes. Caretaker genes act as sensors of DNA damage and participate in the repair process. Their inactivation allows other genetic defects to accumulate and leads to genetic instability. In contrast, the so-called 'gatekeepers' directly control the progression of the cell cycle and their inactivation is thought to be sufficient to promote tumor growth.[60,61]

During the past decade many of the cellular and biochemical functions of the BRCA1- and BRCA2-proteins have been discovered. Together these suggest how BRCA1 and BRCA2 might play a role in carcinogenesis. For BRCA1 these roles include DNA-repair, protein ubiquitylation, chromatin remodeling and cell cycle checkpoint control. BRCA2 is involved in double-strand break DNA repair through homologous recombination, but little else is known about its function. These issues have been discussed in detail in several reviews.[62-65]

A rare form of Fanconi anemia (FA; FANCD1) was shown to be caused by biallelic mutations in *BRCA2*.[66] FA is a recessive disease of childhood that is characterized by specific birth defects, abnormal skin pigmentation, progressive bone-marrow failure and cancer susceptibility. Mutations in several genes can cause this condition, but all lead to chromosomal instability, which is similar to the chromosomal instability seen in *BRCA2*-deficient mice.[67] However, mutations in other FA genes are unlikely to be a major cause of highly penetrant breast cancer predisposition.[68,69]

Other studies have shown that in rare cases, children with medullablastoma or Wilms' tumor also carry two truncating *BRCA2* mutations.[70] Homozygosity for *BRCA1*-inactivating mutations, however, results in embryonic lethality, confirming the functional differences between the two proteins.

The prevalence of heterozygous carriers of high risk mutations in the general Caucasian population has been estimated to be about one in 1000 for *BRCA1*, and one in 750 for *BRCA2*.[71] However, in certain populations, this can be much higher due to the occurrence of founder mutations. For example, *BRCA2* analysis on 3,085 individuals from the same Ashkenazi Jewish population showed a carrier frequency of 1.52% for the 6174delT mutation.[72] This mutation appears to be restricted to the Ashkenazim, and has only once been reported in a person of proven non-Ashkenazi Jewish heritage.[73]

Germline mutations in *BRCA1* or *BRCA2* confer strong lifetime risks of breast cancer and ovarian cancer, together with smaller risks to some other cancer types.[54,74] Within the setting of multiple-case families, the cumulative risk of breast cancer at age 70

years in *BRCA1* and *BRCA2* mutation carriers was 85% and 84%, respectively, and of ovarian cancer 63% and 27%, respectively.[75] However, a more recent meta-analysis on 22 population-based and hospital-based studies showed that the average cumulative risks in *BRCA1*-mutation carriers by age 70 years were 65% for breast cancer and 39% for ovarian cancer. The corresponding estimates for *BRCA2* were 45% and 11%. In addition, in the American population, the estimated breast cancer and ovarian cancer risk at age 70 years are respectively 46% and 39% for *BRCA1* carriers and 43% and 22% in *BRCA2* carriers (Figure 1 and 2). The relative risks of breast cancer declined significantly with age for *BRCA1*-mutation carriers.[74,76] For *BRCA2*-mutation carriers this trend was also observed by Chen et al.[76] but not by Antoniou et al.[74] The estimates based on multiple-case families may have been enriched for mutations of higher risk and/or other familial risk factors, which modify *BRCA1* and *BRCA2* cancer susceptibility. Segregation analyses have produced significant evidence for a modifying effect of other genes on the risk of breast cancer in *BRCA1* and *BRCA2* mutation carriers, explaining the reported differences between population based estimates for *BRCA1*- and *BRCA2*-penetrance and estimates based on high-risk families.[71] For example a C/G polymorphism in the 5' untranslated region of *RAD51* was found to modify both breast and ovarian cancer risk in carriers of a germline *BRCA2* mutation (OR, 3.2; 95% CL, 1.4–40; $P = 0.01$).[77,78] A length-variation of the polyglutamine repeats in the estrogen receptor co-activator *NCO3A* influences breast cancer risk in carriers of *BRCA1* and *BRCA2* (OR, 1.96; 95% CI, 1.25–3.08; P for trend = 0.0036).[79,80] The androgen receptor also has a length-polymorphism, which inversely correlated with the transactivation function of the AR and has been shown to influence age at onset in carriers of BRCA1 in one study[79], but not in others.[81,82] Other unconfirmed modifiers of risk include rare alleles at the *HRAS1* repeat, modifying ovarian cancer risk in *BRCA1* carriers[83], and the variant progesteron receptor allele named *PROGINS*, modifying ovarian cancer risk in *BRCA1/2* carriers with no past exposure to oral contraceptives.[84] Thus, women with the same mutation may differ in their risk profiles, depending on their genetic background. The family history remains therefore an important parameter in translating standard risk estimates to individual patients.

For both *BRCA1* and *BRCA2* it has been shown that cancer risks are influenced by the position of the mutation within the gene sequence.[85,86] Women with a mutation in the central region of the *BRCA1* gene were shown to have a lower breast cancer risk than women with mutations outside this region. The ovarian cancer risk associated with mutations upstream this central region was higher than that associated with

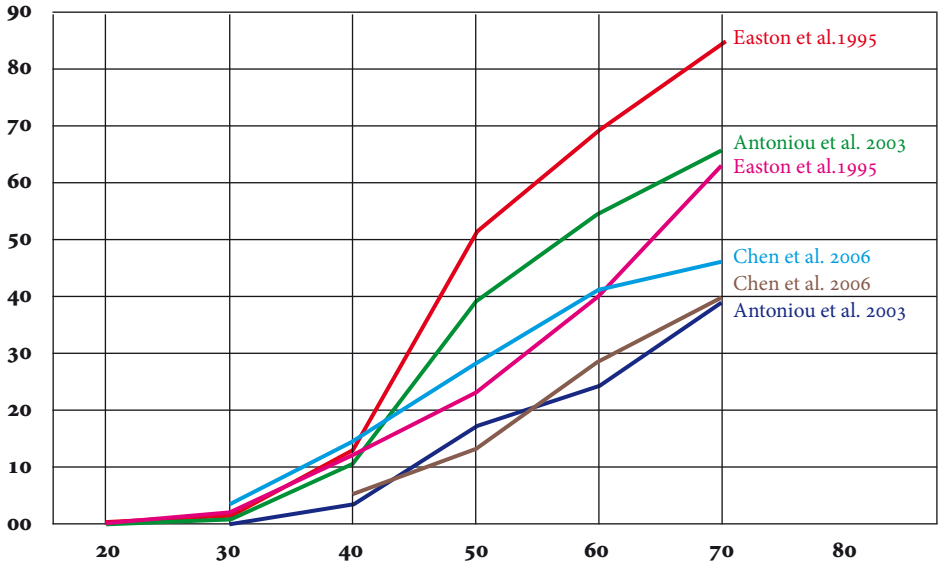


Fig. 1. Cumulative breast and ovarian cancer risk in BRCA1-mutation carriers as a function of age.
The red and pink line respectively represent family-based breast and ovarian cancer risk estimates (Easton et al.[274]). The green / light blue and dark blue / brown lines respectively represent population-based breast and ovarian cancer risk estimates (Antoniou et al.[74] (green/dark blue-line); Chen et al.[76] (light blue/brown-line)).

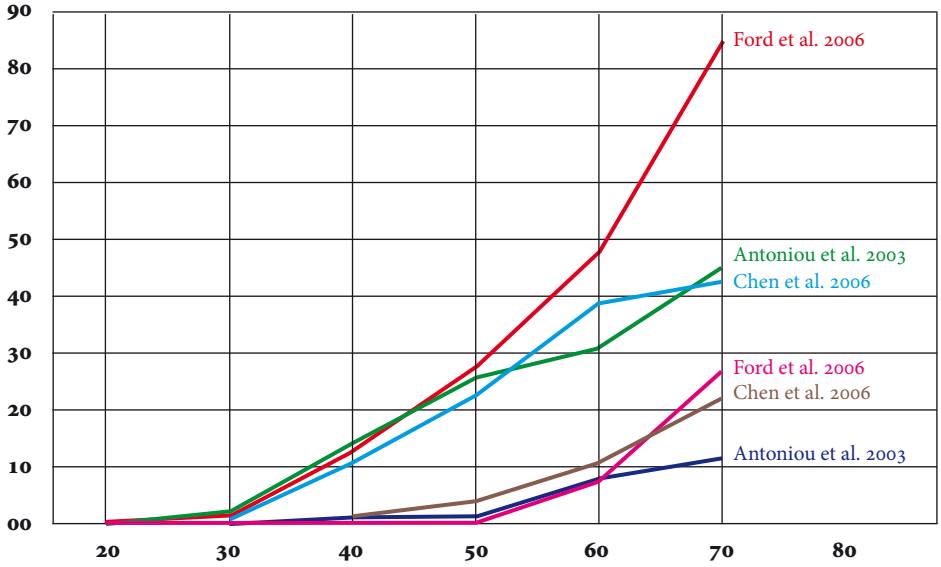


Fig. 2. Cumulative breast and ovarian cancer risk in BRCA2-mutation carriers as a function of age.
The red and pink line respectively represent family based breast and ovarian cancer risk estimates (Ford et al.[75]. The green / light blue and dark blue / brown lines respectively represent population-based breast and ovarian cancer risk estimates (Antoniou et al.[74] (green/dark blue-line); Chen et al.[76] (light blue/brown-line)). X-axis: age.

TABLE 4

Relative cancer risk (RR) for sites other than breast and ovary in BRCA1 and BRCA2 mutation carriers.

BRCA1 Location	RR	95% CI	study	BRCA2 Location	RR	95% CI	Study
Colon	4.11	2.36-7.15	1	Pharynx	7.3	2.0 to 18.6	3
Cervix	3.72	2.26-6.10	2	Pancreas	5.9	3.2 to 10.0	3
uterus	2.65	1.69-4.16	2	Pancreas	3.51	1. 87-6.58	4
pancreas	2.26	1.26-4.06	2	Bones	14.4	2.9 to 42.1	3
prostate	3.33	1.78-6.20	1	Prostate	2.5	1.6 to 3.8	3
prostate	1.82	1.01-3.29	2	Prostate	4.65	3.48-6.22	4
				Melanoma	0.1	0.01-0.2	3
				Melanoma	2.58	1.28-5.17	4
				Gastric	1.2	0.6-2.0	3
				Gastric	2.59	1.46-4.61	4
				Gall bladder	-	-	3
				Gall bladder	4.97	1. 50-16.52	4

1: Ford et al.[301] 2: Thompson et al.[90] 3: van Asperen et al.[87] 4: The Breast Cancer Linkage Consortium.[89]

mutations downstream this region. For *BRCA2*, mutations in the central region (OCCR; ovarian cancer cluster region) were associated with a higher risk of ovarian cancer than mutations outside this region, whereas mutations in the OCCR were associated with a lower breast cancer risk than mutations outside the OCCR.

In addition to a predominantly high increased risk to female breast cancer and ovarian cancer, *BRCA1*- or *BRCA2*-mutation carriers are at increased risk to 'other cancers' as well. An increased relative risk to colon cancer, cervix cancer, uterus, pancreas and prostate has been suggested in *BRCA1*-mutation carriers. In *BRCA2*-mutation carriers an increased relative risk to male breast cancer, gall bladder and bile ducts cancer, gastric cancer, malignant melanoma, pancreas, prostate, bone and pharynx cancer has been observed (Table 4).[75,87-90]

3.1.2. TP53 (Li-Fraumeni Syndrome)

The *TP53* gene is located on chromosome 17p13.1, and encodes a protein involved in many overlapping cellular pathways that control cell proliferation and homeostasis,

such as cell cycle, apoptosis and DNA-repair. The expression of the *TP53* gene is activated in response to various stress signals, including DNA damage. Loss of *TP53* function is thought to suppress a mechanism of protection against accumulating of genetic alterations (tumor suppressor).[91] Germline mutations in *TP53* are very rare: fewer than 400 families with germline mutations have been reported worldwide. Li-Fraumeni syndrome (LFS)(MIM: 151623)[92] is characterized by multiple primary neoplasms in children and young adults, with a predominance of soft tissue sarcomas, osteosarcomas, breast cancer and an increased incidence of brain tumors, leukaemia and adrenocortical carcinomas. Multiple primary tumors are frequently seen in Li-Fraumeni patients. The rarity and high mortality of the Li-Fraumeni syndrome precluded formal linkage analysis. The alternative approach was to select the most plausible candidate gene. Because tumor suppressor genes had been found to be associated with familial neoplasms, the *TP53* gene was a good candidate gene for LFS, because inactivating mutations therein had been associated with sporadic osteosarcomas, soft tissue sarcomas, brain tumors, leukemia's, and carcinomas of the lung and breast. Furthermore, transgenic mice carrying a mutant *TP53* gene have an increased incidence of osteosarcomas, soft tissue sarcomas, adenocarcinomas of the lung, and adrenal and lymphoid tumors, all tumors that occur as part of LFS.[92] Mutations in the *TP53* gene account for roughly 70% of families fulfilling the classical criteria for Li-Fraumeni syndrome (e.g. one patient with a sarcoma diagnosed <45 years with a first degree relative with any cancer diagnosed <45 years and an additional 1st or 2nd degree relative diagnosed with cancer <45 years or a sarcoma at any age).[93-96] Mutations in *TP53* are less common in breast cancer / sarcoma families not fulfilling these classical criteria.[96] Susceptibility to cancer in Li-Fraumeni families follows an autosomal dominant pattern of inheritance[97] and among families with a known germline *TP53* mutation the probability of developing any invasive cancer (excluding carcinomas of the skin) approaches 50% by the age of 30, compared to an age adjusted population incidence of cancer of 1%. It is estimated that more than 90% of *TP53* mutation carriers will develop cancer by the age of 70.[92] One of the most frequently occurring cancers in Li-Fraumeni families is breast cancer with an estimated penetrance in *TP53* mutation carriers of 28%-56% by the age of 45 years.[96,98,99] The peak incidence for breast cancer is between 20 and 40 years, in contrast to the other frequent occurring neoplasms, which mainly develop in young children, suggesting that hormonal stimulation of the mammary glands in puberty is an important cofactor.

Somatic mutations in *TP53* are reported in 20-60% of human breast cancers.[58] A

strong association was observed between *TP53* mutation and LOH at the *TP53* locus, in agreement with its tumor suppressor function.[100] Hypermethylation of the *TP53* gene seems not to play a major role in breast cancer.[101]
Germline mutations in *TP53* are rarely detected in families selected solely on the occurrence of breast and/or ovarian cancer,[102] and are found at very low prevalence (<0.5%) among early-onset cases of breast cancer.[58,103]

3.1.3. PTEN (The Cowden syndrome)

Cowden Syndrome (CS) (MIM: 158350) is an uncommon autosomal dominant disorder characterized by multiple hamartomas of the skin, breast, thyroid, gastrointestinal tract, central nervous system, and a high risk of breast, uterine and non-medullary thyroid cancer. Multiple trichilemmomas, papillomatosis, acral keratosis and benign tumors of the hair follicle are the most characterized neoplasms of the skin. Other features associated with CS are macrocephaly and gangliocytoma of the cerebellum (Lhermitte-Duclos disease).

A linkage genome scan was performed to localize the gene for CS.[104] The authors examined a total of 12 families, and obtained a maximum lod score of 8.92 at theta = 0.02 with the marker D10S573 located on 10q22-q23. They stated that the neurologic and neoplastic features of CS are consistent with the possibility that the Cowden gene is a tumor suppressor gene. The chromosomal region containing the CS gene was known to contain a tumor suppressor gene (*PTEN*) that had been found to be mutated in sporadic brain, breast, and prostate cancer and consequently germline mutations in the *PTEN* gene in 4 of 5 families with Cowden syndrome were found.[105] The prevalence of CS is estimated to be 1: 300 000. Mutations in the *PTEN* gene are present in about 80% of CS families.[105-107,107,108] Especially truncating *PTEN* mutations in CS families are associated with cancer.[109] Women carrying a *PTEN*-mutation have a 25-50% (2-4 fold) lifetime breast cancer risk. The majority of Cowden syndrome related breast cancers occur after the age of 30-35 years.[110,111] Also, breast cancer at young age has been observed in male carriers of a germline *PTEN* mutation with the classical CS phenotype, suggesting an increased risk for males as well.[112] However, no mutations in the *PTEN* gene have been detected in breast cancer families without features of CS.[113,114] Also in sporadic breast cancer patients, germline and somatic mutations in the *PTEN* gene are rare.[115,116] In addition, although LOH at the *PTEN* locus is found in 11-41% of sporadic breast cancers, no somatic mutations have been observed in the remaining allele.[117]

3.1.4. LKB1/STK11 (Peutz-Jegher Syndrome)

The *LKB1/STK11* –gene is located on chromosome 19p13.3, contains 12 exons and encodes a transcript of ~1.3 kb, which acts as a tumor suppressor. Germline mutations in the serine/threonine kinase gene (*LKB1/STK11*) causes Peutz-Jeghers syndrome (PJS) (MIM: 175200). To localize the susceptibility locus for Peutz-Jeghers syndrome, comparative genomic hybridization (CGH) and targeted linkage analysis, combined with loss of heterozygosity (LOH) study were used.[118] They demonstrated a high-penetrance locus in distal 19p with a multipoint lod score of 7.00 at marker D19S886 without evidence of genetic heterogeneity. The study demonstrated the power of CGH combined with LOH analysis in identifying putative tumor suppressor loci. In comparative genomic hybridization, a single hybridization allows DNA copy number changes in the whole genome of a tumor to be assessed in comparison with normal tissue DNA.[119] Within a distance of 190 kb proximal to D19S886, the marker with the highest lod score in the study of Hemminki et al.,[118] a novel human gene encoding the serine/threonine kinase *STK11* was identified and characterized.[120] In a three-generation PJS family, they found an *STK11* allele with a deletion of exons 4 and 5 and an inversion of exons 6 and 7 segregating with the disease. They concluded that germline mutations in *STK11*, probably in conjunction with acquired genetic defects of the second allele in somatic cells, caused the manifestations of PJS.

There is still much controversy on the exact prevalence of PJS. The estimates range from 1:8,900 to 1:280,000 (*The Johns Hopkins guide for patients and families: Peutz-Jeghers syndrome*, copyright 2001; http://www.hopkins-i.org/multimedia/database/hccIntro_111_PJS-Book.pdf). Not in all patients a germline mutation in *LKB1/STK11* is found, suggesting a heterogeneous basis for the disease. PJS is an autosomal dominant disorder characterized by a specific form of hamartomatous polyps (polyps with a muscular core) of the gastrointestinal tract and by melanine pigmentation of the lips, perioral region, the buccal mucosa, fingers, and toes. The polyps are most commonly seen in the small bowel but can occur throughout the gastrointestinal tract and at other sites such as the kidney, ureter, gall bladder, bronchus and nasal passage.[121,122] An elevated risk of gastrointestinal malignancies, breast cancer, pancreas, ovary, uterus, cervix, lung and testicular cancers is recognized in patients with PJS.[123-125] The clinical features of PJS vary within and between families, especially with respect to cancer risk. Overall, the probability of developing cancer by age 65 is estimated to be about 50%. The risk of breast cancer by age 65 ranges between 29% and 54%.[126,127] It's suggested that *LKB1/STK11* can play the role of a tumor suppressor gene

in sporadic breast cancer, and low expression of the *LKB1/STK11* protein is significantly associated with a shorter survival.[128] However in 62 primary breast cancers in patients without PJS, no somatic mutations were found in *LKB1* gene and LOH on 19p13 was observed in only 8%,[129] suggesting only a role in breast cancer susceptibility in patients with PJS.

3.1.5. CDH1/E-Cadherin (HDGC-syndrome)

The *E-cadherin* gene (*CDH1*) is located on chromosome 16q22.1 and contains 14 exons. The mature protein product belongs to the family of cell-cell adhesion molecules and plays a fundamental role in the maintenance of cell differentiation and the normal architecture of epithelial tissues. Genetic linkage analysis in affected members of three New Zealand Maori families with early-onset, histologically poorly differentiated, high-grade, diffuse gastric cancer demonstrated significant linkage to markers flanking the gene for the calcium-dependent cell-adhesion protein E-cadherin (*CDH1*). Sequencing of the *E-cadherin* gene revealed a G>T nucleotide substitution in the donor splice consensus sequence of exon 7, leading to a truncated gene product.[130] Thus, germline *CDH1* truncating mutations are associated with hereditary diffuse gastric cancer syndrome (HDGC-syndrome) (MIM: 192090).

The pattern of inheritance of the disease is consistent with an autosomal dominant susceptibility with incomplete penetrance. In HDGC families, women carrying a *CDH1* mutation have an estimated cumulative risk of diffuse gastric cancer by 80 years of 83%. The lifetime risk of developing breast cancer was estimated at 20-40%.[131-134] Somatic *CDH1* mutations are frequently found in infiltrating lobular breast cancer and in-situ lobular breast cancer (LCIS) in contrast to breast cancers of other histopathological subtype.[132,135,136] Germline mutations in *CDH1* are often found in combination with loss of heterozygosity of the wildtype *E-Cadherin* locus in the tumor, underscoring its role as a tumor suppressor.[132] Today most breast tumors reported in HDGC families are of the lobular subtype. One family with a germline *CDH1* mutation was described as a 'lobular breast cancer family'.[137] Therefore, it has been suggested that *CDH1* mutation screening should be offered to isolated cases of diffuse gastric cancer (DGC) in individuals ages <35 years and for families with multiple cases of lobular breast cancer, with any history of DGC or unspecified gastrointestinal malignancies.[137,138] However, others have failed to detect *CDH1* germline mutations in breast cancer families.[139,140]

3.2. Known low to moderate-risk breast cancer susceptibility genes

3.2.1. ATM

The *ATM* gene is located on chromosome 11q22-23 and contains 63 exons. The ATM protein plays a central role in sensing and signalling the presence of DNA double-strand breaks. In the unirradiated cell nucleus, *ATM* is held inactive, which is dissociated by rapid intermolecular autophosphorylation after irradiation.[141] This initiates cellular ATM kinase activity, which has many substrates including the protein products of *TP53, BRCA1* and *CHEK2*. Carriers of homozygous or compound heterozygous mutations in the *ATM* gene suffer from the rare recessive disorder ataxia-telangiectasia (AT) (MIM: 208900). AT is characterized by cerebellar degeneration (ataxia), dilated blood vessels in the eyes and skin (telangiectasia), immunodeficiency, chromosomal instability, increased sensitivity to ionising radiation and a highly increased susceptibility to cancer, in particular leukaemia's and lymphomas. The estimated incidence of AT is 1:40,000 to 1:100,000 with a carrier frequency of 1:100 to 1:200. Studies based on relatives of AT patients have suggested that female heterozygous carriers are at increased risk of breast cancer.[142-144] The estimated relative risk of breast cancer in obligate AT-heterozygotes range between 1.3 and 13 in the different studies conducted.[145] More recent estimates are in the order of 2.3,[146,147] with relatively narrow 95% confidence intervals. To date there is much controversy about the exact role of germline *ATM* mutations in breast cancer risk. Studies of sporadic and familial breast cancer have failed to consistently demonstrate an elevated prevalence of germline *ATM* gene variants among breast cancer cases relative to controls.[148,149] Initial reports of substantial increased risks of breast cancer (comparable with mutations in *BRCA1* and *BRCA2*) with specific variants in *ATM* (for example IVS10-6T>G)[150,151] have not been replicated in subsequent studies.[152,153]

It was hypothesized that the existence of two distinct classes of *ATM* mutations (truncating and missense) might explain some of the contradictory data on cancer risk. Some missense mutations encode stable, but functionally abnormal proteins that could compete in complex formation with the normal ATM protein, resulting in a dominant-negative cellular phenotype. In contrast, truncating mutations produce an unstable ATM protein so that heterozygote individuals still maintain 50% of wildtype ATM activity, resulting in an almost normal phenotype.[154,155] However, an analysis of 20 missense *ATM* mutations provided little support for an association of *ATM* missense mutation and breast cancer.[156] Thompson et al.[146] also found no evidence for a difference in risk of breast or other cancer according to the type of *ATM* mutation, while the risk estimate of Renwick et al.[147] was based mainly on truncating

mutations. Haplotype analysis could also reveal a role for common variants in the *ATM* gene in causing breast cancer. Five biallelic haplotype tagging single nucleotide polymorphisms (SNP's) have been estimated to capture 99% of the haplotype diversity in Caucasian populations. In the Nurses Health Study, there was no evidence that common haplotypes of *ATM* are associated with breast cancer risk.[157] When confirmed, this could suggest that less common variation in *ATM* is involved in increasing breast cancer risk, which can only be addressed in much larger studies. A possible example of such a variant is the c.7271T>G (V2424G), with an allele frequency of approximately 0.2% among cases and a substantially elevated breast cancer risk.[151,152,158] In conclusion, a role for the *ATM* gene in breast cancer susceptibility is plausible but the exact association remains unclear, and most probably comprises only a modest role in familial breast cancer susceptibility.

3.2.2. TGFβ1

The *TGFβ1*-gene is located on chromosome 19q13.1 and contains 7 exons and very large introns. TGFβ is a multifunctional peptide that controls proliferation, differentiation, and other functions in many cell types. TGFβ acts synergistically with TGFA in inducing transformation. It also acts as a negative autocrine growth factor. Dysregulation of *TGFβ* activation and signalling may result in apoptosis. Many cells synthesize TGFβ and almost all of them have specific receptors for this peptide.

For most normal cell types, TGFβ acts as a potent inhibitor of proliferation and migration and promotes apoptosis, properties associated with tumor suppression.[159,160] However, in cells in which these suppressor functions of the *TGFβ* signalling pathway are overridden, TGFβ may induce cellular changes associated with malignant progression,[161] invasion,[162] and angiogenesis.[163,164] These studies support a model in which TGFβ inhibits the development of early, benign lesions but promotes invasion and metastasis when the tumor suppressor activity is overridden by oncogenic mutations in other pathways.[165]

To date, several somatic mutations that disrupt the *TGFβ*-signalling pathway have been reported in human breast tumors.[166-168] On the basis of these data it was hypothesized that polymorphisms affecting the function of genes in the *TGFβ*-signalling pathway might also play a significant role in the development of breast cancer and the incidence of breast cancer associated with various SNP's in the *TGFβ1* gene was examined. A large combined case control study (3987 patients and 3867 controls) showed that the promotor SNP, *C-509T*, and the *T+29C* signal-peptide SNP (encoding Leu10Pro) are in very strong linkage disequilibrium and are both signifi-

cantly associated with increased incidence of invasive breast cancer in a recessive manner (respectively OR (TT versus C-carrier) =1.25, 95% confidence interval (CI) 1.06-1.48, P = 0.009 and OR (ProPro versus Leu-carrier) = 1.21, 95% CI 1.05-1.37, P = 0.01). Whereas the Leu10Pro signal peptide substitution potentially affects TGFβ1 secretion in contrast to the *C-509T* SNP it was suggested that the observed association was caused by the Leu10Pro SNP.[169]

3.2.3. CASP8

The *CASP8* gene is located on chromosome 2q33-q34, contains 13 exons and the protein product spans 51,2 kb. Caspases are important mediators of the apoptotic process. Death receptor-mediated apoptosis provokes the formation of the death-inducing signalling complex (DISC), comprising the death receptors, adaptor proteins as well as the initiator caspase 10 (CASP10) and caspase 8 (CASP8). It has been shown that a germ-line homozygous missense mutation (R248W) in *CASP8* causes the autosomal recessive autoimmune lymphoproliferative syndrome type IIB (MIM: 607271). This syndrome is characterized by lymphadenopathy and splenomegaly associated with an immunodeficiency. The immunodeficiency is characterized by recurrent sinopulmonary and herpes simplex virus infection with poor response to immunization due to defects in activation of T-lymfocytes, B-lymfocytes and natural killer cells.[170]

Because of the involvement in initiation of apoptosis, it was hypothesized that *CASP8* and *CASP10* might act as low-penetrance familial breast cancer susceptibility genes. Surprisingly, combined analysis of two different studies showed that one missense variant (D302H) in *CASP8* was associated with a reduced risk of breast cancer in a dose-dependent manner. The combined odds ratios (OR) for breast cancer was 0.83 (95% confidence interval = 0.74 to 0.94) for the DH heterozygote and 0.58 (95% CI= 0.39 to 0.88) for the HH homozygote.[171] Recently the Breast Cancer Association Consortium (BCAC) confirmed these findings. They included data from 9-15 studies, comprising 11,391-18,290 cases and 14,753-22,670 controls and found evidence of an association with breast cancer for *CASP8* D302H (with odds ratios (OR) of 0.89 (95% CI = 0.84-0.92, $P_{trend} = 1.1 \times 10^{-7}$) and 0.74 (95% CI = 0.62-0.87, $P_{trend} = 1.1 \times 10^{-7}$) for heterozygotes and rare homozygotes respectively, compared with common homozygotes).[172]

The functional effect, if any, of the aspartate-to-histidine change at residue 302 in caspase-8 is as yet unknown. A different study showed that the *CASP10* V410I variant was also significantly associated with a decreased familial breast cancer risk (OR =

0.62, 95% CI = 0.43-0.88, *P* = 0.0076). In individuals carrying the protective alleles of both *CASP10* (I410) and *CASP8* (H302) the breast cancer risk was even more reduced (OR= 0.37, 95% CI =0.16-0.83, *P*=0.013).[173]

3.2.4. CHEK2

The *CHEK2* gene is located on chromosome 22q12.1 and contains 15 exons. Several pseudogenes, encompassing exons 10-14 of the gene, are scattered throughout the genome. CHEK2 is a G2 checkpoint kinase that plays an important role in DNA repair and it is activated in response to ionising radiation through phosphorylation by ATM. Activation of *CHEK2* also phosphorylates other key cell cycle proteins, including BRCA1 and p53. The role of *CHEK2* in breast cancer susceptibility was first suggested by the identification of the truncating mutation 1100delC, which eliminates kinase activity, in an individual with Li-Fraumeni syndrome without a *TP53* mutation. The possibility that this gene is only contributing to the breast cancer cases within LFS families rather than LFS per se has been raised.[174] The frequency of 1100delC has been estimated in healthy control populations, and was found to be approximately 1%.[175,176] Among unselected patients with breast cancer, its prevalence was found to be approximately 1.5- to 3-fold higher than in controls. Among breast cancer cases selected from families that were not linked to *BRCA1* and *BRCA2* prevalences between 4.9% and 11.4% were found depending on the total number of breast cancer cases in the families.[175-177] Segregation analysis estimated that *CHEK2*1100delC* conferred an increased risk of breast cancer of approximately 2-fold in noncarriers of *BRCA1/2* mutations.[178,179]

These results suggest that *CHEK2*1100delC* is not a high penetrance mutation, but rather a relatively common variant conferring a more moderate risk of breast cancer, which may make a significant contribution to familial clustering of breast cancer. As it is enriched among multiple-case families, but unable to explain all breast cancer in families with at least one carrier case, it has been suggested to interact with other, as yet unknown breast cancer susceptibility alleles.[177] Other variants in *CHEK2* have also been considered to be involved in causing breast cancer risk. Whereas some studies have excluded this possibility,[180,181] others have implicated slightly increased risks associated with 157^T and IVS2+1G > A.[182,183] The 157^T protein, which compromises cellular responses to ionising radiation and shows deficiency in substrate recognition in vivo, was expressed at normal levels in tumor tissues as well as in cultured cells. The 157^T protein was stable and it dimerized with the wild-type CHEK2 co-expressed in human cells. These functional properties of the 157^T protein suggest that

this variant may have negative effect on the pool of normal CHEK2 protein in heterozygous carrier cells by formation of heterodimers with wild-type CHEK2. The 157^{T} variant may be associated with breast cancer risk, but the risk is probably lower than for 1100delC.

Patients carrying the *CHEK2*1100delC* mutation developed breast cancer earlier than non carriers[177,184] and have a eightfold risk of developing contralateral breast cancer when compared with matched controls.[184,185] There is no specific histological subtype described for *CHEK2*-related breast tumors.[184,186] Immunohistochemically, *CHEK2* related breast tumors show in most cases an absent CHEK2 protein staining and are more often negative for luminal cytokeratin 19 staining compared to familial non-*BRCA1/2* and *BRCA1* related breast tumors.[177,187]

3.2.5. BARD1

The *BRCA1-associated ring domain 1* (*BARD1*) gene is located on chromosome 2q34-q35 and contains 11 exons. The BARD1protein was discovered in a yeast two-hybrid screen as a binding partner of BRCA1.[188] BRCA1 and BARD1 form a functional heterodimer through the binding of their RING-finger domains. This interaction is thought to stabilize both proteins, as the respective monomers are unstable.[189,190] BARD1 and BRCA1 have several features in common: similar protein structure, the embryonic lethality of their respective knockout mice, induction of genetic instability when depleted from cells, both proteins have a RING domain, a nuclear export signal at their N termini and two tandem BRCA1 corboxy-terminal (BRCT) domains. The BRCA1-BARD1 interaction is required for several of the cellular and tumor-suppressor functions of BRCA1. However, BARD1 has also been described in tumor suppressive functions independent of BRCA1, by mediating between genotoxic stress and p53-dependent apoptosis.[191] The *BARD1* gene has been reported to be targeted by somatic mutations in breast and ovarian cancers,[192] and has been considered a possible candidate to be involved in cancer susceptibility. In a screen of an Italian cohort of familial breast and ovarian cancers that were not associated with *BRCA1* and *BRCA2* gene mutations, five alterations in *BARD1* were discovered,[193] including 1139del21 and Cys557Ser.

Recently, a Nordic collaborative study of the *BARD1* Cys557Ser allele consisting of altogether 2906 breast and/or ovarian cases and 3591 controls from Finland, Iceland, Denmark, Sweden and Norway provided further evidence that *BARD1* Cys557Ser confers a slightly increased risk of female breast cancer. The frequency of the *BARD1* Cys557Ser variant appeared to be increased among patients from breast/ovarian

cancer families. Significant difference was obtained compared to controls (6.8% vs. 2.7%; P=0.000003; OR=2.6; 95% CI=1.7-4.0).[194] So, in conclusion there seems to be an association between specific *BARD1* mutations and breast and ovarian cancer, but this accounts for only a small fraction of cases of familial breast cancer overall.

3.2.6. The Fanconi Pathway other than FANCD2 (BRIP1 and PALB2)

Fanconi anemia (FA) is an inherited disorder associated with progressive aplastic anemia, multiple congenital abnormalities and predisposition to malignancies including leukemia and solid tumors.[195] The developmental abnormalities include radial aplasia, hyper pigmentation of the skin, growth retardation, microphthalmia and malformation of the kidneys. FA is inherited mainly as an autosomal recessive trait, but is genetically heterogeneous. Analysis of cell lines from different FA patients led to the discovery of at least 13 groups, named FA-A, B, C, D1, D2, E, F, G, I, J, L, M, and N with the corresponding genes named as *FANCA-FANCN*. The interest in the FA pathway by breast cancer researchers was stimulated by the discovery that the gene for *FANCD1* is *BRCA2*. As described before, mono-allelic mutations in *BRCA2* causes susceptibility to breast and other cancers, whereas bi-allelic mutations cause Fanconi anemia. The phenotype of biallelic *BRCA2* mutations differs from other Fanconi anemia subtypes, most notably with respect to the high risk of childhood solid tumors, particularly Wilms tumor and medulloblastoma, which occur very rarely in other Fanconi anemia subtypes.[66,70,196,197]

To date, there have been several studies of the other known FA-genes in relation to breast cancer susceptibility. No clear pathogenic mutations were detected in *FANCA, FANCC, FANCD2, FANCE, FANCF, FANCG* and *FANCL*.[68,198,199] However, in the gene that is variously known as *BACH1/BRIP1/FANCJ* (located at 17q22-24, containing 20 exons) two missense mutations in early onset familial breast cancer cases was found.[200] *BRIP1* encodes a DEAH helicase that interacts with the BRCT domain of *BRCA1* and has BRCA1-dependent DNA-repair and Checkpoint functions.[200,201] Inactivating mutations in *BRCA1* predispose to breast cancer. Inactivation of *BRIP1* results in abrogation of certain BRCA1 function, and therefore it is plausible that inactivating *BRIP1* mutations also predispose to breast cancer.[202]

Unfortunately several other studies from different populations cold not confirm this finding.[198,203-207] However, recently a truncating mutation in *BRIP1* was identified in 9 out of 1,212 individuals with breast cancer from *BRCA1/2* mutation-negative families but in only 2 out of 2.081 controls (P=0.003).[208] They estimated that *BRIP1* mutations

confer a modest relative risk of breast cancer of 2.0 (95% CI=1.2-3.2, P=0.012), similar to truncating variants of *CHEK2* and *ATM*.

The protein PALB2 (for 'partner and localizer of BRCA2') was recently identified as a nuclear partner of BRCA2. PALB2 co localizes with BRCA2, promoting its localization and stability in key nuclear structures, which in turn facilitates BRCA2 functions in DNA repair.[209] The gene encoding the PALB2 protein is located at 16p12 and contains 13 exons. Because of the existence of individuals with the BRCA2-Fanconi phenotype who lacked *BRCA2* mutations, the possibility of a role of *PALB2* (functionally related to *BRCA2*) in FA was raised. And consequently, pathogenic *PALB2* mutations were identified in families affected with FA and cancer in early childhood, demonstrating that bi-allelic *PALB2* mutations cause a new subtype of Fanconi anemia, FANCN.[210] Prompted by these observations, Rahman et al.[211] investigated whether monoallelic *PALB2* mutations confer susceptibility to breast cancer. They identified truncating *PALB2* mutations in 10 out of 923 (1.1%) individuals with familial breast cancer compared with 0 out of 1,084 (0%) controls (P=0.0004). When considering families with both male and female breast cancer *PALB2* mutations were found in 6.7%. Although numbers were low, it suggests that *PALB2* mutations may confer a high risk of male breast cancer, which is also a hallmark of *BRCA2*.

The authors estimated that *PALB2* mutations confer a modest relative risk of breast cancer of 2.3 (95% CI = 1.4-3.9, P = 0.0025).

4. Genetics of familial breast cancer

4.1. Attributable risks

How much of the familial risk is currently explained by the known genes? *BRCA1* and *BRCA2* appear to be the two major factors among families with multiple cases of early-onset breast cancer. Germline *BRCA1* mutations are found in 80% of families with at least 4 cases of breast cancer diagnosed before the age of 60 and at least one case of ovarian cancer.[75] This reflects the high risks conferred by *BRCA1* mutations to both breast and ovarian cancer (see section 3.1.1). Likewise, *BRCA2* mutations are strongly associated with families with a case of male breast cancer. Among families in which female breast cancer is the only major cancer phenotype, *BRCA1* and *BRCA2* mutations are less often encountered, unless the number of cases diagnosed under 60 is very high (i.e., six or more). These estimates derive from a highly selected group of families, selected to be sufficiently informative for linkage analyses, and are therefore subject to strong upward bias. Nonetheless, similar findings have been made on clinic-based families from a variety of different ethnic backgrounds.[212,213] On avera-

ge, *BRCA1* and *BRCA2* mutations are found in approximately 25% of the families who self-refer to a Cancer Family Clinic, with higher occurrences among families with cases of ovarian cancer or male breast cancer. Mutations in the other high risk cancer susceptibility genes *TP53* (Li-Fraumeni Syndrome), *PTEN* (Cowden syndrome), *CDH1* (HDGC-syndrome) and *LKB1* (Peutz-Jegher Syndrome) are also associated with breast cancer but germline mutations in these genes are very rare and are not found in patients with breast cancer in the absence of the other clinical stigmata of these cancer syndromes.[113,115,214] It is thus obvious that *BRCA1* and *BRCA2* are unable to explain all the observed familial clustering.

4.2. Segregation analyses

The observation of large extended kindred's with many cases of early-onset breast cancer is a strong indication that one or more highly penetrant autosomal dominant genes for breast cancer may exist. Many studies have used segregation analysis in large numbers of families with breast cancer to derive genetic models that could explain the observed familial aggregation. Many of these analyses found support for a model in which susceptibility to breast cancer was explained by a rare dominant disease allele conferring a high lifetime risk of the disease.[215-217] A widely used model in linkage analyses has been the model by Claus et al,[215] which specifies a dominant allele with a population frequency of 0.003 and a penetrance of 80% by age 70. The identification of *BRCA1* and *BRCA2* by linkage analysis in multiple case families in the 1990's confirmed the existence of such high penetrance alleles.[218,219]

Using data from both a population-based series of breast cancer cases and high risk families in the UK, with information on *BRCA1* and *BRCA2* mutation status, the genetic models that can best explain familial breast cancer outside *BRCA1* and *BRCA2* families were investigated.[71] The allele frequency of *BRCA1* was estimated to be approximately 0.05% and slightly higher estimates were derived for *BRCA2*. The best fitting model for the residual non-*BRCA1/2* familial aggregation of breast cancer was a polygenic model, although a model with a single recessive allele produced a similar fit.[71] A comparable study used three-generation families ascertained from women with breast cancer diagnosed at age <40 years, obtained from population cancer registries in Australia.[220] A residual dominantly inherited risk of female breast cancer, in addition to that derived from mutations in *BRCA1* and *BRCA2*, was suggested. However, this analysis also suggested that there is a substantial recessively inherited risk of early-onset breast cancer of 86% by age 50. Of note, when considering only the population-based cases, the UK-dataset also produced a recessive model as the

best-fitting single gene model for *BRCAX*, with a disease allele frequency of 24% and a penetrance of 42% by age 70.[221] However, a polygenic model gave a similarly good fit. The dominant model gave a somewhat worse fit although the difference was not significant. But when the known effects of parity on breast and ovarian cancer risk were included in the model, the polygenic model fits best.[221] These findings suggest that several common, low penetrance genes with multiplicative effects on risk may account for the residual non-*BRCA1/2* familial aggregation of breast cancer, although Mendelian inheritance of an autosomal dominant or recessive allele cannot be ruled out at this stage.

Due to the recent discovery of low to moderate breast cancer susceptibility genes, the question rises, how many of the observed familial clustering could be explained by combinations of these genes. Unfortunately, in medical journals few (if any) publications have appeared on this topic. However, it's clear that also the known low to moderate breast cancer susceptibility genes will not explain all the remaining familial clustering.

4.3. Linkage analyses

Family-based linkage studies have been very successful in mapping genes that underlie monogenic disorders, including common cancers. *BRCA1* was the first locus found to be linked to breast cancer in early onset multiple-case families.[222] After this, it was quickly established that linkage to *BRCA1* extended to families in which both breast and ovarian cancer were prevalent.[223,224] In contrast, families with multiple cases of female breast cancer and at least one case of male breast cancer were clearly not linked to *BRCA1*.[225] Linkage analysis of male breast cancer families then led to the discovery of *BRCA2* on 13q12.[219,226] However, attempts to localize further genes associated with an inherited predisposition to breast cancer have not been successful to date. The lack of a clear phenotype that could indicate the presence of another major breast cancer gene may be one of the reasons for this failure.

A number of linkage studies have analysed candidate regions, which were derived from the genetic analysis of breast tumors (Table 5). For example, the short arm of chromosome 8 is known to be frequently deleted in sporadic breast cancer,[227] and CGH analysis of familial cases highlighted the long arm of chromosome 13 to be lost in several cases belonging to a single family.[228] Although suggestive LOD scores were found in these studies,[229-232] none were greater than three (the commonly accepted level of statistical significance), and none were confirmed in studies of independent collections of families.[233-235]

TABLE 5

Summary of different published linkage studies

Study	number of families or cases	model	lod	alpha	position
1	11	dom	1.43		8q
2	1	dom	1.99		9q34
3	1	dom	1.85		6q
4	8	dom	2.51		8p12-22
5	4	dom	2.97		8p12-22
6	31	dom	0.03	0.03	8p12-22
7	77	dom	3.46	0.65	13q21
8	128	dom	-11.0		13q21
9*	14	dom	1.12		9q21
		npl	3.20		2q32
10*	150	dom	1.21	0.18	2 (17)
		npl	1.10		2 (16)
		dom	1.80	0.18	4 (79)
		rec	1.04		5 (169)
		npl	1.56		14 (44)
		dom	1.15	0.06	22 (41)
	4 cases <50	dom	2.38	0.5	2 (17)
		dom	1.57	0.28	4 (66)
		dom	1.12	0.35	10 (89)
		dom	1.43	0.12	22 (41)
11**	SNP analysis		P=0.00038		17
			P=0.0006		8p12
			P=0.000007		15
			P=0.000007		9

1: King et al.[302] 2: Skolnick et al.[303] 3: Zuppan et al.[304] 4: Kerangueven et al.[229] 5: Seitz et al.[232]
6: Rahman et al.[233] 7: Kainu et al.[228] 8: Thompson et al.[234] 9: Huusko et al.[236] 10: Smith et al.[235]
11: Ellis et al.[305]
*: genome-wide linkage study. **: genome-wide SNP-analysis.
dom: linkage analysis assuming a dominant model.
rec: linkage analysis assuming a recessive model.
npl: linkage analysis assuming a non-parametric model.
lod: lod-score under heterogeneity.

To date, only two genome-wide linkage scans have been reported in multiple-case non-*BRCA1/2* breast cancer families. Huusko et al.[236] studied 14 high-risk Finnish breast cancer families in which a role for *BRCA1* or *BRCA2* was excluded by mutation analysis (DGGE, SSCP or CSGE), protein truncation test and linkage analysis. All families had at least three breast cancer cases with DNA available for genotyping. The age of diagnosis and the occurrence of ovarian cancer were not used as exclusion criteria. Suggestive linkage was seen at marker D2S364 (2q32) with a parametric two-point LOD score of 1.61 (theta=0), and an LOD score of 2.49 in nonparametric analyses. This finding was not replicated in a much larger study of 149 non-*BRCA1/2* breast cancer families performed by the Breast Cancer Linkage Consortium (BCLC)[235] (see chapter 5.1). These families were selected for linkage analysis when they had at least 3 cases of breast cancer under 60, and no cases of ovarian or male breast cancer. The strongest linkage signal in this study was found on the short arm of chromosome 4 (LOD 1.80; α=0.18). When the analysis was restricted to families with at least four breast cancer cases diagnosed before age 50 a LOD-score of 2.38 was found on chromosome 2 (2p24-25). To provide some protection against model mis-specification, lod scores were also calculated under a recessive model and using an allele sharing approach (non-parametric linkage analysis). These approaches, however, identified no further strong linkage signals. This study represents by far the largest genome wide linkage screen for breast cancer susceptibility loci to date.

The failure to detect strong linkage signals might be explained in several ways. First, it might reflect extensive locus heterogeneity, in which multiple high-risk loci underlie the same disease phenotype. Accordingly, each locus explains only a small proportion of families, which severely limits the statistical power of the study.

Second, the genetic model used for linkage analysis may not have been the correct one. It is possible that many genes are involved, each conferring only a small risk of the disease. In that case, phenocopies and incomplete penetrance causes a problem, as the carrier status of a disease allele cannot be definitively inferred from disease status. Within each family, different combinations of genes could be involved in individual breast cancer susceptibility. Hence, if there are still moderate to high penetrance breast cancer genes to be detected, it is clear that each will explain only a small proportion of families. A possible way of addressing the genetic heterogeneity problem and the associated loss of statistical power might thus be to find variables that allow the sub-classification of families into more homogeneous groups. This could possibly be achieved by a better definition of the tumor characteristics in the multiple case families not due to *BRCA1* or *BRCA2*.

5. Tumor characteristics

5.1. Pathology

It is now well established that breast tumors arising in women carrying a *BRCA1* mutation have distinct histopathological features. Histopathologically the *BRCA1* related tumors are generally of higher grade, showing pushing margin growth patterns and a high proportion of lymphocytic infiltration compared with sporadic breast cancer and familial non-*BRCA1/2* breast cancer.[237-239] Interestingly, breast tumors associated with *BRCA1* hypermethylation are histopathologically similar to those that are caused by germline mutations in *BRCA1*, in that they are high grade, infiltrating ductal breast cancers that do not express ER.[240,241] Other studies have suggested that *BRCA1* tumors are larger and more often associated with axillary lymph node involvement,[242-244] although the evidence for these associations is less convincing than for grade. The majority of *BRCA1*-associated tumors are infiltrating ductal, but there is a significantly higher frequency of tumors classified as medullary or atypical medullary type than in noncarriers (21% vs 2%). Ductal carcinoma in situ (DCIS) adjacent to invasive cancer is observed less frequently while the frequency of lobular neoplasia in situ (LCIS) is similar when compared to controls.[245]

No specific histological type is thought to be associated with *BRCA2*. The only factors found to be significant for *BRCA2* were tubule score, fewer mitoses and continuous pushing margins.[237] The lobular type is associated with mutations in the *E-cadherin* gene (*CDH1*).

5.2. Loss of heterozygosity

Loss of heterozygosity (LOH), the loss of a normal, functional allele at a heterozygous locus, is the most common type of somatic alteration found in primary human breast tumors.[246] Consistent LOH in a genomic region implicates the presence of tumor-suppressor genes or other genes related to tumor pathogenesis.[247-249] In germline *BRCA1/2* mutation carriers complete loss of the wildtype allele (LOH) is a common mechanism of inactivation,[250] which is consistent with Knudson's two-hit theory for tumor-suppressor genes. *BRCA1* related tumors also show frequent LOH at 4q and 5q, and those from families linked to *BRCA2* on 6q.[251,252]

Despite the hundreds of LOH studies of sporadic breast cancer, the number and identity of tumor-suppressor genes relevant to this disease remain largely unknown.[253] It was concluded that finding tumor-suppressor genes might require 'brute force' approaches, presumably involving analysis of many tumors. One such approach is represented by a pooled analysis of 151 published LOH studies of breast cancer (>15.000

tumors). They observed a preferential loss in specific regions of chromosomes 7q, 16q, 13q, 17p, 8p, 21q, 3p, 18q, 2q, and 19p, in descending order of significance. Interestingly, genes causing inherited rare syndromic breast cancer susceptibility were not in regions of substantially elevated loss.[227] In a study, described in chapter 4.1, comprising 100 familial non-*BRCA1/2* related breast tumors LOH frequencies of 40% or greater were found at 1q41, 4p16, 11q23.3, 16p13, 16q24, 17p12, 21q22, 22q11 and 22q13, with the highest frequency at 22q13.[187] Except for 22q, many of these chromosomal sites have also been highlighted in analyses of sporadic breast tumors. The same study identified loci (on chromosome 2, 3, 6, 12, 13, 21 and 22) at which LOH was found significantly more often within families than expected on the basis of overall LOH frequency at that given locus in all families. Unfortunately, in an attempt to address the problem of genetic heterogeneity, selecting families based on these LOH-findings did not increase LOD-scores on the loci identified by LOH. However, it remains possible that families in which multiple breast tumors show LOH at the same locus are caused by a shared genetic defect on another chromosome.

5.3. Comparative genome hybridisation (CGH)

Current approaches for detecting LOH can be sensitive to other sources of allelic imbalance, for example amplification.[253] To distinguish between these, LOH-data should be combined with (array-) CGH. This might be relevant because we do not know at this stage whether other breast cancer susceptibility genes act according to Knudson's two-hit inactivation model. It is conceivable, as was found for the *MET* oncogene in hereditary papillary renal carcinomas, that trisomy (or copy-number gain) of the mutant allele contributes to susceptibility.[254]

With metaphase comparative genomic hybridisation (CGH) analysis a distinct *BRCA1* classifier could be determined. Specific somatic genetic aberrations on chromosome 3p (losses), 3q (gain) and 5q (losses) could distinguish *BRCA1* related tumors from control tumors with a sensitivity of 96% and a specificity of 76%.[255] However, metaphase CGH analysis could not reliably distinguish between *BRCA2*-associated breast tumors and control tumors or *BRCA1*-associated breast tumors.[256] Based on array-CGH analysis (which has a higher resolution than metaphase CGH) 169 significant BAC clones were identified which enabled discrimination between *BRCA1*, *BRCA2* and sporadic tumors to some degree. Using hierarchical clustering methods, *BRCA1*-associated tumors were tightly clustered and separated from sporadic cases, whereas *BRCA2*-tumors showed a somewhat higher similarity with the sporadic cases, although they still displayed a genomic profile of their own (30% of *BRCA2*-tumors

clustered within the control or *BRCA1*-group).[257] All studies showed that *BRCA1*-associated tumors have the highest frequency of copy number alterations. In familial non-*BRCA1/2* associated tumors a significant higher incidence of 8q-gains, 19p-gains, 19q-gains and 8p-losses was observed with metaphase CGH compared to sporadic tumors.[258]

5.4. Immunophenotype, global gene expression

Many studies have shown that *BRCA1* tumors are immunohistologically more often negative for ER, PR and Her2Neu expression, *TP53* mutated and positive for cytokeratin 5/6 compared with sporadic tumors and familial non-*BRCA1/2* tumors (*BRCAX*).[187,259,260] When compared with *BRCA2* tumors these differences are also observed for ER, Her2Neu and Cytokeratine 5/6.[187] *BRCAX*-related tumors are significantly more often positive for BCL2 compared with *BRCA1*- and *BRCA2*-related tumors.[187,260] (see chapter 4.1). Figure 3 Is an illustration of a typical *BRCA1*-related immunohistochemical staining pattern.

Table 6, provides a list of significant immunohistochemical differences, found in 4 different studies, between *BRCA1*, *BRCA2*, non-*BRCA1/2* familial tumors and tumors unselected for family history.

It appears to be difficult to distinguish *BRCAX* tumors from sporadic tumors and *BRCA2* tumors. Differences found in one study were not confirmed by the other studies. This is partly explained by different selection criteria for the *BRCAX* group, the use of slightly different antibodies or the number of different antibodies used. However, it can also reflect the extensive heterogeneity in the *BRCAX* group.

Gene expression profiling of sporadic cases enabled discrimination of five different tumor subtypes; one basal-like, one *ERBB2*-overexpressing, two luminal-like and one normal breast-tissue-like subgroup. These tumor subtypes may represent different biological entities and might originate from different cell types. A basal-like gene expression pattern has been associated with *BRCA1* carriers.[261] In addition, the study by Hedenfalk et al.[262] showed that the expression patterns from 15 fresh frozen tumors from seven non-*BRCA1/2* families clustered within their respective families. They even showed that the *BRCAX* subgroups were not only separated from one another but also from the *BRCA1* and *BRCA2* tumors. To date, this is the only study in which clustering of non-*BRCA1/2* breast tumors was seen. It would be interesting to see if this observation could be confirmed and extended to larger number of cases. These findings could indicate that genetic predisposition to breast cancer might

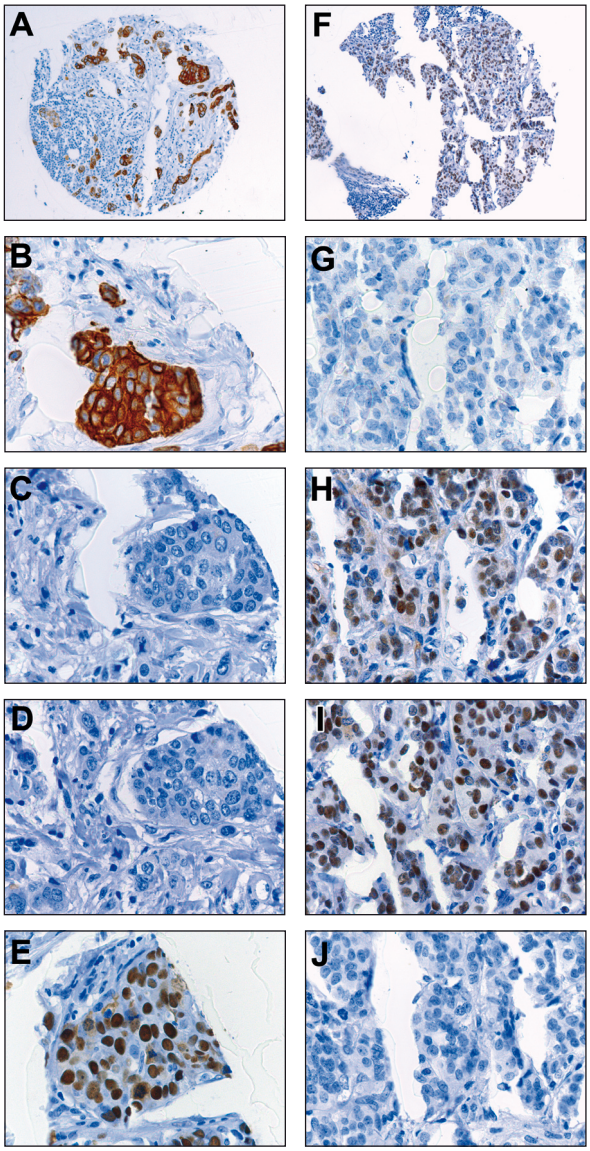


Fig. 3. Immunohistochemical staining results of a BRCA1- and BRCA2- related breast tumor on a tissue microarray. The samples A, B, C, D and E are from one *BRCA1*- (2315del5) tumor and the samples F, G, H, I and J of one *BRCA2*- (6648insA) tumor, both on the same paraffin tissue microarray block. A and F provide an overview of the analyzed biopsy cores. B: a typical strong positive cytokeratin 5/6 staining pattern, C: a typical absent estrogen receptor (ER) protein expression, D: a typical absent progesterone receptor (PR) protein expression and E: a typical strong TP53 protein expression as is in generally seen in *BRCA1*-related tumors. G: an absent cytokeratin 5/6 protein expression, H: a strong ER protein expression, I: a strong PR protein expression and J: an absent TP53 protein expression. Magnification x10 in A and F and x40 in B, C, D, E, G,

TABLE 6

Summary of immunohistochemical staining pattern differences found when comparing BRCA1-breast tumors with BRCA2- and familial non-BRCA2/1 (BX) breast tumors, and BX-tumors with BRCA2- and sporadic breast tumors

Antibody	staining	Summery of published immunohistochemical differences between tumor groups.							
		B1 vs BX	**ref**	**B1 vs B2**	**ref**	**B2 vs BX**	**ref**	**BX vs Sp**	**ref**
ER	absent	S	2,3,4	S	3	S NS	4 3		
PR	absent	S	2,3,4					S NS	4 2
Her2Neu	absent	S NS	3 2,4	S	3	S NS	2 3,4	S NS	2 4
Bcl2	absent	S	2,3			S	3		
P53	strong	S	2,4					S NS	
P-CD	strong	S	2						
Cycline D1	absent	S	3			S	3		
Cytokeratin 5/6	strong	S	3	S	3				
Ki-67	absent	S	2			S	2,3	S	
Chek2	strong	S NS	1 3			S NS	1 3		

References; 1: Honrado et al.[306] (Comparison between 74 BRCA1-tumors, 71 BRCA2-tumors, 108 non-BRCA1/BRCA2-tumors and 288 sporadic tumors). 2: Palacios et al.[260] (Comparison between 20 BRCA1, 18 BRCA2, 37 non-BRCA1/BRCA2-tumors). 3: Oldenburg et al.[187] (Comparison between 31 BRCA1, 21 BRCA2, 100 non-BRCA1/BRCA2-tumors). 4: Eerola et al.[307] (Comparison between 51 BRCA1, 59 BRCA2, 152 non-BRCA1/BRCA2-tumors and 862 sporadic tumors). Absent: absent protein expression.
Strong: strong protein expression. S: significant (significant difference found between indicated tumor groups). NS: nonsignificant

preferentially give rise to distinct subtypes (as is seen for *BRCA1* related breast tumors) and that the separation of the heterogeneous group of *BRCAX* breast cancers into more homogeneous subgroups may be possible. If so, incorporating tumor characteristics into genome-wide linkage analysis could identify linkage signals that are not evident using breast cancer as a whole as the disease endpoint.

In one such an attempt to find distinct subgroups (using LOH and immunohistochemistry), cases from families with a high probability of segregating a breast cancer susceptibility gene but with a minimal residual probability that this is due to *BRCA1* or *BRCA2*, were selected. Unfortunately, cluster analysis of the separate and combined data did not result in subgroups that would allow useful subclassification of the families for further linkage analysis. In addition, when using the phenotype categories described by Abd-El-Rehim[263] it was noted that different tumors within the same family frequently belonged to different phenotype categories, indicating that it is unlikely that the basal/luminal phenotype has a strong genetic basis in these cases[187] (see chapter 4.1). However, the possibility that array CGH and expression profiling could define distinct subgroups of familial breast cancer still deserves further exploration (see chapter 4.2).

CHAPTER 3.

PUTATIVE CANDIDATE GENES

3.1. EXTENDING THE P16-LEIDEN TUMOUR SPECTRUM BY RESPIRATORY TRACT TUMOURS

R.A. Oldenburg[1,2], W.H. de Vos tot Nederveen Cappel[3], M. van Puijenbroek[4], A. van den Ouweland[2], E. Bakker[1], G. Griffioen[3], P. Devilee[1], C.J. Cornelisse[4], H. Meijers-Heijboer[2], H.F.A. Vasen[5] and H. Morreau[4].

J Med Genet. Mar;41(3):e31. (2004)

[1] *Center of Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands*
[2] *Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands*
[3] *Department of Gastroenterology, Leiden University Medical Center, Leiden, The Netherlands*
[4] *Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands*
[5] *The Netherlands Foundation for the Detection of Hereditary Tumors, Leiden, The Netherlands*

KEY POINTS

- We studied eight different familial atypical multiple mole melanoma families with co-segregation of a p16-Leiden germline mutation.
- One family harbours an extraordinarily high number of tumours, comprising, breast, lung, and colon cancers, and oral squamous cell carcinomas (OSCC). In this family it seems that at least three of four lung cancer patients (one unknown), both OSCC patients, and only one of five individuals with breast cancer (two unknown) were carrying the p16-Leiden germline mutation. Immunohistochemical testing for p16 was performed and loss of heterozygosity (LOH) of the p16-Leiden wild type allele was analysed in different tumours. Additionally, four breast carcinomas and four lung tumours of eight p16-Leiden mutation positive patients from the seven remaining families were analysed.
- Immunohistochemistry of p16 was negative in all four analysed lung carcinomas.

LOH of the wild type p16 allele was present in one of three carcinomas tested. In both OSCC's, p16 immunohistochemistry was negative and LOH of the wild type allele was present in the one case analysed. Furthermore, immunohistochemistry of p16 was negative in one of five analysed breast tumours of mutation positive patients and only this tumour showed LOH of the wild type p16 allele.

- Our results suggest that the p16-Leiden germline mutation may be involved in susceptibility to lung cancer and OSCC development in some patients. There is no evidence for a dominant role of the p16-Leiden germline mutation in the development of breast cancer, although an interaction with as yet unidentified modifying factors cannot be ruled out.

INTRODUCTION

Familial atypical multiple mole melanoma (FAMMM; OMIM #155601) is characterised by the familial occurrence of melanoma of the skin in combination with multiple atypical precursor naevi.[1–4] The disease is inherited as an autosomal dominant trait, with germline mutations in the p16 (*CDKN2A*) gene having been reported in at least a quarter of FAMMM families. Previously, we reported an increased risk of pancreatic carcinoma in Dutch FAMMM families with a 19 bp deletion in exon 2 of the *CDKN2A/p16* gene (p16-Leiden; OMIM #600160.0003).[4]

Recently a patient with three carcinomas of the pharynx and oral cavity with a germline heterozygous *p16-Leiden mutation* was reported.[5] All three tumours showed inactivation of the retained wild type allele, with the somatic event being aberrant promoter methylation. Two other reports also described the occurrence of head and neck or oral squamous cell carcinomas (OSCC) in families with different *p16 germline* mutations.[6,7] A relationship between p16 germline mutations and breast cancer has also been suggested, although in the families studied, *BRCA1* and *BRCA2* mutations were not excluded.[8,9]

We studied a FAMMM family (EMC13769; Fig. 1) with co-segregation of the *p16-Leiden* germline mutation, with an extraordinary number of tumours comprising OSCC's, lung tumours, breast carcinomas, and colorectal carcinomas. We determined the mutation status in the various patients and investigated by loss of heterozygosity (LOH) analysis of the wild type allele in the tumours, in combination with immunohistochemistry, whether a causal relationship exists between the *p16-Leiden* mutation and the development of the different tumour type. Insufficient tissue was available for methylation studies. We additionally studied four breast tumours and four

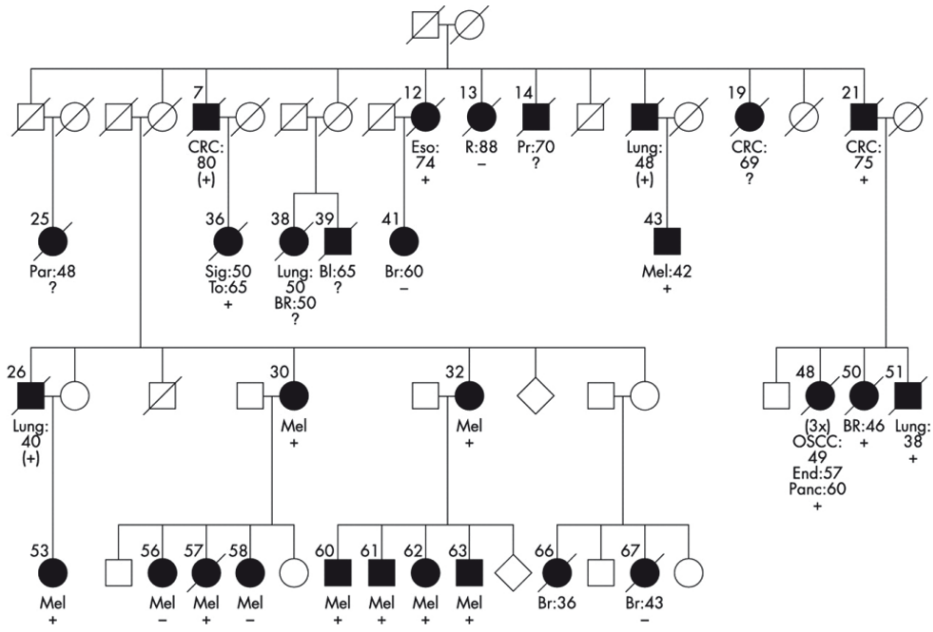


Figure 1; Pedigree of the family EMC13769.

Subject number appears above the symbol, age of diagnosis follows the diagnosis. Mel, melanoma; OSCC, oral squamous cell carcinoma. Cancer of the: Bl, bladder; Br, breast; CRC, colorectum; Eso, oesophagus; End, endometrium; Lung, lung; Panc, pancreas; Par, parotic gland; Pr, prostate; R, rectum; Sig, sigmoid; To, tongue. +, p16-Leiden positive; -, p16-Leiden negative; (+), obligate carrier; ?, p16-Leiden carrier status unknown.

lung tumours from eight other patients (from seven other families), all of whom carried a germline *p16-Leiden* mutation.

MATERIALS AND METHODS

Patients

Blood samples and/or paraffin embedded tumour samples were obtained for DNA solation from available subjects that had developed a carcinoma, to determine their *p16-Leiden* mutation status. Unavailable subjects with *p16-Leiden* positive offspring were classified as 'obligate carriers'. Informed consent was given by family members themselves or by their relatives, in case of deceased subjects. Tumours were pathologically verified whenever possible.

Tumour analysis

Paraffin embedded tumour tissues were obtained, and revision of histology was

performed. Areas of highest tumour density were selected for further molecular analysis. Serial sections were produced for immunohistochemical analysis.

DNA isolation

Genomic DNA of normal and tumour tissue was isolated from formalin fixed paraffin embedded material, resuspended in 96 µl of PK-1 lysis buffer (50 mmol/l KCL, 10 mmol/l Tris pH 8.3, 2.5 mmol/l MgCl2, 0.45% NP40, 0.45% Tween 20, 0.1 mg/ml gelatine) containing 5% Chelex beads (Biorad, Hercules, CA, USA) and 5 µl proteinase K (10 mg/ml), and incubated for 12 h at 56°C. The suspension was incubated for 10 minutes at 100°C, centrifuged, and the supernatant carefully decanted.

Polymerase chain reaction amplification

The *p16-Leiden* deletion comprises 19 bp and removes nucleotides 225–243 of exon 2.[10] Genomic DNA from tumour and normal tissue was subjected to PCR amplification using labelled primers containing the 225–243 region; *p16*-forward-TET M1 (tumour) or FAM M1 (normal), sequence 5'-ATGATGGGCAGCGCCCGAGT-3' and *p16*-reverse A2, sequence 5'-ACCAGCGTGTCCAGGAAG-3' (Life Technologies). The total volume per reaction was 12 µl including 5 µmol of each primer (stock forward and reverse primer), a mix of 0.25 µl dNTP (10 mmol/l), 1.2 µl magnesium chloride (20 mmol/l), 1.2 µl bovine serum albumin (1 mg/ml), 1.2 µl AmpliTaq Gold buffer (without MgCl2) and 0.25 µl AmpliTaq Gold DNA polymerase, 10 ng of normal or tumour DNA, and H2O. The following conditions were used: 33 cycles of 1 minute at 96°C, 2 minutes at 55°C, 1 minute at 72°C, and a delayed extension step of 7 minutes at 72°C in a GeneAmp 9700 thermocycler (Applied Biosystems, Foster City, CA, USA). Mixtures of 24 µl dionised formamide, 1 µl TAMRA 500 (Applied Biosystems) and 1.2 µl of PCR product were run on a ABI 310 Genetic analyser (Applied Biosystems) for 20 minutes with run profile GS STR POP4 (1.0 ml) C and analysed with GENESCAN 3.1 computer software (Perkin-Elmer Corp).

Loss of heterozygosity analysis

Owing to the 19 bp deletion, we could specifically analyse the fate of the wild type allele in terms of LOH. Analysis of LOH was possible when both normal and tumour tissue was available. LOH was scored when there was loss of intensity of one allele in the tumour sample with respect to the matched wild type allele from normal tissue. The quotient of the peak height ratios from normal and tumour DNA served as the allelic imbalance factor (AIF); that is, the ratio of the peak height at 101 bp of the

deleted allele and the peak height at 120 bp of the wild type allele. The threshold for allelic imbalance was defined as 40% reduction of one allele, agreeing with an AIF of <=0.59 or >1.3. The threshold for retention was defined to range from 0.76 to 1.3 as previously empirically determined.[11] AIF's of 0.60–0.75 and 1.3–1.69 were considered to belong to a so-called grey area, for which no definitive decision has been made.

Immunohistochemical testing for p16

Tissue sections (4 µm) were prepared on APES coated slides, and dried overnight in a 37°C oven. Sections were deparaffinised in xylene (3x5 minutes). Endogenous peroxidase was blocked by incubation in methanol/H2O2 0.3% for 20 minutes and sections were rehydrated with ethanol and distilled water. Antigen retrieval for p16 immunostaining was performed by microwaving in boiling 0.01 mol/l sodium citrate buffer (pH 6.0) for 10 minutes. After cooling for 2 hours and washing (2x5 minutes) in PBS, the sections were incubated overnight at room temperature with mouse anti-human p16 (1:500, clone JC8; Neomarkers Fremont, CA, USA) with tonsil tissue as positive control. Sections were subsequently washed (3x5 minutes in PBS) and incubated (30 minutes) with biotinylated secondary antibody in PBS/BSA 1%, washed (3x5 minutes in PBS) and incubated (30 minutes) with a horseradish peroxidase/streptavidin complex (SABC). Diaminobenzidine-tetrahydrochloride (DAB) was used as a chromogen, followed by counterstaining with haematoxylin. As a negative control, the primary antibody was omitted. Expression was scored by microscopic examination. Loss of p16 expression was scored when nuclei of tumour cells stained negative and nuclei of normal (stromal) cells stained positive (internal positive control).

BRCA1 and BRCA2 mutation screening

As described above, we were able to obtain tumour material of five *p16-Leiden* carriers with breast cancer. Three (NFDHT 1–3, table 1Go) had no first or second degree relative with breast cancer. The other two (EMC 13769 No 50 and LUMC 152, table 1Go) had several relatives with breast cancer diagnosed before the age of 60 years. Complete *BRCA1* and *BRCA2* mutation analysis was performed in the suspect families (EMC 13769 and LUMC 152) and found to be negative. We screened for germline mutations frequently detected in the Dutch population. Protein truncation tests[12] were also performed for PCR fragments of exon 11, and denaturating gradient gel electrophoresis was performed for the remaining exons and exon/intron junctions

of *BRCA1* and *BRCA2*. Additionally we screened for the deletions of exon 13 (3.8 kb) and exon 22 (510 bp) of *BRCA1*.[13]

Microsatellite instability

Microsatellite instability was analysed in a diagnostic setting as previously described using markers D2S123, D5S346, D17S250, BAT25, BAT26, and BAT40,[14] and immunohistochemical testing for MLH1, MSH2, and MSH6 was performed.[15]

RESULTS

Lung cancer

We analysed four different *p16-Leiden* families (Table 1, Fig. 1) with one or more cases of lung cancer. Family EMC13769 (Fig. 1) harbours four cases of lung cancer. One subject was a proven carrier of a germline *p16-Leiden* mutation (subject 51), two subjects are obligate carriers, and the *p16-Leiden* carrier status remains unknown for one (subject 38). The p16 immunohistochemistry analysis in the tumour of subject 51, a smoker, tested negative, and LOH of the wild type allele was found. The three other (NFDHT) families harbour 4 *p16-Leiden* mutation carriers with documented lung cancers. The immunohistochemistry analysis for p16 was negative in three analysed lung tumours. LOH of the wild type allele was ambiguous in one tumour, and in one tumour (carcinoid) retention was found (Table 1). In the other two tumours no normal tissue was available to perform the analysis.

Oral squamous cell carcinoma (OSCC)

Two subjects of family EMC13769 had a tumour originating in the oral cavity—that is, one tongue carcinoma (subject 36 at 65 years of age) and one subject with three primary OSCC's (subject 48 at 49 years). Immunohistochemical analysis of the tongue carcinoma was negative for p16 but lacked an internal positive control, and LOH analysis was not possible. Immunohistochemical analysis of the one of the three OSCC's from subject 48 (Fig. 1) tested negative for p16, and LOH of the wild type allele in this tumour was found (Table 1).

Breast cancer

We analysed five families with breast cancer. Family EMC 13769 shows five cases of breast cancer. Only one was carrying the *p16-Leiden* mutation (subject 50). Germline mutations in *BRCA1* & *BRCA2* were excluded for subjects no 41, 50 and 67. The p16 protein in the tumour from EMC13769 subject 50 stained positive and no LOH

TABLE 1

Results of LOH and immunohistochemical analysis in P16-Leiden mutation carriers.

Family	Subject no.	Anatomical site	Age at diagnosis	p16-IHC	Internal controle	Tumour (%)	LOH
Tumours originating in the lung and oral cavity							
EMC13769	36	OSCC (Tongue)	65	-	-		NA
	48	OSCC (1X)	49	-	+	>50	Yes
	51	Lung (adenocarcinoma)	38	-	+	>50	Yes
NFDHT4	1	Lung (SCC)	61	NA			A
NFDHT4	2	Lung (SCC)	48	-	NP		NA
NFDHT5		Lung (carcinoid)	46	-	+	70-80	R
NFDHT6		Lung (SCC)	56	-	+		NA
Tumours originating in the breast							
EMC13769	50	Breast*	46	+		>30	R
LUMC152		Breast*	41	-	+	50-60	Yes
NFDHT1		Breast	42	+		30	R
NFDHT2		Breast	47	+ †		30	R
NFDHT3		Breast	46	+ †			NA
Tumours orig**inating in the digestive tra**ct							
EMC13769	21	Colon‡	75	-	+	<30	R
	36	Sigmoid	52	+		>30	R

EMC - Erasmus MC; LUMC - Leiden University Medical Center;
NFDHT - Netherlands Foundation for the Detection of Hereditary Tumours;
OSCC - oral squamous cell carcinoma; SCC - squamous cell carcinoma;
', No staining of tumour cells or internal control cells; NP - no internal control cells identified;
R - retention of the wild type allele; A - ambiguous;
NA - not analysed; **BRCA1* and *BRCA2* tested negative; † few positive tumour nuclei;
‡ microsatellite instability analysis: immunohistochemistry for MLH1, MSH2, and MSH6 positive.

was found (Table 1). Of the four additional typed breast carcinomas from *p16* mutation carriers from the families LUMC 152 and NFDHT 1–3 (Table 1), only one showed expression loss of the p16 protein with LOH of the wild type allele, although in two of four other analysed breast carcinomas only a few tumour nuclei stained positive (with the retention of the p16 wild type allele in one, the other not tested).

Digestive tract

Family EMC13769 harbours six cases of carcinomas of the digestive tract. Of the two tumours analysed (both patients had a germline *p16-Leiden* mutation), one tumour stained positive and one negative. Neither showed LOH (Table 1), nor microsatellite instability (microsatellite stable phenotype of the tumours with normal expression of MLH1, MSH2, and MSH6).

DISCUSSION

All lung and oral cavity tumours studied developed (most likely) in *p16-Leiden* mutation carriers. For two persons we cannot rule out the possibility that the *p16-Leiden* germline mutation in their offspring came from the non-bloodline spouses. However, as this family does not come from the 'Dutch region' where multiple *p16-Leiden* mutation carriers have been identified, we think that they are most probably obligate carriers of the same *p16-Leiden* mutation. The age of onset in most patients is unusually young and abrogation of p16 seems present in all analysed cases (4/4), a ratio that seems higher than that encountered in sporadic lung cancer (36–45%).[16] The *p16-Leiden* mutation might therefore indeed predispose carriers to an increased risk of lung and oral cavity carcinomas. With respect to lung cancer, this is supported by two other important observations. Firstly, an increased cumulative risk of developing lung cancer in male *p16-Leiden* mutation carriers was found compared with the general Dutch population (14.3% v 8.9%).[4] Secondly, *Cdkn2a* is the most likely candidate for the lung tumour susceptibility locus pulmonary adenoma progression gene 1 (*PAPG1*) in mice.[17,18] *PAPG1* has been mapped to a 1.5 cM segment on chromosome 4, which contains the *Cdkn2a* gene that encodes p16INK4a. *Cdkn2a* is polymorphic between the lung tumour resistant mouse strain BALB/cJ and the lung tumour susceptible A/J strain, and the resistant allele is preferentially lost in lung tumours of p16INK4a heterozygous mice. Additionally, germline deletion of the gene in mice leads to increased tumour size and notable histological signs of malignant progression.[17]

Sufficient information on the smoking habits of most subjects in our study was lacking. However, smoking may have contributed to the unusually early age of onset of three tumours, although one of the tumours is classified as an adenocarcinoma, a type not typically associated with smoking.

Our study does not provide evidence for a dominant role of *p16-Leiden* in the development of breast cancer. Breast cancer seems also statistically not increased in our cohort studied[4] However, in view of the early onset of breast cancer in our *p16-Lei-*

den positive cases, we cannot rule out a role for the gene in tumour progression, either due to haploinsufficiency or total abrogation of p16 as seen in one of our cases (LUMC152). Recently, it has been postulated for other genes that mutation or loss of a single allele may be sufficient to play an important role in progression towards cancer.[19] Furthermore, an interaction with as yet unidentified modifying factors (genetic and/or environmental) has yet to be elucidated.

Both analysed tumours from the digestive tract showed no LOH; however, one stained negative. In this case methylation might have inactivated the wild type allele, which is a frequent event in sporadic colon cancer.[20] The role of the *p16-Leiden* germline mutation in the development of colon cancer needs further research.

In conclusion, the *p16-Leiden* mutation not only seems to predispose to melanoma and pancreatic tumours but also to head and neck tumours[5–7], and tumours of the lung in some families. Promoter methylation[5] or loss of the wild type allele seems to be the mechanism for the 'second genetic hit'. Clinical criteria for *p16* germline mutation screening should be adapted accordingly.

ELECTRONIC-DATABASE INFORMATION

Online Mendelian Inheritance in Man (OMIM), www.ncbi.nlm.nih.gov/Omim/ (for FAMMM (OMIM 155601) and p16-Leiden (OMIM 600160.0003).

ACKNOWLEDGEMENTS

The research was supported by the Netherlands Organization for Health, Research and Development (ZonMw). We would like to thank Mrs C. van der Drift for her help with this study.

REFERENCE LIST

1. **Greene MH, Tucker MA, Clark WH Jr, et al.** Hereditary melanoma and the dysplastic nevus syndrome: the risk of cancers other than melanoma. *J Am Acad Dermatol 1987 Apr;16:792–7.*
2. **Bergman W, Watson P, de Jong J, et al.** Systemic cancer and the FAMMM syndrome. *Br J Cancer 1990 Jun;61:932–6.*
3. **Goldstein AM, Fraser MC, Struewing JP, et al.** Increased risk of pancreatic cancer in melanoma-prone kindreds with p16INK4 mutations. *N Engl J Med 1995;333:970–4.*
4. **Vasen HF, Gruis NA, Frants RR, et al.** Risk of developing pancreatic cancer in families with familial atypical multiple mole melanoma associated with a specific 19 deletion of p16 (p16-Leiden). *Int J Cancer 2000;87:809–11.*
5. **Schneider-Stock R, Giers A, Motsch C, et al.** Hereditary p16-Leiden mutation in a patient with multiple head and neck tumours. *Am J Hum Genet 2003;72:216–18.*

6. **Whelan AJ, Bartsch D, Goodfellow PJ.** Brief report: a familial syndrome of pancreatic cancer and melanoma with a mutation in the CDKN2 tumor-suppressor gene. *N Engl J Med 1995;333:975–7.*
7. **Yarbrough WG, Aprelikova O, Pei H, et al.** Familial tumor syndrome associated with a germline nonfunctional p16INK4a allele. *J Natl Cancer Inst 1996;88:1489–91.*
8. **Borg A, Sandberg T, Nilsson K, et al.** High frequency of multiple melanomas and breast and pancreas carcinomas in CDKN2A mutation-positive melanoma families. *J Natl Cancer Inst 2000;92:1260–6.*
9. **Plna K, Hemminki K.** Re: High frequency of multiple melanomas and breast and pancreas carcinomas in CDKN2A mutation-positive melanoma families. *J Natl Cancer Inst 2001;93:323–5.*
10. **van der Velden PA, Sandkuijl LA, Bergman W, et al.** A locus linked to p16 modifies melanoma risk in Dutch familial atypical multiple mole melanoma (FAMMM) syndrome families. *Genome Res 1999;9:575–80.*
11. **Cleton-Jansen AM, Callen DF, Seshadri R, et al.** Loss of heterozygosity mapping at chromosome arm 16q in 712 breast tumours reveals factors that influence delineation of candidate regions. *Cancer Res 2001;61:1171–7.*
12. **Hogervorst FB, Cornelis RS, Bout M, et al.** Rapid detection of BRCA1 mutations by the protein truncation test. *Nat Genet 1995;10:208–12.*
13. **Petrij-Bosch A, Peelen T, van Vliet M, et al.** BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nat Genet 1997;17:341–5.*
14. **Boland CR, Thibodeau SN, Hamilton SR, et al.** A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res 1998;58:5248–57.*
15. **Hendriks Y, Franken P, Dierssen JW, et al.** Conventional and tissue microarray immunohistochemical expression analysis of mismatch repair in hereditary colorectal tumours. *Am J Pathol 2003;162:469–77.*
16. S**anchez-Cespedes, Decker PA, Doffek KM, et al.** Increased loss of chromosome 9p21 but not p16 inactivation in primary non-small cell lung cancer from smokers. *Cancer Res 2001;61:2092–6.*
17. **Zhang Z, Wang Y, Herzog CR, et al.** A strong candidate gene for the Papg1 locus on mouse chromosome 4 affecting lung tumor progression. *Oncogene 2002;21:5960–6.*
18. **Demant P.** Cancer susceptibility in the mouse: genetics, biology and implications for human cancer. *Nat Rev Genet 2003;4:721–34.*
19. **Fodde R, Smits R.** Cancer biology. A matter of dosage. *Science 2002;298:761–3.*
20. **Burri N, Shaw P, Bouzourene H, et al.** Methylation silencing and mutations of the p14ARF and p16INK4a genes in colon cancer. *Lab Invest 2001;81:217–29.*

3.2. THE CHEK2*1100delC VARIANT ACTS AS A BREAST CANCER RISK MODIFIER IN NON-BRCA1/BRCA2 MULTIPLE-CASE FAMILIES

Rogier A. Oldenburg[1,4], Karin Kroeze-Jansema[1], Jaennelle Kraan[1], Hans Morreau[2], Jan G. M. Klijn[5], Nicoline Hoogerbrugge[7], Marjolein J. L. Ligtenberg[7], Christi J. van Asperen[1], Hans F. A. Vasen[8], Carel Meijers[6], Hanne Meijers-Heijboer[4], Truuske H. de Bock[3], Cees J. Cornelisse[2] and Peter Devilee[1,2]

Cancer Research. 2003 Dec 1;63(23): 8153-7;

1. Centre for Human and Clinical Genetics, and
2. Departments of Pathology and
3. Medical Decision Making, Leiden University Medical Centre, Leiden;
4. Departments of Clinical Genetics,
5. Medical Oncology, and
6. Pathology, Erasmus MC, Rotterdam;
7. Department of Human Genetics, University Medical Centre, Nijmegen; and
8. Netherlands Foundation for the Detection of Hereditary Tumors, Leiden, the Netherlands

ABSTRACT

The frame-shifting mutation 1100delC in the cell-cycle-checkpoint kinase 2 gene (*CHEK2*) has been reported to be associated with familial breast cancer in families in which mutations in *BRCA1* and *BRCA2* were excluded. To investigate the role of this variant as a candidate breast cancer susceptibility allele, we determined its prevalence in 237 breast cancer patients and 331 healthy relatives derived from 71 non-*BRCA1/BRCA2* multiple-case early onset breast cancer families. Twenty-seven patients (11.4%) were carrying the *CHEK2*1100delC* variant. At least one carrier was found in 15 of the 71 families (21.1%). There was no evidence of cosegregation between the variant and breast cancer, but carrier patients developed breast cancer earlier than did noncarriers. We studied CHEK2 protein expression in 111, and loss of heterozygosity at *CHEK2* in 88 breast tumors from these patients. Twelve of 15 tumors from carriers showed absent protein expression as opposed to 3 of 76 tumors from noncarriers ($P < 0.001$). *CHEK2* loss of heterozygosity was associated with absence of protein expression but not with 1100delC carrier status. Thus, selecting for breast

cancer cases with a strong familial background not accounted for by *BRCA1* or *BRCA2* strongly enriches for carriers of *CHEK2*1100delC*. Our results support a model in which *CHEK2*1100delC* interacts with an as yet unknown gene (or genes) to increase breast cancer risk.

INTRODUCTION

First-degree female relatives of a breast cancer patient have an 2-fold increased risk to develop breast cancer.[1] Germ-line mutations in the *BRCA1* and *BRCA2* genes account for <5% of this familial risk.[2, 3] To explain the remainder of familial risk, various genetic models have been proposed. Models incorporating a single third hypothetical gene, *BRCA3*, or a number of common low penetrance genes with additive effect seem to fit equally well, although the latter fitted best when the known effects of parity on breast cancer risk were included.[3, 4] A mutation 1100delC in *CHEK2* has been proposed recently to be a low-penetrance breast cancer susceptibility allele.[5, 6] *CHEK2* is located on chromosome 22 and encodes the human orthologue of yeast *Cds1* and *Rad53*, which are G2 checkpoint kinases.[7] *CHEK2* is involved in cell cycle control and DNA repair through its ability to phosphorylate p53, Cdc25c, and *BRCA1*. The *CHEK2*1100delC* variant is a protein-truncating mutation that abrogates the kinase activity of the protein. It occurs in 0.3–1.4% of healthy control individuals,[5, 6, 8] but in about double that frequency among unselected cases of breast cancer. It is even further enriched among breast cancer cases with a positive family history in which *BRCA1* and *BRCA2* mutations have been excluded. Up to 5.5% of such cases may be carrying the *CHEK2*1100delC* variant, although it apparently incompletely segregates with breast cancer in the families of these cases.[5] Other variants in *CHEK2* seem to be very rare and are not enriched among familial breast cancer cases.[9–11] We have embarked recently on a genome-wide linkage search for new breast cancer susceptibility genes in a highly selected group of breast cancer families. Phenotypic and genotypic criteria[12] have minimized the probability that these families harbor mutations in *BRCA1* or *BRCA2*, but have selected for families that are caused by other high penetrant genes. Here, we investigate the role of the *CHEK2*1100delC* variant as a cause of breast cancer in these families.

MATERIALS AND METHODS

Families.

Families were ascertained through the Clinical Genetic Centres in Leiden, Rotterdam, and Nijmegen, as well as through the Netherlands Foundation for the Detec-

tion of Hereditary Tumors. Families were eligible if there were at least three cases of breast cancer diagnosed before the age of 60 from whom genotypes could be determined or inferred. DCIS or LCIS before the age of 60 as first primary cancer were also considered eligible diagnoses. Families with cases of ovarian cancer or male breast cancer were excluded, and occurrences of other cancer types were ignored. Seventeen of these 71 families were also part of the previous study identifying the *1100delC*CHEK2* variant as a low-penetrance breast cancer susceptibility gene.[5] The 71 families selected contained a total of 384 breast cancer patients, 297 of which diagnosed before the age of 60, 2 of which occurred in spouses (excluded from the statistical analysis), and 5 of which had in situ cancer (4 DCIS and 1 LCIS) only. There was one family where the third case diagnosed under 60 was an in situ cancer (combined DCIS/LCIS at age 53).

Pathology reports were retrieved for 260 patients (68%). For another 84 patients, diagnoses were confirmed by medical records, and retrieval of pathology reports was still in progress at the time when this study was finalized. For the remaining 40 cases, breast cancer diagnoses were ascertained by family interview only. Blood samples and paraffin-embedded tumor tissues were collected after obtaining written informed consent. The institutional ethical committees of all of the hospitals involved approved this study.

BRCA1 and BRCA2 Mutation Testing.

In each family, the youngest breast cancer patient from whom a blood sample was available was tested for mutations in the *BRCA1* and *BRCA2* genes (and for many families the next youngest as well). The different Clinical Genetic Centers applied a variety of methodologies. The large central exons (exon 11 in *BRCA1* and *BRCA2*, exon 10 of *BRCA2*) were scanned by protein truncation tests.[13] The small exons were scanned for mutations by denaturing gradient gel electrophoresis or direct sequencing. All of the laboratories specifically assayed the presence of large founder deletions in *BRCA1* by deletion junction-PCR.[14] For cases where scanning was still in progress at the time of sampling for the purpose of this research, we performed conformation-sensitive gel electrophoresis[15] covering all of the coding regions of both genes. This identified 10 different variants of uncertain clinical significance and 12 different polymorphisms. None of these were cosegregating with breast cancer or the *CHEK2*1100delC* variant.

Genotyping of the CHEK2*1100delC Variant.

The DNA sequence of exon 10 of *CHEK2*, where the 1100delC resides, is present in multiple homologous copies in the genome. For PCR, we used oligonucleotides 10F (5' TGT CTT CTT GGA CTG GCA GA; Fam-labeled) and 10R (5' ATC ACC TCC TAC CAG TCT GTG C), which specifically amplify the functional copy of *CHEK2*, relative to the nonfunctional pseudogenes.[16] The reaction volume of 10 μl contained 20 ng of genomic DNA, 1 μl 10' SuperTaq buffer (HT Biotechnology LTD.), 1 mM dNTPs, 300 mM of each primer, and 0.1 units of Silverstar DNA polymerase (Eurogentec). Annealing temperature was 65°C, and the PCR ran for 38 cycles. The resulting PCR-products were analyzed on an ABI3700, in fragment analysis mode. The wild-type allele runs as a 291-bp fragment and the mutant allele as a 290-bp fragment, which are readily separated into two peaks of about equal signal intensity in this assay. All of the positive samples were confirmed by sequencing as described previously.[5]

LOH Analysis.

LOH at the *CHEK2* locus was investigated by comparing the genotypes in normal and tumor DNA at four flanking markers, D22S420, D22S315, D22S280, and D22S283. CHEK2 maps between D22S315 and D22S280, which span an interval of 7 Mb. Four punches (5 mm long and 0.6 mm in diameter) were taken from paraffin-embedded tumor tissues, in the area where the tumor was located. These punches generally contain >50% tumor cells. DNA was isolated from these punches as described previously.[17] Allelic imbalance was defined as the ratio of allele intensities in the normal versus the tumor DNA. An AIF of 1.70 was scored positive.[18] LOH at the *CHEK2* locus was scored positive when the AIF- pattern was such that at least one proximal and one distal marker showed AIF 1.70 without interruption by a marker showing an AIF <1.70.

Tissue Array and Immunohistochemical Analyses.

All of the tumor samples were embedded in standard paraffin blocks. On the respective H&E-stained sections, a representative tumor area was selected. Two to four tissue cores (0.6 mm in diameter; Beecher Instruments, Silver Spring, MD) were punched from the designated area using a biopsy needle and arrayed into the recipient blocks. Using a tape-transfer system (Instrumedics, Hackensack, NJ), 4-μm sections were transferred to glass slides. For antigen retrieval, the deparaffinized sections were boiled in a microwave for 15 min in citrate buffer (pH 6.0) before incubation with a mouse monoclonal antibody, NCL-CHK2 (Novocastra Laboratories, LTD.,

Newcastle, United Kingdom), directed against the human CHEK2 protein. After this the slides were incubated with a second step antibody streptavidin-biotin labeled (Labvision) for 90 min. Two independent pathologists evaluated the staining results without prior knowledge of the mutation status of *CHEK2*. The tumors were scored as having an absent, weak, moderate, or high CHEK2 protein expression depending on the intensity of the staining regardless of the proportion of tumor cells falling in this category. When no staining was found, an absent protein expression was scored.

Statistical Analysis.

Prevalences, clinical characteristics of patients, and tumors were compared between groups by 2 tests. All of the tests of statistical significance were two-sided. A t-test was used to compare mean ages of onset between carriers and noncarriers. Additionally, Kaplan-Meier age of onset probability curves were estimated and differences were tested by the log-rank test. To obtain an impression of the size of the effect of a *CHEK2*1100delC* mutation on age of onset, a Cox-regression analysis was performed.

RESULTS

We investigated 71 families with a phenotype of early onset breast cancer, defined as having at least 3 cases diagnosed before the age of 60, and no cases of ovarian or male breast cancer. Mutations in *BRCA1* and *BRCA2* were excluded in at least the youngest breast cancer case from which a blood sample was available. These families contained a total of 384 breast cancer patients. We collected DNA samples from 237 patients, including all of those with in situ cancer, as well as from 331 family members without breast cancer and 54 spouses. Of the 622 individuals we were thus able to assay for the presence of the *CHEK2*1100delC* variant, we found 41 (6.6%) to be carriers (Table 1). The prevalence among breast cancer patients was 11.4% (27 of 237), which was significantly higher than the prevalence of the variant in healthy female family members (6 of 212; $2 = 12.047$; df = 1; $P < 0.001$). Three carriers were known with in situ cancer (2 DCIS and 1 LCIS). Fifteen families (21.1%) had at least 1 positive individual for this variant. One of these was a family in which the only identifiable carrier was a woman with in situ cancer (DCIS; Fig. 1). The proportion of families in which at least 1 individual carried the *CHEK2* variant increased to 31.8% in families with >5 breast cancer patients diagnosed under 60 (Table 1). However, this trend was not statistically significant ($2 = 2.6$; df = 2; $P = 0.272$). In addition, *CHEK2*-posi-

TABLE 1

CHEK2*1100delC prevalences

Description	Total	CHEK2+	%
All sampled individuals	**622**	**41**	**6.6**
Male	154	8	5.2
Female	468	33	7
All sampled breast cancer cases	**237**	**27**	**11.4**
Cases diagnosed under 60	194	24	12.4
Cases diagnosed 60 or over	43	3	7.0
Cases with in situ cancer only	5	3	60.0
Healthy family members	**331**	**14**	**4.2**
Males	119	8	6.7
Females	212	6	2.8
Spouses [a]	**54**	**0**	**0**
Male	35	0	0
Female	19	0	0
All families	**71**	**15**	**21.1**
3 cases < 60	30	4	13.3
4 cases < 60	19	4	21.1
>= 5 cases < 60	22	7	31.8

[a] Two of these individuals were diagnosed with breast cancer.

TABLE 2

LOH at CHEK2

LOH at CHEK2	Number of cases	CHEK2 carriers	%
Positive[a]	11	3	27.3
Suspected[b]	29	5	17.2
Negative	29	3	10.3
Unknown[c]	20	3	15.0
Totals	89	14	14.0

[a] Cases in which at least one proximal and one distal marker showed AIF 1.70 without interruption by a marker showing an AIF < 1.70.

[b] Cases in which LOH was found only proximal or distal of CHEK2.

[c] Cases in which one of the reactions failed.

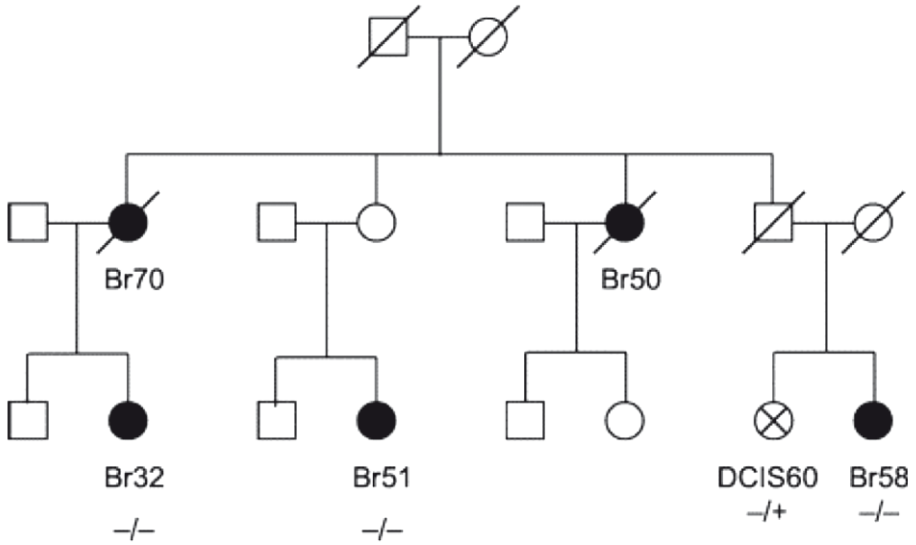


Fig. 1. Pedigree of family RUL154.
Filled symbols are individuals diagnosed with breast cancer, the age at diagnosis is given below the symbol. -/+ indicates that the individual carries the *CHEK2*1100delC* variant; -/- indicates the individual does not carry this variant.

tive families had on average slightly more blood-sampled cases than *CHEK2*-negative families (3.8 versus 3.2; data not shown). Although not a statistically significant difference, this indicates that the odds of detecting the variant is dependent on the number of blood-sampled breast cancer cases in a family.

In the 15 *CHEK2*1100delC*-positive families we defined the youngest carrier breast cancer case as the index patient. Under the null hypothesis of complete random Mendelian inheritance, we predicted that 12.9375 of the 54 affected relatives would be carrier of the variant. We observed 12 carriers, so that the null hypothesis could not be rejected. We performed LOH analysis in 89 archival breast tumor tissues from 88 breast cancer cases from these 71 families, at four markers mapping to either side of *CHEK2* (Table 2). LOH at *CHEK2* was found in 11 tumors, 3 of which derived from 2 *CHEK2*1100delC* carriers. In all 3 of the tumors, we could demonstrate that the lost allele was derived from the nontransmitting parent (data not shown). Although the 1100delC variant occurred 2.7 times more frequently among cases showing LOH at *CHEK2*, this difference was not statistically significant ($\chi^2 = 1.239$; df = 2; $P = 0.538$).

A tissue microarray with 111 tumors from 111 cases was stained with a mouse monoclonal antibody against the human CHEK2 protein. Examples of obtained staining

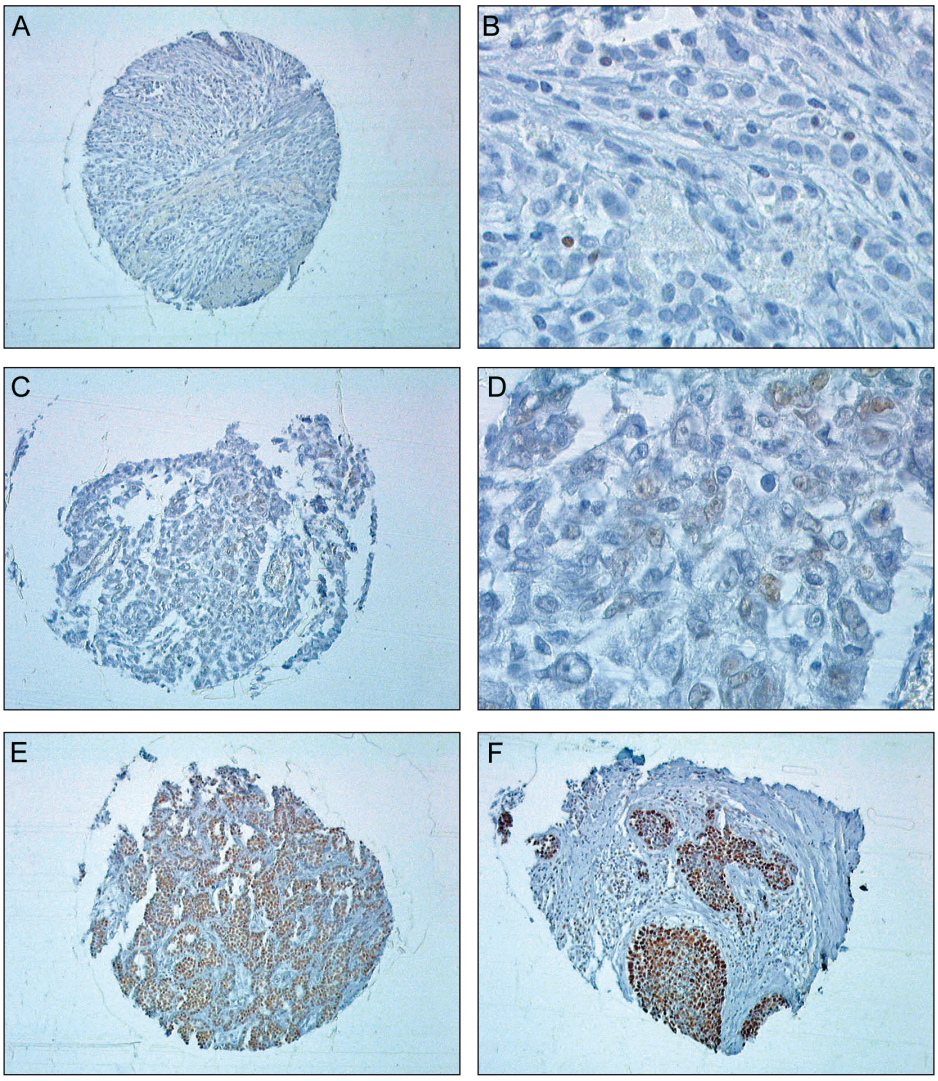


Fig. 2. Immunohistochemical staining of chek2 in human breast tumors on a tissue microarray.
The samples shown are from four different tumors and represent the four different scoring categories used here. A and B, absent protein expression in a tumor from a chek2*1100delC carrier. Note the scattered strongly staining normal epithelial cells as positive internal control (B). C–F, represent tumors from noncarriers. C and D, weak protein expression. E, moderate protein expression. F, high expression. Magnification x25 in A, C, E, and F. and x100 in B and D.

TABLE 3

Chek2 protein expression according to 1100delC carrier status and LOH

Variable	CHEK2 protein expression				Total
	Absent	Weak	Moderate	Strong	
CHEK2 +[a]	12	2	1	0	15
CHEK2 -[b]	3	41	27	5	76
LOH +[c]	3	7	0	0	10
LOH suspected[d]	7	12	8	1	28
LOH -[e]	0	11	14	2	27
LOH unknown[f]	4	9	5	1	19

[a] CHEK2 +, carriers of the 1100delC variant.

[b] CHEK2 -, noncarrier.

[c] LOH+, at least one proximal and one distal marker showed AIF 1.70 without interruption by a marker showing an AIF < 1.70.

[d] LOH suspected, one distal or proximal marker showed an AIF < 1.70 while the closest marker on the other side of CHEK2 was uninformative.

[e] LOH -, no LOH was found.

[f] LOH unknown, one of the reactions failed.

patterns are shown in Fig. 2. As noted in a previous study[19] there was considerable variability in the percentage of normal cells that were positive. CHEK2 protein expression was absent in 12 of 15 tumors from *CHEK2*1100delC* carriers (80.0%; Table 3). False-negative staining was considered unlikely, because in 6 of 12 tumors from *CHEK2*1100delC* carriers the stromal component stained normally.

Notably, the one tumor showing moderate protein expression was an in situ carcinoma (DCIS) from a patient from family RUL154 (Fig. 1). In comparison, only 3 of 76 tumors (3.9%) from noncarriers showed an absent CHEK2 protein expression ($2 = 52.709$; df = 3; $P < 0.001$). For 37 tumors, protein expression and LOH data were available. CHEK2 protein expression was absent in 3 of 10 tumors with *CHEK2*-LOH, 2 of which were from *CHEK2*1100delC* carriers. The other 7 tumors with *CHEK2*-LOH all showed a weak CHEK2 protein expression. In comparison, all 27 of the tumors, which retained heterozygosity at *CHEK2*, showed some degree of protein expression ($2 = 15.879$; df = 6; $P = 0.014$). The mean age of diagnosis of the first primary tumor of *CHEK2*1100delC* carrier patients was not significantly different from that in noncarriers (48.3 versus 50.6 years; $P = 0.30$). However, any age difference may have been

masked by our selection for early onset breast cancer. Indeed, in a Kaplan-Meier analysis the age of onset distribution between the two groups was different ($P < 0.0001$). It is unlikely that this effect is confounded by differences in tumor grade because the percentage of grade III tumors was higher in noncarriers than in carriers (22 of 81 versus 1 of 9). A Cox-regression analysis revealed an odds ratio of 2.1 (95% confidence interval, 1.393–3.166; $P < 0.001$) for carriers to develop breast cancer relative to noncarriers (derived from *CHEK2*1100delC* positive and *CHEK2*1100delC* negative families). Among the 237 genotyped breast cancer patients in our cohort, 35 (14.8%) were known to have had a second primary
breast cancer. Five of these (14.3%) were positive for the *CHEK2* variant. Of the 202 patients with one primary breast cancer, 22 tested positive (10.9%). This difference was not statistically significant.

DISCUSSION

We found the *CHEK2*1100delC* variant in 11.4% of the breast cancer cases belonging to a highly selected group of families. This prevalence was substantially higher then reported previously by others. Two earlier studies[5, 6] selected familial breast cancer cases from families that were not linked to *BRCA1* or *BRCA2*, and found a prevalence of 5.1% and 5.5%, respectively. The families we studied are highly selected in several ways. First, they contain at least 3 breast cancer cases diagnosed before age 60 (the average number of breast cancer cases per family was 5.4). Second, they were selected against cases of ovarian and male breast cancer. Third, they all tested negative for mutations in *BRCA1* and *BRCA2*. On the basis of population incidence, the odds that 3 cases in a family occur under 60 by chance alone are very low, and, thus, they likely have a genetic basis. Hence, in this group of families we suspect an enrichment of a gene (or genes) other than *BRCA1* and *BRCA2* that may confer substantial breast cancer risks.[12] However, because we and others[5, 6] found no or weak evidence for cosegregation between *CHEK2*1100delC* and breast cancer, *CHEK2* is an unlikely candidate for such a gene. It is possible that other, more high-risk mutations in *CHEK2* exist that could account for these cases, but this has thus far not been substantiated by more comprehensive mutation scanning of the gene (9, 10, 20, 21). A more likely explanation for the data presented here is a model in which *CHEK2*1100delC* interacts with an as yet unknown rare gene (or genes) to confer breast cancer risks comparable with those conferred by *BRCA1* or *BRCA2*. Selecting for families caused by this rare gene would also enrich for *CHEK2*1100delC* carriers, which would act like a modifier of the breast cancer risk. The CHEK2 Consortium, studying families of

Dutch, German, United Kingdom, and North American origin, found the prevalence of the 1100delC variant to increase in families with 4 cases,[5] but the Finnish study found the highest prevalence among non-*BRCA1/2* cases with a moderate family history.[6] We also found weak evidence for increasing prevalence of *CHEK2*1100delC* among families with a more extensive family history of breast cancer. Even among populations with an apparently overall lower prevalence of the 1100delC variant,[8] this enrichment is observed. The higher allele frequency in Northern Europe as opposed to North America might be due to a founder effect of *CHEK2*1100delC*. The proposed risk modifying effect of *CHEK2*1100delC* is also supported by our finding that carriers in our families develop breast cancer systematically earlier than do noncarriers. Although this may be a peculiarity of this selected group of patients, a similar age-effect has been noted for genetic variants in *AR*, *HRAS1*, *RAD51*, and *AIB1* in carriers of *BRCA1* or *BRCA2* mutations.[22, 23, 24, 25] Alternatively, breast cancer in these families has a polygenic basis involving multiple interacting low-penetrance alleles,[26] one of which is the *CHEK2*1100delC* variant. The *CHEK2*1100delC* is approximately twice as prevalent among unselected breast cancer cases than among controls, suggesting it is a low-risk allele in its own right.[5, 6] In keeping with this, we found that *CHEK2*1100delC* is associated with breast cancer, but it was unable to explain the majority of breast cancer cases in these families. A role for *CHEK2* inactivation in breast tumor development is nonetheless supported by the highly significant association we found between *CHEK2*1100delC* carrier status and an absence of protein expression in the breast tumors. This confirms results obtained by others[6, 19] irrespective of minor differences in interpretation of immunohistochemical staining patterns among these studies. It would also explain the slightly earlier age of onset of breast cancer in 1100delC carriers, as these individuals only need to inactivate the wild-type allele whereas noncarriers would need to inactivate both copies of the gene. Paradoxically, the breast tumors of *CHEK2* carriers do not significantly more frequently show LOH at *CHEK2*. Hence, LOH may not be the only mechanism inactivating the wild-type allele, although the association between LOH and an absent protein expression we observed does indicate it is involved in some cases. Alternative mechanisms include promoter hypermethylation[27] and somatic mutations, but the roles of both appear to be marginal in breast cancer.[19, 28] Conceivably, other components of the pathway(s) regulating the expression and/or stability of CHEK2 protein are disturbed in these cases. An association with bilateral disease, but only a marginal trend toward earlier age of diagnosis was reported in one study.[6] In our cohort of cases we found an association between *CHEK2* carrier

status and earlier age of diagnosis but not between carrier status and multiple primary tumors. This could be a peculiarity of the selected families. Conceivably, many cases not carrying the *CHEK2* variant are carriers of another gene defect that predisposes them strongly to develop breast cancer. In combination with a long retrospective follow-up time, this may have masked the subtle effect of *CHEK2* on risk. In conclusion, we find a strong association between *CHEK2*1100delC* prevalence and breast cancer family history. Our results provide support for the hypothesis that this variant modifies the cancer risk conferred by an as yet unknown gene (or genes). Given the cancer occurrence in the families described here, this gene is expected to cause breast cancer risks comparable with those conferred by *BRCA1* and *BRCA2*. At this point it is in our opinion not appropriate to offer a predictive test for *CHEK2* in a clinical setting. The exact relative risk conferred by *CHEK2*1100delC* is not clear, but likely modest in comparison with *BRCA1* and *BRCA2*. In addition, estimates of breast cancer risk are difficult to make in these families, because the type of interaction (multiplicative or additive) and the role of other factors are presently unknown. Selecting for families with at least one carrier of the *CHEK2*1100delC* might reduce the genetic heterogeneity likely to exist among non-*BRCA1/BRCA2* families and facilitate the mapping of this breast cancer susceptibility gene by classical linkage analysis.

ACKNOWLEDGMENTS

We thank Klaas G. van der Ham for technical assistance with photography of immunohistochemistry results.

REFERENCE LIST

1. **Pharoah P. D. P., Day N. E., Duffy S., et al.** Family history and the risk of breast cancer: A systematic review and meta-analysis. *Int. J. Cancer, 71: 800-809, (1997)*
2. **Peto J., Collins N., Barfoot R., et al.** Prevalence of BRCA1 and BRCA2 gene mutations in patients with early- onset breast cancer. *J. Natl. Cancer Inst., 91: 943-949, (1999)*
3. **Antoniou A. C., Pharoah P. D., McMullan G., et al.** Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. *Genet. Epidemiol., 21: 1-18, (2001)*
4. **Cui J., Antoniou A. C., Dite G. S., et al.** After BRCA1 and BRCA2-what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer. *Am. J. Hum. Genet., 68: 420-431, (2001)*
5. **Meijers-Heijboer H., Van den Ouweland A., Klijn J., et al.** Low-penetrance susceptibility to breast cancer due to CHEK2*1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat. Genet., 31: 55-59, (2002)*
6. **Vahteristo P., Bartkova J., Eerola H., et al.** A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. *Am. J. Hum. Genet., 71: 432-438, (2002)*

7. **Bartek J., Falck J., Lukas J.** CHK2 kinase-a busy messenger. *Nat. Rev. Mol. Cell Biol., 2: 877-886, (2001)*
8. **Offit K., Pierce H., Kirchhoff T., et al.** Frequency of CHEK2*1100delC in New York breast cancer cases and controls. *BMC. Med. Genet., 4: 1-4, (2003)*
9. **Allinen M., Huusko P., Mantyniemi S., et al.** Mutation analysis of the CHK2 gene in families with hereditary breast cancer. *Br. J. Cancer, 85: 209-212, (2001)*
10. **Sodha N., Bullock S., Taylor R., et al.** CHEK2 variants in susceptibility to breast cancer and evidence of retention of the wild type allele in tumours. *Br. J. Cancer, 87: 1445-1448, (2002)*
11. **Schutte M., Seal S., Barfoot R., et al.** Variants in CHEK2 other than 1100delC do not make a major contribution to breast cancer susceptibility. *Am. J. Hum. Genet., 72: 1023-1028, (2003)*
12. **Ford D., Easton D. F., Stratton M., et al.** Breast Cancer Linkage Consortium Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. *Am. J. Hum. Genet., 62: 676-689, (1998)*
13. **Hogervorst F., Cornelis R., Bout M., et al.** Rapid detection of BRCA1 mutations by the Protein Truncation Test. *Nat. Genet., 10: 208-212, (1995)*
14. **Petrij-Bosch A., Peelen T., Van Vliet M., et al.** BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nat. Genet., 17: 341-345, (1997)*
15. **Korkko J., Annunen S., Pihlajamaa T., et al.** Conformation sensitive gel electrophoresis for simple and accurate detection of mutations: comparison with denaturing gradient gel electrophoresis and nucleotide sequencing. Proc. *Natl. Acad. Sci. USA, 95: 1681-1685, (1998)*
16. **Sodha N., Houlston R. S., Williams R., et al.** A robust method for detecting CHK2/RAD53 mutations in genomic DNA. *Hum. Mutat.,19:173-177, (2002)*
17. **Vos C. B., Ter Haar N. T., Peterse J. L., et al.** Cyclin D1 gene amplification and overexpression are present in ductal carcinoma in situ of the breast. *J. Pathol., 187: 279-284, (1999)*
18. **Cleton-Jansen A. M., Callen D. F., Seshadri R., et al.** Loss of heterozygosity mapping at chromosome arm 16q in 712 breast tumors reveals factors that influence delineation of candidate regions. *Cancer Res., 61: 1171-1177, (2001)*
19. **Sullivan A., Yuille M., Repellin C., et al.** Concomitant inactivation of p53 and Chk2 in breast cancer. *Oncogene, 21: 1316-1324, (2002)*
20. **Vahteristo P., Tamminen A., Karvinen P., et al.** p53, CHK2, and CHK1 genes in Finnish families with Li-Fraumeni syndrome: further evidence of CHK2 in inherited cancer predisposition. *Cancer Res., 61: 5718-5722, (2001)*
21. **Miller C. W., Ikezoe T., Krug U., et al.** Mutations of the CHK2 gene are found in some osteosarcomas, but are rare in breast, lung, and ovarian tumors. *Genes Chromosome Cancer (Phila.), 33: 17-21, (2002)*
22. **Phelan C. M., Rebbeck T. R., Weber B. L., 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)*
23. **Rebbeck T. R., Kantoff P. N., Krithivas K., et al.** Modification of BRCA1-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat. *Am. J. Hum. Genet., 64: 1371-1377, (1999)*
24. **Rebbeck T. R., Wang Y. T., Kantoff P. W., et al.** Modification of BRCA1- and BRCA2-associated breast cancer risk by AIB1 genotype and reproductive history. *Cancer Res., 61: 5420-5424, (2001)*
25. **Levy-Lahad E., Lahad A., Eisenberg S., et al.** A single nucleotide polymorphism in the RAD51 gene modifies cancer risk in BRCA2 but not BRCA1 carriers. *Proc. Natl. Acad. Sci. USA, 98: 3232-3236, (2001)*
26. **Weber B. L., Nathanson K. L.** Low penetrance genes associated with increased risk for breast cancer. *Eur. J. Cancer, 36: 1193-1199, (2000)*
27. **Esteller M.** CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene, 21: 5427-5440, (2002)*

28. **Ingvarsson S., Sigbjornsdottir B. I., Chen H. P., et al.** Mutation analysis of the CHK2 gene in breast carcinoma and other cancers. *Breast Cancer Res., 4: NIL (2002)*

3.3. EVALUATION OF LINKAGE OF BREAST CANCER TO THE PUTATIVE *BRCA3* LOCUS ON CHROMOSOME 13q21 IN 128 MULTIPLE CASE FAMILIES FROM THE BREAST CANCER LINKAGE CONSORTIUM

Deborah Thompson, Csilla I. Szabo, Jon Mangion, Rogier A. Oldenburg, Fabrice Odefrey, Sheila Seal, Rita Barfoot, Karin Kroeze-Jansema , Dawn Teare, Nazneen Rahman, Hélène Renard KConFab Consortium, Graham Mann, John L. Hopper, Saundra S. Buys, Irene L. Andrulis, Ruby Senie, Mary B. Daly, Dee West, Elaine A. Ostrander, Ken Offit, Tamar Peretz, Ana Osorio, J. Benitez, Katherine L. Nathanson, Olga M. Sinilnikova, Edith Olàh, Yves-Jean Bignon, Pablo Ruiz, Michael D. Badzioch, Hans F. A. Vasen, Andrew P. Futreal, Catherine M. Phelan, Steven A. Narod, Henry T. Lynch, Bruce A. J. Ponder, Ros A. Eeles, Hanne Meijers-Heijboer, Dominique Stoppa-Lyonnet, Fergus J. Couch, Diana M. Eccles, D. Gareth Evans, Jenny Chang-Claude, Gilbert Lenoir, Barbara L. Weber, Peter Devilee, Douglas F. Easton, David E. Goldgar, and Michael R. Stratton.

Proc Natl Acad Sci U S A. 2002 Jan 22;99(2):827-31.

ABSTRACT

The known susceptibility genes for breast cancer, including *BRCA1* and *BRCA2*, only account for a minority of the familial aggregation of the disease. A recent study of 77 multiple case breast cancer families from Scandinavia found evidence of linkage between the disease and polymorphic markers on chromosome 13q21. We have evaluated the contribution of this candidate '*BRCA3*' locus to breast cancer susceptibility in 128 high-risk breast cancer families of Western European ancestry with no identified *BRCA1* or *BRCA2* mutations. No evidence of linkage was found. The estimated proportion (α) of families linked to a susceptibility locus at D13S1308, the location estimated by Kainu et al. [(2000) *Proc. Natl. Acad. Sci. USA* 97, 9603–9608], was 0 (upper 95% confidence limit 0.13). Adjustment for possible bias due to selection of families on the basis of linkage evidence at *BRCA2* did not materially alter this result (α = 0, upper 95% confidence limit 0.18). The proportion of linked families reported by Kainu et al. (0.65) is excluded with a high degree of confidence in our dataset [heterogeneity logarithm of odds (HLOD) at α = 0.65 was −11.0]. We conclude that, if a susceptibility gene does exist at this locus, it can only account for a

small proportion of non-*BRCA1/2* families with multiple cases of early-onset breast cancer.

INTRODUCTION

Several genes are known to predispose to breast cancer. In the context of large multiple case families, the *BRCA1* and *BRCA2* genes are numerically the most important, accounting for most families segregating both early-onset breast cancer and ovarian cancer. However, as many as 60% of families with site-specific female breast cancer cannot be explained by *BRCA1* and *BRCA2*.[1,2] Moreover, population studies have demonstrated that these genes only account for 15% of the overall familial risk of breast cancer.[3,4] Even after allowing for other susceptibility genes that confer increased risk in the context of familial cancer syndromes, including *TP53* (Li Fraumeni), *PTEN* (Cowden), and *ATM* (ataxia telangiectasia), at least 80% of familial breast cancer risk is not explained by known genes, suggesting that other important susceptibility genes remain to be mapped. Outside the context of these specific syndromes, known genes other than *BRCA1*/*BRCA2* do not appear to account for a substantial proportion of high-risk breast cancer families. Linkage analysis in a set of 56 families with 3 or more cases of breast cancer yielded no evidence for a significant role of *PTEN*, although an attributable fraction of up to 35% could not be ruled out in a family set of this size.[5] However, direct mutation testing of the *PTEN* gene in a subset of these families has failed to identify any mutations, lending further support to the linkage results indicating that this locus is unlikely to account for a significant fraction of hereditary breast cancer.

To date, few additional candidate breast cancer susceptibility loci have been identified in families not attributable to any of the known genes. A potential susceptibility locus on chromosome 8p12–8p22 was identified through targeted linkage analysis of a region of frequent loss in breast tumors.[6,7] However, our analysis of a larger family series did not support the contribution of a putative gene at this locus to more than a small proportion [HLOD = 0.03, α= 0.03, upper 95% confidence limit (CL) 0.30] of high-risk families.[8]

These findings illustrate the difficulties inherent in efforts to identify additional susceptibility genes for a disease with high population prevalence. First, breast cancer is a genetically heterogeneous disease, and it is likely that there are multiple genes remaining to be identified among non-*BRCA1*/*BRCA2* families, with any one accounting only for a small proportion of such families. Second, in moderate-size families with a mixture of cases diagnosed at early and late ages, chance familial clustering of

cases may confound linkage-based approaches. Finally, penetrances of additional breast cancer susceptibility genes are likely to be lower than those associated with *BRCA1* and *BRCA2*.[9] Thus, analysis of a large family series with stringent selection criteria is required to achieve sufficient statistical power for unambiguous localization of novel susceptibility loci and meaningful evaluation of candidate genomic regions. To surmount these obstacles, our international collaborative group [Breast Cancer Linkage Consortium (BCLC)] has accrued, and continues to accrue, a collection of families appropriate to address the problem.

Recently, Kainu et al.[10] reported evidence for a novel breast cancer susceptibility locus on chromosome 13q21. They studied 77 families with multiple cases of breast cancer from Finland, Sweden, and Iceland in which no germline *BRCA1* or *BRCA2* mutations had been identified. Families were not specifically selected for early onset disease, nor were they excluded if one or more cases of ovarian cancer were present.

Initial analysis by comparative genomic hybridization (CGH) of tumors from 23 of these families and 14 others not analyzed further by linkage identified loss of 13q21–31 as a frequent and early event. Consistent loss of 13q21 in all five tumors from one family delineated a minimal region of haplotype sharing in these individuals as the target locus for a susceptibility gene. However, no evidence was presented for specific loss of the wild-type allele in these tumors, as would be expected for the underlying genetic model (inactivation of a tumor suppressor gene).

Genetic linkage analysis using 23 microsatellite markers from this region revealed supportive evidence of linkage to breast cancer. A maximum multipoint HLOD of 3.46 was found at marker D13S1308, with an estimated 65% of families linked. This marker lies ≈25 cM distal to *BRCA2* on chromosome 13q. Simulation studies to account for the possible confounding of linkage results by the proximity of these loci indicated that the linkage was unlikely to be the result of unidentified *BRCA2* mutations in a subset of families. However, the evidence for linkage was confined to a single pair of tightly linked markers (D13S1308/D13S1296) in this region, with linkage evidence dropping off quite rapidly surrounding this peak; indeed markers flanking a 2.1-cM region surrounding this peak yielded negative two-point LOD scores at recombination fractions up to 20%.

We present results from our attempt to confirm this linkage result through analysis of our series of 128 breast cancer families. In the remainder of this article, we refer to this locus as '*BRCA3*,' the quotation marks serving to emphasize the uncertainty regarding the existence and location of one or more such susceptibility loci.

TABLE 1

Summary of the families used in the 13q21 analysis

Age of diagnosis	Number of breast cancer cases in family				
	<3	3	4	5	>5
<50 years	51	48	19	5	5
<60 years	0	58	39	14	17
All cases:	0	26	36	25	41
Cases sampled/ genotyped	26	68	20	9	5

METHODS

Families.

Families were ascertained from cancer genetics or oncology centers in Europe (United Kingdom, Germany, Spain, Netherlands, France, and Israel), the United States, Australia, and Canada. One family was from Mexico. All families were Caucasian except the Mexican family that was of mixed European–Amerindian descent. Only families in which at least three women were diagnosed with breast cancer under age 60 years were eligible for the study. We excluded families in which cases of either ovarian cancer or male breast cancer were observed, because these phenotypes are strong predictors of *BRCA1* or *BRCA2* mutations.[1] Within these 128 families a total of 650 women were affected with breast cancer (median 5 per family); 56% of these cases were diagnosed under age 50. Samples from 409 affected individuals and 293 unaffected relatives were available for genotyping. Table 1 shows the characteristics of the families in more detail.

Entries are the number of families with the specified number of breast cancer cases of the indicated diagnostic criteria and sample availability.

Exclusion of *BRCA1* and *BRCA2*.

At least one breast cancer case from each family was screened for mutations in *BRCA1* and *BRCA2*, including all coding exons and splice junctions; in general, the sampled case with the youngest age at diagnosis was screened. This screening was performed using a variety of methods, including heteroduplex analysis (HDA), conformation sensitive gel electrophoresis (CSGE), and direct sequencing. Families from The Netherlands were also screened for the large genomic rearrangements that are known Dutch founder mutations, as these would not be detected by standard PCR-based screening methods. Other families were also tested for population-specific mutations, where

TABLE 2

Summary of markers used in the analysis

Marker	Map position, cM	Centers typed	Multipoint LOD score	
			Homogeneity	Heterogeneity (alpha = 0.65)
S1444	23.3	I	—	—
S1700	23.5	I	—	—
S260	23.7	S,L,I	−40.65	−14.25
S171	25.1	S	−36.07	−13.04
S1493	25.8	I	−33.83	−12.42
S267	26.9	S	−30.79	−11.51
S1293	26.9	I	−30.79	−11.51
S153	45.6	S	−32.54	−9.35
S788	45.6	I	−32.54	−9.35
S1317	51.0	L	−33.88	−10.35
S1262	51.0	I	−33.88	−10.35
S1308	52.6	S,L,I	−38.00	−11.03
S1296	52.6	I,L	−37.64	−10.93
S1291	53.2	L	−35.00	−10.06
S800	55.3	I	—	—
S166	55.3	S	—	—

I - IARC; S – ICR, Sutton; L – Leiden University. Based on published marker locations from Marshfield Medical Research Foundation (http://research.marshfieldclinic.org/genetics/). Note that *BRCA2* is at position 24.8 on this map.

appropriate. Overall, we estimate that, taken together, these methods have an average sensitivity of 0.70.[1,11]

Genotyping.

Genotyping was carried out at the International Agency for Research on Cancer (62 families), Institute of Cancer Research (49 families), and University of Leiden (17 families). Genotypes were generated for 16 microsatellite markers within a 32-cM region of chromosome 13q21 spanning both *BRCA2* and the putative '*BRCA3*' locus (see Table 2). Not all centers genotyped all markers; Table 2 gives details on which

loci were genotyped at each center. Microsatellite repeats were amplified from peripheral blood lymphocyte genomic DNA by standard methods using published primer sequences (The Genome Database, http://gdbwww.gdb.org/). PCR conditions were specific to each genotyping center, as was fragment analysis. Internal consistency of allele sizing was achieved at each center by incorporating samples with known allele sizes on each gel. A common DNA sample (CEPH-1347-02) was typed to ensure consistency of allele sizing between centers. Allele frequencies were calculated separately for each center from the pedigree genotypes by using DOWNFREQ software, Version 1.1 (available through http://linkage.rockefeller.edu/soft/).

Statistical Analysis.

We performed standard parametric linkage analyses, essentially identical to our previous analyses of linkage in breast cancer families (e.g., refs. 1, 5, and 8) and to the analysis conducted by Kainu et al.[10] These analyses assume the model of susceptibility to breast cancer based on the segregation analysis of Claus et al.[9] Under this model, susceptibility to breast cancer is conferred by a dominant allele with population frequency of 0.003. The risk of breast cancer by age 80 is assumed to be 0.80 in carriers and 0.08 in noncarriers. Risks are modeled in seven age categories (<30, 30–39, 40–49, 50–59, 60–69, 70–79, and 80+) as described in Easton et al.[12]

Multipoint linkage analyses were carried out using the programs GENEHUNTER (V. 2.0-B; ref. 13), VITESSE,[14] and FASTLINK.[15] GENEHUNTER was used where possible because it can analyze large numbers of polymorphic loci simultaneously and hence all of the markers we used could be incorporated into a single analysis. However, 33 families were too large to be accommodated by GENEHUNTER without discarding informative individuals. For these families we computed multipoint LOD scores by using either VITESSE (29 families) or FASTLINK (four families with multiple founders). The analyses assumed the intermarker distances as shown in Table 2.

We used the multipoint LOD scores for each family to compute heterogeneity LOD scores, using the standard admixture model, and hence estimated the proportion of families (α) linked to the putative '*BRCA3*' locus by maximizing the heterogeneity LOD score. A 95% confidence interval for αwas derived by computing the values of the heterogeneity LOD score that were within 0.83 (corresponding to a *Z* value of 1.96) of its maximum value. Ninety-nine percent confidence intervals were also computed.

Because the putative '*BRCA3*' locus on 13q21 is linked to *BRCA2*, we performed a

further analysis to allow for the possibility that preferential selection for families unlinked to *BRCA2* may have biased the results against linkage at '*BRCA3*.' In this analysis, we computed multipoint heterogeneity LOD scores at the candidate '*BRCA3*' locus, conditional on the LOD scores at *BRCA1* and *BRCA2*, according to the formula:

$$\text{LOD}(\theta_3) = \log_{10}\left[\frac{\alpha_1(1-\mu_1)10^{\text{LOD}_1(\theta_1)}+\alpha_2(1-\mu_2)10^{\text{LOD}_{2:3}(\theta_2)}+\alpha_3 10^{\text{LOD}_{2:3}(\theta_3)}+1-\alpha_1-\alpha_2-\alpha_3}{\alpha_1(1-\mu_1)10^{\text{LOD}_1(\theta_1)}+\alpha_2(1-\mu_2)10^{\text{LOD}_2(\theta_2)}+\alpha_3 10^{\text{LOD}_2(\theta_3)}+1-\alpha_1-\alpha_2-\alpha_3}\right]$$

In this formula α_1, α_2, and α_3 are the proportions of families meeting the eligibility criteria that are linked to *BRCA1*, *BRCA2* and '*BRCA3*,' respectively, and μ is the sensitivity of *BRCA1*/2 mutation screening. For the purposes of these analyses, α_1 and α_2 were set to 0.15 and μ to 0.7. $\text{LOD}_1(\theta_1)$ and $\text{LOD}_2(\theta_2)$ are the LOD scores at *BRCA1* and *BRCA2*, respectively, whereas $\text{LOD}_{2:3}(\theta_2)$ and $\text{LOD}_{2:3}(\theta_3)$ are the LOD scores at *BRCA2* and '*BRCA3*,' respectively, based on markers typed at both loci; $\text{LOD}_2(\theta_3)$ is the LOD score for '*BRCA3*' calculated using only markers at *BRCA2*. This calculated LOD score is the likelihood for the linkage data at '*BRCA3*' conditional on the existing linkage and mutation evidence at *BRCA1* and *BRCA2*, and hence corrects (albeit conservatively) for any bias in the '*BRCA3*' evidence produced by exclusion of families linked to *BRCA2*.

RESULTS

Total LOD scores were strongly negative throughout the 8-cM interval between D13S153 and D13S1291 (Table 2 and Fig. 1). At the location of '*BRCA3*' estimated by Kainu et al.,[10] D13S1308, the total LOD score was –38.00. Based on the admixture model, the estimated proportion of linked families (α) was 0, with an upper 95% confidence limit of 0.13. The estimated α was also zero for all possible positions in the interval D13S153-D13S1291. Of the 128 families, only four had a multipoint LOD score of greater than 0.5 at D13S1308, the highest of which was 0.67 (one additional family achieved a LOD score of 1.55 at a more distal marker, D13S800). Twelve families achieved LOD scores less than –1 at D13S1308.

We reanalyzed the data conditioning on the genotyping data at *BRCA1* and *BRCA2*. In

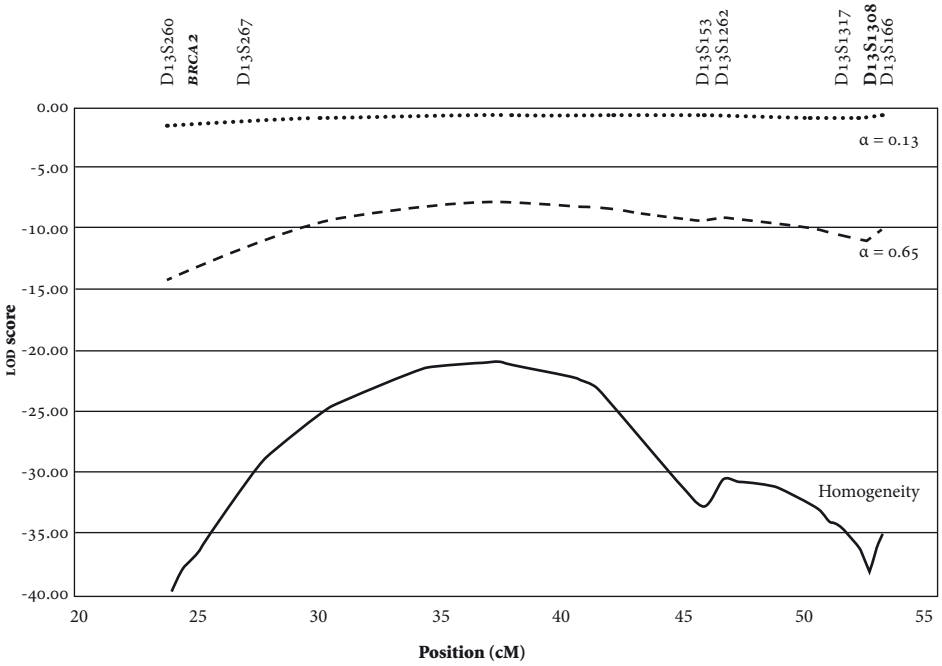


Figure 1. Multipoint LOD scores for the 128 families analyzed are shown graphically. The solid line represents scores obtained under the assumption of homogeneity; the dashed line assumes the proportion of linked families (α) to be 65%, as estimated by Kainu et al.[10]; and the dotted line represents the 95% upper confidence interval (α = 0.13).

this analysis the total LOD score was −25.08. In the heterogeneity analysis based on these conditional LOD scores, the estimated proportion of families linked to '*BRCA3*' was again 0, with an upper 95% confidence limit of 0.18. In the 95 families that could be analyzed with GENEHUNTER, we also analyzed the data by using the nonparametric method[13] to evaluate haplotype sharing among affected women. Again, no significant evidence of linkage was found (data not shown).

Discussion

Our results clearly conflict with those reported by Kainu et al.[10] Using a set of multiple case female site-specific breast cancer families analyzed for a similar set of markers within the candidate region and subjected to comparable statistical analysis, we found no evidence of linkage to 13q21. The proportion of linked families (65%) reported by Kainu et al.[10] is excluded with a high degree of statistical significance (the heterogeneity LOD score at α= 0.65 was −11.03 in our dataset). This is true even after

a conservative correction for possible bias due to potential exclusion of families linked at the *BRCA2* locus (conditional LOD at α= 0.65 was −7.64). In addition, under both unconditional and conditional analyses, the estimated proportion of linked families was 0, with upper 95% confidence intervals of 13% and 18%, respectively, indicating that if there is a susceptibility locus on 13q, it is likely to account for only a minority of breast cancer families. The paper of Kainu et al.[10] did not provide confidence limits on their estimated proportion of linked families. However, based on their LOD scores given under homogeneity and 65% heterogeneity, and assuming confidence intervals that are symmetrical about the best estimate, we have estimate a lower 95% confidence limit for α of 0.31. Thus the 95% confidence limits for the two studies do not overlap. Moreover, even when using a more stringent criteria of 99%, the upper confidence limit for our estimated proportion of linked families is 0.19 for the unconditional analysis and 0.26 for the analysis conditioning on *BRCA2* markers, further indicating a minor role, if any, for this locus.

There were some differences in selection criteria between the two studies. Our study was restricted to families in which at least three cases of breast cancer were diagnosed below age 60, whereas Kainu et al.[10] included families with three cases diagnosed at any age. Thus, our families may be more heavily selected for genes conferring high risk. It is perhaps noteworthy that the initial hypothesis-generating family analyzed by comparative genomic hybridization (CGH) in Kainu et al.[10] would not have qualified for our study because only two of the five cases were diagnosed under age 60. However, in the subset of 51 families with less than three cases diagnosed under age 50 (Table 1), there is also considerable evidence against linkage to this locus (multipoint LOD = −8.06; HLOD = 0; upper 95% CI for α = 24%; HLOD for α of 65% = −3.57). Thus it is unlikely that difference in age criteria can explain the differences in results between the two studies.

An additional difference in selection criteria was exclusion of families with any cases of ovarian cancer in our series, given the close association of this disease with *BRCA1* and *BRCA2*. Although no *BRCA2* mutations were identified in the family set of Kainu et al., the combination of detection methods applied to screening families have detection sensitivities of ≈0.70.[1,11] Thus, although simulated linkage results allowing for up to 25% of the families in the dataset of Kainu et al.[10] being due to undetected *BRCA2* mutations only exceeded the observed maximal lod score in 1 of 3,000 replicates, it is not known to what extent the seven families with ovarian cancer contributed to the observed overall LOD score.

The families in our study were drawn from Western Europe, or in descendent popu-

lations in North America and Australia, whereas the families studied by Kainu et al.[10] were from the Nordic countries. Although we have not specifically examined the ethnic origins of each family in our set, it is anticipated that the set of families from the United States and Canada (n = 43) are more ethnically heterogeneous, although most, if not all, are of Western European origin. Only a small minority of all of the families in our set are likely to be of Scandinavian origin, most notably the families ascertained in Minnesota, Seattle, and other parts of the Midwest, which have a high concentration of families descendent from emigrants of Sweden and Norway. One might speculate that the difference in the results observed is due to a population specific founder effect i.e., an excess of some specific mutation in '*BRCA3*' in the Nordic populations.

We believe this to be unlikely. The different Nordic populations have different population histories and do not originate from a single small founder population. Although closely related, the Swedish, Icelandic, and (to a lesser extent) Finnish populations are also genetically similar to English and Dutch populations.[17] If the observed linkage were due to a susceptibility allele that had reached a high frequency in the Swedish and Finnish populations, this allele would also be expected to occur at a detectable frequency in the British and Dutch families. On the other hand, if the linkage is the result of several different mutations in the candidate '*BRCA3*' gene, the expectation would be that (as in the case of *BRCA1* and *BRCA2*) mutations would also occur in the British, Dutch, and other populations, albeit the set of mutations might be different. Under either model, we would have expected to observe similar evidence of linkage in our families. Indeed, even when the prevalence of a population specific founder mutation has led to a specific susceptibility gene accounting for the majority of families of a hereditary cancer syndrome [e.g., *BRCA2* in the Icelandic population accounting for 61.4% of breast cancer families[18]; >50% of hereditary nonpolyposis colon cancer (HNPCC) families in the Finnish population attributable to two specific MLH1 mutations[19]], these same genes account for a substantial fraction of families with the same cancer syndrome in other populations (breast cancer reviewed in[20]; HNPCC[19]).

We conclude therefore that any contribution of a locus at chromosome 13q21 to familial breast cancer is likely to be small in breast cancer families of European origin. Further linkage studies in large series of multiple case families, or targeted association studies in large series of breast cancer cases and controls, will be needed to identify remaining genes underlying familial aggregation of the disease.

ACKNOWLEDGEMENTS

We wish to acknowledge the database management support of C. Bonnardel (International Agency for Research on Cancer). We thank Prof. David Anderson (ret.) for having originally collected some of the families used in this study. This study is funded in part by the Cancer Research Campaign (CRC); the Association for International Cancer Research (AICR); the Dutch Cancer Society, The Swiss Bridge Foundation; KConFab is supported by the National Health and Medical Research Council (Australia) and the National Breast Cancer Foundation (Australia). National Institutes of Health Grants CA69446, CA69467, CA69417, CA69631, CA69398, and CA69638 support the Cooperative Family Resource for Breast Cancer Studies (CFRBCS), which is coordinated by Daniela Seminara, Epidemiology Branch, National Cancer Institute, Bethesda, MD. CMP is supported by US Army Grant DAMD17-00-1-0478.

REFERENCE LIST

1. **Ford, D., Easton, D. F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D. T., Weber, B., Lenoir, G., Chang-Claude, J., et al.** *Am. J. Hum. Genet. 62, 334-345. (1998)*
2. **Serova, O., Mazoyer, S., Puget, N., Dubois, V., Tonin, P., Shugart, Y., Goldgar, D., Narod, S. A., Lynch, H. T. & Lenoir, G. M.** *Am. J. Hum. Genet. 60, 486-495. (1997)*
3. **Peto, J., Collins, N., Barfoot, R., Seal, S., Warren, W., Rahman, N., Easton, D. F., Evans, C., Deacon, J. & Stratton, M. R.** *J. Natl. Cancer Inst. 91, 943-949. (1999)*
4. **Anglian Breast Cancer Study Group** *Br. J. Cancer 83, 1301-1308. (2000)*
5. **Shugart, Y. Y., Cour, C., Renard, H., Lenoir, G., Goldgar, D., Teare, D., Easton, D., Rahman, N., Gusterton, R., Seal, S., et al.** *J. Med. Genet. 36, 720-721. (1999)*
6. **Kerangueven, F., Essioux, L., Dib, A., Noguchi, T., Allione, F., Geneix, J., Longy, M., Lidereau, R., Eisinger, F., Pebusque, M. J., et al.** *Oncogene 10, 1023-1026. (1995)*
7. **Seitz, S., Rohde, K., Bender, E., Nothanagel, A., Kolble, K., Schlag, P. M. & Scherneck, S.** *Oncogene 14, 741-743. (1997)*
8. **Rahman, N., Teare, M. D., Seal, S., Renard, H., Mangion, J., Cour, C., Thompson, D., Shugart, Y., Eccles, D., Devilee, P., et al.** *Oncogene 19, 4170-4173. (2000)*
9. **Claus, E. B., Risch, N. & Thompson, W. D.** *Am. J. Hum. Genet. 48, 232-242. (1991)*
10. **Kainu, T., Jou, S.-H., Desper, R., Schäffer, A. A., Gillanders, E., Rozenblum, E., Freas-Lutz, D., Weaver, D., Stephan, D., Bailey-Wilson, J., et al.** *Proc. Natl. Acad. Sci. USA 97, 9603-9608. (2000)*
11. **Eng, C., Brody, L. C., Wagner, T. M., Devilee, P., Vijg, J., Szabo, C., Tavtigian, S. V., Nathanson, K. L., Ostrander, E. & Frank, T. S.** *J. Med. Genet 38, 824-833. (2001)*
12. **Easton, D. F., Bishop, D. T., Ford, D., Crockford, G. P. & the Breast Cancer Linkage Consortium** *Am. J. Hum. Genet. 52, 678-701. (1993)*
13. **Kruglyak, L., Daly, M., Reeve-Daly, M. P. & Lander, E. S.** *Am. J. Hum. Genet. 58, 1347-1363. (1995)*
14. **O'Connell, J. R. & Weeks, D. E.** *Nat. Genet. 11, 402-408. (1995)*
15. **Cottingham, R.W., Jr., Idury, R. M. & Schaffer, A. A.** *Am. J. Hum. Genet. 53, 252-263. (1993)*

16. **Terwilliger, J. D., Shannon, W. D., Lathrop, G. M., Nolan, J. P., Goldin, L. R., Chase, G. A. & Weeks, D. E.** *Am. J. Hum. Genet. 61, 430-438. (1997)*
17. **Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A.** *T*he History and Geography of Human Genes *Princeton Univ. Press, Princeton, NJ (1994)*
18. **Thorlacius, S., Olafsdottir, G., Tryggvadottir, L., Neuhausen, S., Jonasson, J. G., Tavtigian, S. V., Tulinius, H., Ogmundsdotti, r. H. M. & Eyfjord, J. E.** *Nat. Genet. 13, 117-119. (1996)*
19. **Holmberg, M., Kristo, P., Chadwicks, R. B., Mecklin, J. P., Jarvinen, H., de la Chapelle, A., Nystrom-Lahti, M. & Peltomaki, P.** *Hum. Mut. 11, 482. (1998)*
20. **Szabo, C. I. & King, M. C.** *Am. J. Hum. Genet. 60, 1013-1020. (1997)*

CHAPTER 4

MOLECULAR PROFILING

4.1. CHARACTERIZATION OF FAMILIAL NON-BRCA1/2 BREAST TUMORS BY LOSS OF HETEROZYGOSITY AND IMMUNOPHENOTYPING

R.A. Oldenburg[1,2], K. Kroeze-Jansema[1], H. Meijers-Heijboer[2], C.J. van Asperen[1], N. Hoogerbrugge[3], I. van Leeuwen[4], H.F.A. Vasen[4], A.M Cleton-Jansen[5], J. Kraan[1], J.J. Houwing-Duistermaat[6], H. Morreau[5], C.J. Cornelisse[5], P. Devilee[1,5]

Clin Cancer Res. 2006 Mar 15;12(6):1693-700.

[1] *Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden*
[2] *Dept. of Clinical Genetics, Erasmus Medical center, Rotterdam*
[3] *Dept. of Human Genetics, University Medical Center Nijmegen*
[4] *Netherlands Foundation for the Detection of Hereditary Tumors, Leiden*
[5] *Dept. of Pathology, Leiden University Medical Center, Leiden*
[6] *Dept. of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden*

ABSTRACT

Purpose

Since the identification of *BRCA1* and *BRCA2*, there has been no major breast cancer susceptibility gene discovered by linkage analysis in breast cancer families. This has been attributed to the heterogeneous genetic basis for the families under study. Recent studies have indicated that breast tumors arising in women carrying a *BRCA1* mutation have distinct histopathological, immunophenotypic and genetic features. To a lesser extent, this is also true for breast tumors from *BRCA2* carriers. This indicates that it might be possible to decrease the genetic heterogeneity among families in which *BRCA1* and *BRCA2* have been excluded with high certainty (*BRCAX* families) if distinct subgroups of *BRCAX*-related breast tumors could be identified.

Experimental Design

Loss of heterozygosity analysis with at least one marker per chromosomal arm (65

markers) was used to characterize 100 breast tumors derived from 92 patients from 42 selected *BRCAX* families. In addition, the immunophenotype of 10 markers was compared to that of 31 *BRCA1*- and 21 *BRCA2*-related breast tumors.

Results and conclusions

The *BRCAX*-related tumors were characterized by more frequent LOH at 22q relative to sporadic breast cancer ($P<0.02$), and differed significantly from *BRCA1*- and *BRCA2*-related tumors in their positivity for Bcl2. However, cluster analyses of the combined data (LOH and immunohistochemistry) did not result in subgroups that would allow meaningful sub classification of the families. On chromosomes 2, 3, 6, 12, 13, 21 and 22 we found markers at which LOH occurred significantly more frequent among the tumors from patients belonging to a single family than expected on the basis of overall LOH-frequencies. Nonetheless, linkage analysis with markers for the corresponding regions on chromosomes 12, 21 and 22 did not reveal significant LOD's

INTRODUCTION

A positive family history remains one of the most important risk factors for breast cancer, with first-degree relatives of patients having an approximately 2-fold elevated risk. About 15% of all patients have a first-degree relative with breast cancer, and although germ-line mutations in *BRCA1* and *BRCA2* account for a substantial proportion of these cases,[1] these mutations explain only 20-25% of the overall excess familial risk.[2,3] Mutations in other genes such as *TP53* and *PTEN* are involved in rare multi-cancer syndromes and contribute very little to this risk. Mutations in *BRCA1* and *BRCA2* are strongly associated with families with at least 4 cases of breast cancer diagnosed before the age of 60 and one or more cases of ovarian cancer or male breast cancer.[1] However, in families with 4 or 5 cases of breast cancer, and no ovarian or male breast cancer cases, *BRCA1* and *BRCA2* mutations were significantly less frequent. Because such a familial clustering is unlikely to have occurred by chance, this has been taken as evidence that other breast cancer susceptibility genes must exist.[4] After the identification of *BRCA1* and *BRCA2*, several chromosomal regions have been implicated by linkage analysis to harbor a breast cancer susceptibility gene. In particular, linkage has been found with markers for 8p12-22 and 13q21,[5,6] but although mutations in *BRCA1* and *BRCA2* were excluded, these studies comprised either small or heterogeneous groups of families. Accordingly, these linkage results have proven difficult to replicate by others in independently collected sets of families.[7,8] It has

been argued that the inability to detect genetic linkage is largely due to a heterogeneous genetic basis for the families under study.[4]

It is now well established that breast tumors arising in women carrying a *BRCA1* mutation have distinct histopathological, immunophenotypic and genetic features.[9-14] This is also true for breast tumors from *BRCA2* carriers, although to a lesser extent. These findings indicate that it might be possible to subgroup the breast tumors derived from patients from families in which *BRCA1* and *BRCA2* have been excluded with high certainty (from now on called *BRCAX* families). This could possibly decrease the genetic heterogeneity within this group of families, and thereby increase the statistical power to detect linkage. Here, we used loss of heterozygosity and immunohistochemical analyses to characterize 100 breast tumors derived from *BRCAX* families. The *BRCAX*-related tumors were characterized by more frequent LOH at 22q relative to sporadic breast cancer, and differed significantly from *BRCA1*- and *BRCA2*-related tumors in their positivity for Bcl2. However, cluster analyses of the combined data (LOH and immunohistochemistry) did not result in subgroups that would allow useful sub classification of the families.

MATERIALS AND METHODS

Family selection

The families were ascertained through the Clinical Genetic Centers in Leiden, Rotterdam, and Nijmegen, as well as through the Netherlands Foundation for the Detection of Hereditary Tumors (STOET). Families were eligible if there were at least three cases of breast cancer diagnosed before the age of 60 from whom genotypes could be determined (N=216) or inferred (N=20). Families with cases of ovarian cancer or male breast cancer were excluded, and occurrences of other types of cancer were ignored. Pathological reports or medical reports were retrieved where available. Blood samples and paraffin-embedded tumor tissues were collected after obtaining written informed consent. The institutional ethical committees of all of the hospitals involved approved this study.

In total we collected 100 breast tumors derived from 92 patients from 42 selected *BRCAX* families. Nine of these 100 breast tumors belong to 8 *CHEK2**1100delC mutation carriers.[15] Although the families under study were not tested for mutations in other breast cancer susceptibility genes (such as p53, E-cadherin and PTEN), they did not show the phenotypic characteristic belonging to these cancer syndromes. We also collected 40 paraffin-embedded tumor samples from sporadic breast cancer

cases unselected for family history or age, and from 31 *BRCA1*-mutation carriers and 21 *BRCA2* mutation carriers.

BRCA1 and BRCA2 mutation Testing

In each family, the youngest breast cancer patient from whom a blood sample was available was tested for mutations in the *BRCA1* and *BRCA2* genes (and for many families the next youngest as well). The joint Clinical Genetic Centers applied a variety of methodologies. The largest central exons (exon 11 in *BRCA1* and *BRCA2,* exon 10 of *BRCA2*) were scanned by protein truncation tests.[16,17] The small exons were scanned for mutations by denaturating gradient gel electrophoreses (DGGE) or direct sequencing. All of the laboratories specifically assayed the presence of large founder deletions in *BRCA1* by deletion junction-PCR.[18] The entire coding sequences of *BRCA1* and *BRCA2* were investigated by conformation-sensitive gel electrophoresis (CSGE) in families that were incompletely scanned at the time of ascertainment.[19] Since 2002, each center offers full sequence analysis and DGGE covering the entire coding regions of both genes, and Multiple Ligation-dependent Probe Amplification (MLPA) to detect large deletions/duplications in *BRCA1*.[20]

Histology

Paraffin embedded tumor tissues were obtained and the breast tumors were histologically classified according to the WHO criteria.[21] An expert pathologist (H. Morreau, MD) assessed type of invasive cancer, histological grade, presence of *in situ* component and the presence of lymphocyte infiltrate. Age of the patient at time of diagnosis was available from pathological and medical reports.

LOH Analysis

On the respective H&E stained sections the areas of highest tumor density were selected. Four to six tissue cores (0.6 mm in diameter, Beecher Instruments, Silver Spring, MD) were punched from the designated area using a biopsy needle. DNA was isolated from these punches as described previously.[17] These punches generally contain >50% tumor cells. Normal DNA was isolated from the blood samples. For the LOH-analysis we used 65 fluorescence-labeled microsatellite markers selected from Weber Screening Set 6 and covering all chromosome arms.[22] Selection criteria were allele product-sizes below 250 bp (because PCR success rates with DNA isolated from paraffin-embedded material drops sharply with larger amplimers) and position in the telomeric half of a chromosome arm (because this will also detect mitotic recom-

bination events[23,24]). The PCR-products were visualized on an ABI PRISM 3700 DNA Analyzer (Applied Biosystems) and analyzed with the Genotyper software version 3.7 NT (Applied Biosystems). The sporadic breast tumors were analyzed only for the six different markers on chromosome 22 and marker D11S15901 on chromosome 11, and the BRCA1-related breast tumors were analysed only for D4S1562 and D5S1471. Allelic imbalance was defined as the ratio of allele intensities in the normal versus the tumor DNA. An AIF (Allelic imbalance factor) of 1.70 or above was scored as 'LOH'.[25] A technical limitation in the interpretation of the allelic imbalance factor is the possible contamination of tumor DNA with non-malignant DNA. Although a biopsy needle to punch tissue cores does not prevent contamination with non-malignant cells, in 80% of the tumor DNA samples we detected at least one AIF >5.0, which is only achievable when relatively high proportions of tumor cells are present in the sample.[26]

Tissue-microarray (TMA)

Breast cancer tissue microarrays were prepared as described previously.[27] From each case three tissue cores were assembled in the TMA. In total 4 TMA blocks were constructed. Three blocks with *BRCAX* samples and one block with tumors samples from *BRCA1* and *BRCA2* mutation carriers.

Immunohistochemistry scoring

Immunohistochemical staining was performed by the labeled Streptoavidin biotin method (DAKO, Glostrup, Denmark) with a heat-induced antigen retrieval step.
One pathologist (HM) and one researcher (RAO) evaluated the immunohistochemical staining results. The percentage of stained nuclei, independent of the intensity, was scored for p53, ER, PR, and Cyclin D1. In the same way, the percentage of cells with cytoplasmic staining was scored for Bcl2. Her2/Neu was assessed in accordance with the DAKO HercepTest guidelines with a score of ≤ 1 considered negative. Cytokeratin 5/6, Cytokeratin 7 and Cytokeratin 19 were scored according to the presence or absence of membranous expression in the invasive component. The Chek2 staining pattern was scored as described earlier.[15] For p53 we used four different categories on the basis of any level of nuclear staining; 1) negative, 2) <25%, 3) 25-75% and 4) >75% positive nuclei. For ER and PR, a case was considered positive when ≥ 10 % of the nuclei stained above background. For Cyclin D1 the cut off limit was 30%. For bcl2 the cut off limit was 70%.

Statistics

Proportions were compared using chi-square statistics. Familial aggregation of LOH status at a marker was tested using a score statistic.[28] This statistic tests for the presence of an additive genetic effect. For this analysis AIF's between 1.3 and 1.7 were regarded as missing. Empirical p-values were computed by permutation of the LOH status among relatives of the same family.

Cluster Analysis

For the hierarchical cluster analysis we used the software programs Cluster and TreeView. The data was normalized, mean centered and average linkage clustering was applied. We renumbered the LOH data of 100 tumors as follows; AIF's>1.70 were scored as '1' (LOH), AIF's between 1.0 and 1.29 (retention of heterozygosity) as '–1', AIFs between 1.3 and 1.7 as '0', and homozygotes as missing. The immunohistochemical data for the different markers was scored as '1' when considered positive and '-1' when considered negative.

Linkage analysis

Genotypes were generated for 19 microsatellite markers on chromosome 12, 5 on chromosome 21, and 12 on chromosome 22. The markers were derived from Linkage Mapping Set version 2 (Applied Biosystems), and amplified from peripheral blood lymphocyte genomic DNA by standard PCR methods. DNA from CEPH 1347-02 was typed as reference to ensure consistency of allele sizing. Allele frequencies for parametric linkage analyses were calculated based on one randomly chosen individual from each family. Multipoint linkage analyses were carried out using the program GENEHUNTER version 2.1-B.[29] We used a model in which susceptibility to breast cancer is conferred by a dominant allele with a reduced penetrance and a population frequency of 0.003.[30,31] The risk of breast cancer by age 80 was assumed to be 0.85 in carriers and 0.096 in non-carriers. Risks are modeled in seven age categories (<30, 30-39, 40-49, 50-59, 60-69, 70-79, and 80+) as described.[31] We used the multipoint LOD-scores for each family to compute heterogeneity LOD scores, using the standard admixture model, and hence estimated the proportion of families (α) linked to the putative 'BRCAX' locus by maximizing the heterogeneity LOD score. Non-parametric linkage analyses were carried out by the program MERLIN version 0.9.12b 32.

TABLE 1

Histologic description of the different groups analyzed

%	BRCAx (n = 100)	Control (n = 40)	BRCA1 (n = 31)	BRCA2 (n = 21)
Ductal carcinoma in situ	1	0		4.8
Ductal carcinoma	81	85	100	95.2
Lobular carcinoma in situ	1	0		
Lobular carcinoma	10	5		
Colloid carcinoma	1	2.5		
Other	1	5		
Unknown	5	2.5		
Grade 1	20.7	29.4	0	5
Grade 2	50	41.2	12.9	45
Grade 3	29.3	29.4	87.1	50
	P < 0.0002*	P < 0.02**	P < .0002***	

DCIS, ductal carcinoma in situ; LCIS, lobular carcinoma in situ

*, grade of BRCA1 tumors versus grade of BRCAx tumors; **, BRCA1 versus BRCA2 tumors;

***, BRCA1 versus control tumors.

RESULTS

Histology

A total of 100 paraffin-embedded breast tumor samples could be retrieved from 92 patients from 42 early onset breast cancer families (*BRCAX*-families, defined as having at least 3 cases diagnosed before the age of 60, and no cases of ovarian or male breast cancer). We previously found 8 patients to carry the *CHEK2**1100delC mutation, representing 9 of these 100 breast tumors.[15]

The histological characteristics of this group of breast cancers, as compared to sporadic breast tumors (N=40) and the breast tumors from *BRCA1* (N=31) and *BRCA2* (N=21) mutation carriers are listed in Table 1. The most common histological type in all groups was infiltrating ductal carcinoma. Contrary to earlier suggestions,[33] lobular carcinoma was not significantly more often found in the *BRCAX* tumors relative to sporadic cancers. The *BRCA1* tumors were of higher grade than *BRCAX* tumors (P<0.001) and the *BRCA2* tumors (P= 0.01). Most of the *BRCAX* tumors were of grade II, and there was a trend towards grade being lower than that of *BRCA2* tumors (P=0.07), which is consistent with previously reported results.[9,11,33]

Immunohistochemistry

Three tissue microarray (TMA) blocks were constructed with 98 of the 100 *BRCAX* tumors, and one with 31 *BRCA1* and 21 *BRCA2*-related breast tumors. All tumors were stained with antibodies against ER, PR, p53, Bcl2, Her2/Neu, Cyclin D1, CHEK2, the basal cytokeratin 5/6 and the luminal cytokeratins 7 and 19, the immunohistochemical markers most commonly studied in *BRCA1*/2-associated breast carcinomas (Table 2). *BRCAX* tumors were significantly different from *BRCA1* tumors for ER ($p<0.001$), PR ($p=0.002$), Her2/Neu ($p=0.02$), Cyclin D1 ($p=0.02$), Bcl2 ($p<0.001$), and the basal CK5/6 ($p=0.0015$) staining. There were also significant differences between the *BRCA1* and *BRCA2* tumors for ER ($p=0.002$), Her2/Neu ($p=0.02$) and the basal CK5/6 ($p<0.001$) staining. *BRCAX* tumors differed significantly from both *BRCA1* and *BRCA2* tumors only for Bcl2 ($p<0.001$), while for CK5/6 this difference was borderline significant ($p=0.09$). As expected, the 9 tumors from *CHEK2*1100delC* carriers were significantly more often negative for CHEK2 staining than *BRCA1*, *BRCA2*, and *BRCAX* tumors. Interestingly they are also significantly more often negative for luminal CK19 staining than *BRCAX* ($p=0.0008$) and *BRCA1* ($p=0.006$) tumors.

We combined the results of the luminal marker (CK19) together with the basal marker (CK5/6) expression to subdivide the *BRCAX* breast tumors into four different cellular phenotypes: 'luminal' (only expression of the luminal marker), 'basal' (expression of the basal marker and no expression of the luminal marker), 'mixed' (expression of the basal marker and expression of the luminal markers) and 'null' (no expression of basal and luminal markers).[34] In this subdivision CK7 was not included, because of the high percentage of tumors that stained positive in all groups. The results demonstrate that a high proportion of *BRCAX* breast carcinomas express the mixed phenotype or have a pure luminal phenotype (Table 3). The *BRCA1* tumors are more often of the mixed phenotype compared with *BRCAX* tumors ($p= 0.0017$) and with *BRCA2* tumors ($p=0.0007$). No significant difference was seen between the *BRCA2*- and *BRCAX*-tumors. The *CHEK2*1100delC* related tumors showed a trend towards the null phenotype. Among the *BRCAX* tumors, the mixed tumors were more often positive for Her2/Neu relative to the luminal group ($p=0.02$), and the pure luminal tumors are more often grade III than the tumors with a null phenotype ($p=0.006$) (data not shown).

Genome-wide loss of heterozygosity (LOH)

The 100 BRCAX tumors were analyzed for LOH with 65 polymorphic markers representing all chromosomal arms. Of the potential 6,500 pair-wise normal/tumor com-

TABLE 2

Immunohistologic results

	BRCAx %	BRCA1 %	BRCA2 %	CHEK2 %	P
P53					
0	16.5	63	78.9	22.2	
<25%	62.6	3.7	10.5	77.8	
25-75%	9.9	7.4	5.3	0	
>75%	11	25.9	5.3	0	
ER					
Negative	33	85.7	41.2	44.4	0.0000009* 0.01†
Positive	67	14.3	58.8	55.6	0.002‡
PR					
Negative	43.5	78.6	61.1	33.3	0.002*
Positive	56.5	21.4	38.9	66.7	0.01†
Her2Neu					
Negative	74.5	96.1	72.2	77.8	0.02*
Positive	25.5	3.9	27.8	22.2	0.02‡
Cycline D1					
Negative	71.3	92.3	87.5	88.9	0.02*
Positive	28.7	7.7	12.5	11.1	
Bcl2					
Negative	39.1	88.9	94.1	33.3	0.000005* 0.0009†
Positive	60.9	11.1	5.9	66.7	0.00003⊙ 0.0009•
CHEK2					
Negative	13	22.3	16.7	66.7	0.04†
Positive	53.3	37	38.9	22.2	0.003#
Strong pos	33.7	40.7	44.4	11.1	0.02•
Cytokeratin 5/6					
Negative	54.3	19.2	76.5	55.6	0.0015* 0.04†
Positive	45.7	80.8	23.5	44.4	0.0002‡
Cytokeratin 7					
Negative	3.3	3.8	0	11.1	
Positive	96.7	96.2	100	88.9	
Cytokeratin 19					
Negative	27.9	11.1	33	66.7	0.0008†
Positive	72.1	88.9	67	33.3	0.006#

ER estrogen receptor, PR progesterone receptor,

* BRCA1 versus BRCAx tumors, † CHEK2 versus BRCA1 tumors, ‡ BRCA1 versus BRCA2 tumors,

⊙ BRCAx versus BRCA2 tumors, • CHEK2 versus BRCA2 tumors, # CHEK2 versus BRCAx tumors.

TABLE 3

Immunophenotype distribution based on the expression of the basal basal cytokeratin 5/6 and the luminal cytokeratin 19

%	Luminal	Basal	Mixed	Zero
BRCAx (n = 91)	35.2	8.8	36.3	19.8
BRCA1 (n = 27)	14.8	7.4	74.1	3.7
BRCA2 (n = 16)	43.8	6.3	18.8	31.3
CHEK2 (n = 9)	11.1	22.2	22.2	44.5
	P = 0.0017*		P = 0.00073**	

* BRCA1 versus BRCAx tumors

** BRCA1 versus BRCA2 tumors

parisons, 1,698 (26.1%) failed due to PCR problems of either the tumor DNA or normal DNA. Of the remaining 4,802, 1,220 (25.4%) were homozygous (not informative). Thus, in total 3,582 (55.1%) informative AIF's could be calculated. Using an AIF of 1.7 or greater as cut-off for LOH, the mean percentage of LOH among the markers was 30% (±6.3%), which is similar to the overall average LOH rate calculated from 151 published LOH studies of breast cancer.[35] LOH frequencies of 40% or greater were found at 1q41, 4p16, 11q22, 11q23.3, 16p13, 16q24, 17p12, 21q22, 22q11 and 22q13 (Figure 1), with the highest frequency found at D22S445 (59%). Whereas many of these chromosomal sites have also been highlighted in analyses of sporadic breast tumors, we did confirm that the percentage of LOH at D22S445 and D22S315 was significantly higher in BRCAX versus the sporadic breast tumors (respectively P<0.02 and P=0.035)(Figure 2). We also confirmed the high levels of allelic imbalance at 4q (7 of 12 informative cases) and 5q (4 of 9 informative cases) in *BRCA*1-related tumors.[36]

In 28 families we were able to assess LOH in at least 2 breast tumors from 2 patients. We tested whether there were loci at which LOH was found significantly more often within families than expected on the basis of overall LOH frequency at this locus in all our families. This was found for markers D2S125 (p=0.007), D3S2409 (p=0.045), D6S1552 (p=0.03), D12S2070 (p=0.02), D13S285 (p=0.02), D21S1255 (p<0.001) and D22S315(p=0.01). Of note, marker D22S445 did not show this familial clustering (p=0.35).

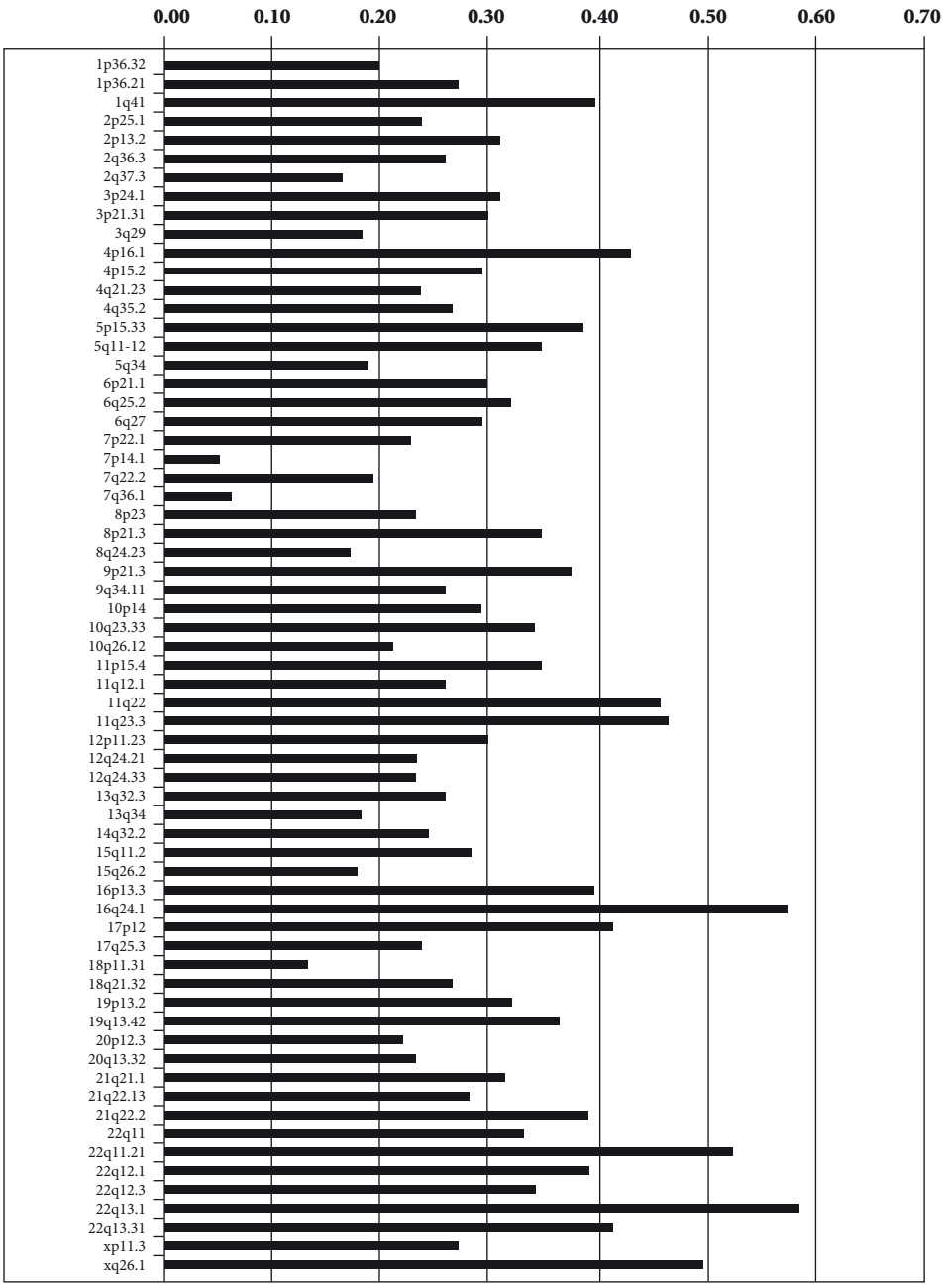


Figure 1. Percentages Loss of heterozygosity (LOH) for the different chromosomes.

The numbers above the graph represent the different chromosomes. A tumor was scored positive for LOH when having an AIF ≥ 1.70.

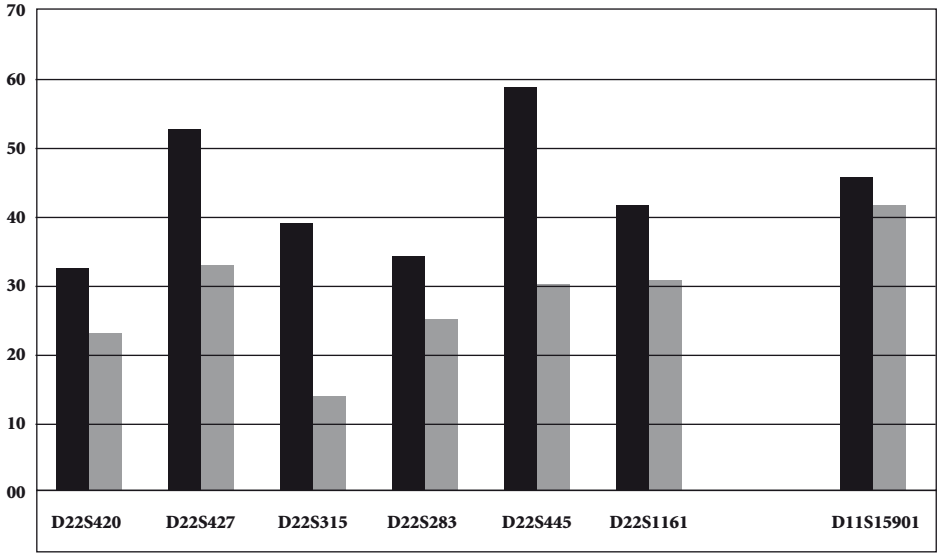


Figure 2

Percentages Loss of heterozygosity (LOH) found for the different microsatellite markers on chromosome 22 and 11. The black columns represent the BRCAx tumors and the grey columns the sporadic tumors.

Cluster analyses

We attempted to use the LOH data of 98 tumors in a hierarchical non-supervised clustering analysis by scoring AIF's>1.70 as '1', AIF's between 1.00 and 1.29 (retention of heterozygosity) as '–1', and AIF's between 1.30 and 1.70 and homozygotes as 'missing' in the software package 'Cluster'. Although the tumors were separated into two groups, these were not readily discernable on the basis of any single marker or combination of markers, nor did the tumors derived from the same family or the *CHEK2**1100delC carriers cluster together (data not shown). Adding the immunophenotyping and histological typing data did not resolve this.

Linkage analysis

We performed a linkage analysis in 55 families, complying with our selection criteria, for chromosomes with either a conspicuous LOH score (#22, at D22S445) or for which LOH showed significant familial clustering (#12, #21). For chromosomes 2, 3, 6 and 13 there were too few families for which linkage and LOH data could be combined to be statistically meaningful. The highest multipoint LOD score at chromosome 21 over all 55 families was -6.37 between markers D21S1256 and D21S1914. At the same locus, the non-parametric LOD (NPL) score was 1.72. Assuming hetero-

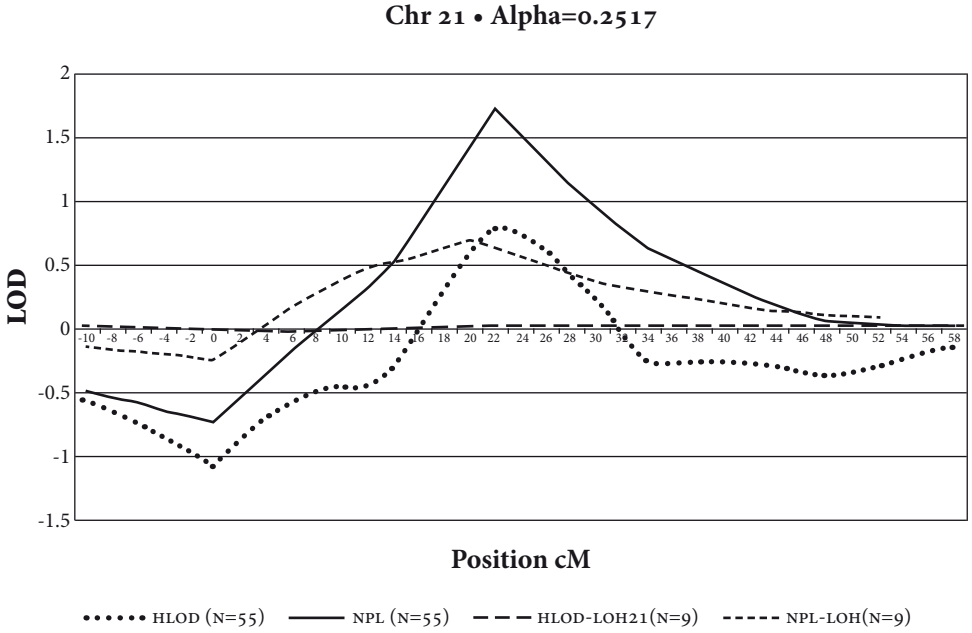


Figure 3

Heterogeneity logarithm of odds (HLOD) and nonparametric linkage scores (NPL) for chromosome 21; Alpha, the proportion of linked families calculated by the program GENEHUNTER; HLOD-LOH21, HLOD found for the 9 families in which the tumor of at least one patient showed LOH at marker D21S1255; NPL-LOH21, NPL-scores found for the 9 families in which the tumor of at least one patient showed LOH at marker D21S1255.

geneity, we found a non-significant heterogeneity LOD (HLOD) score of 0.80 (alpha=0.25). Selecting the 9 families in which the tumor of at least one patient showed LOH at marker D21S1255 decreased both the NPL and HLOD scores (Figure 3). Similar results were obtained for chromosome 12 (data not shown). In agreement with the absence of linkage, we were unable to detect consistent loss of the same parental allele on either #12 or #21 in the tumors from these families.

For chromosome 22, the highest multipoint LOD score was −11.34 between markers D22S303 and D22S315, and under the admixture model the estimated proportion of linked families was 0. When selecting the 12 families in which the tumor of at least one patient showed LOH at marker D22S445, the peak multipoint LOD score under heterogeneity was 0.06 (alpha=0.2) between marker D22S303 and D22S315 (27 cM proximal of D22S445).

DISCUSSION

We have analyzed 100 breast tumors from patients strongly selected for a particular familial background for loss of heterozygosity and immunophenotype analysis. To our knowledge, this is the first study analyzing LOH at all chromosome arms in such an extended and highly selected group of familial tumors. The main purpose of the study was to detect patterns of LOH and/or immunophenotype that would define distinct subgroups of tumors, on the basis of which we would then be able to stratify the families from which they derive. This is one approach to address the genetic heterogeneity problem, which is commonly believed to be the main reason for the inability to detect further moderate- to high-risk breast cancer susceptibility genes.[4,37] For this reason, we have selected cases from families with a high probability of segregating a breast cancer susceptibility gene, but with a minimal residual probability that this is *BRCA1* or *BRCA2*.

In many families we collected tumor tissues from two or more patients, allowing us to analyze whether certain genetic, immunohistochemical and morphological features were more prevalent within families than predicted by chance. We did indeed observe this for LOH with several markers, but not for any of the immunohistochemical markers. However, linkage analysis in the total group of 55 families did not produce significant LOD scores for any of these chromosomes, nor did linkage analysis in subgroups of families selected on basis of these LOH results. This suggests that LOH analysis of familial cases is unlikely to facilitate the detection of new breast cancer susceptibility loci by linkage analysis. It remains possible, however, that families in which multiple breast tumors show LOH at the same locus are caused by a shared genetic defect on another chromosome. A genome-wide linkage search in our families should address this. For example, it has been reported that breast tumors from families linked to *BRCA1* show more frequent LOH on 4q and 5q relative to sporadic breast cancer, which we have confirmed here.[36,38] Hence it might have been possible to detect linkage to BRCA1 among the families in which several tumors show LOH on 4q or 5q, rather than among families only selected on clinical phenotype. Although our LOH analysis covered all chromosome arms, certain LOH events may have escaped detection because of the limited number of markers we have used. LOH analysis with microarrays with 10,000 SNP's could indicate shared LOH regions with more accuracy, as was found for lung cancer.[39] LOH analysis with polymorphic markers detects any imbalance in parental chromosomes, including trisomy,[26] so that our 'LOH' scorings in fact reflect a wide range of different chromosomal aberrations. To distinguish between these, LOH data should be combined with (array-)CGH. This might be relevant

because we do not know at this stage whether other breast cancer susceptibility genes act according to Knudson's two-hit inactivation model.[26,40] It is conceivable, as was found for the *MET* oncogene in hereditary papillary renal carcinomas,[41] that trisomy (or copy-number gain) of the mutant allele contributes to susceptibility.

A better resolution for subgroup analysis of the tumors might be achieved by global gene expression analysis. Many different studies describe the possible classification of the heterogeneous group of sporadic breast cancers in distinct subtypes using microarray techniques.[42,43] Five different subtypes (one basal-like, one *ERBB2*-over-expressing, two luminal-like, and one normal breast tissue-like subgroup) have been recognized.[44] These tumor subtypes may represent different biological entities and might originate from different cell types. Four distinct phenotypes (pure luminal, mixed luminal/basal, pure basal and null) have been defined by immunostaining 1944 sporadic breast tumors with antibodies for both the luminal and basal phenotypes.[34] These subgroups were significantly different in their biological features and clinical course of the disease. In addition, another study[14] showed that the expression patterns from 15 fresh frozen tumor samples from 7 non-*BRCA1/2* families clustered within their respective families, suggesting an underlying common genetic basis. The recently developed DASL-assay technique,[45] which makes gene expression analysis possible in archival paraffin-embedded tissues, may extend this observation to larger numbers of cases.

The hypothesis that genetic predisposition to breast cancer might preferentially give rise to certain subtypes is also supported by histopathological findings in *BRCA1* related tumors. These are generally of higher grade, show pushing margin growth patterns and high lymphocyte infiltration in comparison to sporadic cases.[33] They are also more often estrogen receptor (ER), progesterone receptor (PR) negative, Bcl2-negative, P53-mutated and negative for Her2/Neu amplification (our data, and refs[9,11,46]). In gene expression profiling, a basal-like gene expression pattern has been associated with *BRCA1* carriers.[13] We found most *BRCA1* tumors (81.5%) to belong to the pure basal or mixed phenotype category, based on cytokeratin 5/6 and cytokeratin 19 expression, as opposed to the *BRCA2* tumors which were mostly (75%) of the luminal or null phenotype. Intriguingly, *BRCAX* tumors were almost equally distributed over both categories. However, we noted that different tumors within the same family frequently belonged to different phenotype categories, indicating that it is unlikely that the basal/luminal phenotype is genetically determined in these cases.

The morphological and immunohistochemical results from *BRCAX* breast carcinomas and those arising in *BRCA1* and *BRCA2* mutation carriers are similar to those

recently reported by others.[9,11,47,48] Only Bcl2 displayed a significant difference between *BRCAX* tumors and *BRCA1*- or *BRCA2*-tumors (both p<0.0001), but the proportion of positive *BRCAX* tumors is not conspicuously different from what is observed in series of unselected sporadic breast tumors.[9] In general, the patterns of immunostaining and LOH in *BRCAX* tumors closely resemble those of sporadic breast tumors, with the possible exception of the 'mixed' phenotype (as defined by cytokeratins 5/6, 19) and LOH at chromosome 22. Two recent studies[49,50] have used classical CGH to analyze a small number of BRCAX-related breast tumors. Both these studies too found chromosomal aneuploidy patterns broadly resembling those of sporadic breast tumors, but did not identify chromosome 22 as a frequent target for aneuploidy. Conversely, regions on chromosome 8 and 19, identified by CGH,[50] were not observed by us. It should be noted, however, that a direct comparison of the *BRCAX* cases in these studies and ours is difficult due to differences in the applied selection criteria for *BRCAX* families. For example, the occurrence of ovarian cancer was not used to exclude families in the CGH studies,[49,50] increasing the probability that some are caused by undetected mutations in *BRCA1*.

Thus, in our families a clustering of sporadic, or sporadic-like breast cancer is seen. Yet, it has been argued that such familial clustering is unlikely to occur by chance but instead is more likely to have a genetic basis.[1] Therefore, if our families indeed have a genetic basis, our results suggest that this basis is the same as that for sporadic breast cancer. Analyses of genetic models to explain familial breast cancer have indicated that, after correction for *BRCA1* and *BRCA2*, the polygenic model incorporating multiple interacting low penetrance genes is the most likely explanation.[51-53] Such genes are also suspected to explain a substantial proportion of sporadic breast cancer. If more detailed analyses of this group of patients by high-resolution array-CGH or gene expression profiling confirms that these tumors resemble sporadic tumors very much, than this is in agreement with the idea that the remainder of familial risk to breast cancer is caused in a polygenic way. Finding these genes will be a challenge for years to come, but family studies will remain valuable in this regard because one is enriching for genetic susceptibility,[54] as was convincingly shown with the identification of the *CHEK2**1100delC variant.[15,55]

AKNOWLEDGMENTS

We want to thank Annemieke M. van der Wal for here assistance with the immunohistochemical staining. We want to thank the Dutch Cancer Society (KWF; RUL1999-2021)) and the Center for Medical Systems Biology for the Financial support.

REFERENCE LIST

1. **Ford D, Easton DF, Stratton M, et al.** Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. *Am.J.Hum.Genet. 62:676-89 (1998)*
2. **Peto J, Collins N, Barfoot R, et al.** Prevalence of BRCA1 and BRCA2 gene mutations in patients with early- onset breast cancer. *J.Natl.Cancer Inst. 91:943-9 (1999)*
3. **Easton DF.** How many more breast cancer predisposition genes are there? *Breast Cancer Res. 1:14-7 (1999)*
4. **Narod SA, Foulkes WD.** BRCA1 and BRCA2: 1994 and beyond. *Nature Reviews Cancer 4:665-76 (2004)*
5. **Seitz S, Rohde K, Bender E, et al.** Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: linkage analysis in German breast cancer families. *Oncogene 14:741-3 (1997)*
6. **Kainu T, Juo SH, Desper R, et al.** Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proc.Natl.Acad.Sci.USA 97:9603-8 (2000)*
7. **Rahman N, Teare MD, Seal S, et al.** Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. *Oncogene 19:4170-3 (2000)*
8. **Thompson D, Szabo CI, Mangion J, et al.** Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. *Proc.Natl.Acad.Sci.USA 99:827-31 (2002)*
9. **Palacios J, Honrado E, Osorio A, et al.** Immunohistochemical characteristics defined by tissue microarray of hereditary breast cancer not attributable to BRCA1 or BRCA2 mutations: Differences from breast carcinomas arising in BRCA1 and BRCA2 mutation carriers. *Clinical Cancer Research 9:3606-14 (2003)*
10. **van Beers EH, Van Welsem T, Wessels LF, et al.** Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. *Cancer Res. 65:822-7 (2005)*
11. **Eerola H, Heikkila P, Tamminen A, Aittomaki K, Blomqvist C, Nevanlinna H.** Histopathological features of breast tumours in BRCA1, BRCA2 and mutation-negative breast cancer families. *Breast Cancer Res. 7:R93-100 (2005)*
12. **Lakhani SR, Jacquemier J, Sloane JP, et al.** Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J.Natl.Cancer Inst. 90:1138-45 (1998)*
13. **Van 't Veer LJ, Dai HY, Van de Vijver MJ, et al.** Gene expression profiling predicts clinical outcome of breast cancer. *Nature 415:530-6 (2002)*
14. **Hedenfalk I, Ringner M, Ben Dor A, et al.** Molecular classification of familial non-BRCA1/BRCA2 breast cancer. *Proc.Natl.Acad.Sci.USA 100:2532-7 (2003)*
15. **Oldenburg RA, Kroeze-Jansema K, Kraan J, et al.** The CHEK2*1100delC variant acts as a breast cancer risk modifier in non-BRCA1/BRCA2 multiple-case families. *Cancer Res. 63:8153-7 (2003)*
16. **Hogervorst F, Cornelis R, Bout M, et al.** Rapid detection of BRCA1 mutations by the Protein Truncation Test. *Nature Genet. 10:208-12 (1995)*
17. **Oldenburg RA, de Vos tot Nederveen Cappel WH, van Puijenbroek M, et al.** Extending the p16-Leiden tumour spectrum by respiratory tract tumours. *J.Med.Genet. 41:e31 (2004)*
18. **Petrij-Bosch A, Peelen T, Van Vliet M, et al.** BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nature Genet.17:341-5 (1997)*
19. **Korkko J, Annunen S, Pihlajamaa T, Prockop DJ, Ala-Kokko L.** Conformation sensitive gel electrophoresis for simple and accurate detection of mutations: comparison with denaturing gradient gel electrophoresis and nucleotide sequencing. *Proc.Natl.Acad.Sci.USA 95:1681-5 (1998)*

20. **Hogervorst FBL, Nederlof PM, Gille JJP, et al.** Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method. *Cancer Res. 63:1449-53 (2003)*
21. Tumours of the breast. *In* **F. A. Tavassoli and P. Devilee *(eds.)*,** World Health Organization: Tumours of the Breast and Female Genital Organs Lyon *IARC Press, 2003*
22. **Ghebranious N, Vaske D, Yu A, Zhao C, Marth G, Weber JL.** STRP screening sets for the human genome at 5 cM density. *BMC. Genomics 4:6 (2003)*
23. **Gupta PK, Sahota A, Boyadjiev SA, et al.** High frequency in vivo loss of heterozygosity is primarily a consequence of mitotic recombination. *Cancer Res. 57:1188-93 (1997)*
24. **Tischfield JA.** Loss of heterozygosity or: how I learned to stop worrying and love mitotic recombination. *Am.J.Hum.Genet. 61:995-9 (1997)*
25. **Cleton-Jansen AM, Callen DF, Seshadri R, et al.** Loss of heterozygosity mapping at chromosome arm 16q in 712 breast tumors reveals factors that influence delineation of candidate regions. *Cancer Res. 61:1171-7 (2001)*
26. **Devilee P, Cleton-Jansen AM, Cornelisse CJ.** Ever since Knudson. *Trends Genet. 17:569-73 (2001)*
27. **Kononen J, Bubendorf L, Kallioniemi A, et al.** Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nature Med. 4:844-7 (1998)*
28. **Houwing-Duistermaat JJ, Derkx BH, Rosendaal FR, Van Houwelingen HC.** Testing familial aggregation. *Biometrics 51:1292-301 (1995)*
29. **Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES.** Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am.J.Hum.Genet. 58:1347-63 (1996)*
30. **Claus EB, Risch N, Douglas W.** Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am.J.Hum.Genet. 48:232-42 (1991)*
31. **Easton D, Bishop D, Ford D, Crockford G,** Breast Cancer Linkage Consortium. Genetic linkage analysis in familial breast and ovarian cancer: Results from 214 families. *Am.J.Hum.Genet. 52:678-701 (1993)*
32. **Abecasis GR, Cherny SS, Cookson WO, Cardon LR.** Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat.Genet. 30:97-101 (2002)*
33. **Lakhani SR, Gusterson BA, Breast Cancer Linkage Consortium, et al.** The pathology of familial breast cancer: histological features of cancers in families not attributable to mutations in BRCA1 or BRCA2. *Clin.Cancer Res. 6:782-9 (2000)*
34. **Abd El-Rehim DM, Pinder SE, Paish CE, et al.** Expression of luminal and basal cytokeratins in human breast carcinoma. *J.Pathol. 203:661-71 (2004)*
35. **Miller BJ, Wang D, Krahe R, Wright FA.** Pooled analysis of loss of heterozygosity in breast cancer: a genome scan provides comparative evidence for multiple tumor suppressors and identifies novel candidate regions. *Am.J.Hum.Genet. 73:748-67 (2003)*
36. **Tirkkonen M, Johansson O, Agnarsson BA, et al.** Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. *Cancer Res. 57:1222-7 (1997)*
37. **Nathanson KL, Weber BL.** 'Other' breast cancer susceptibility genes: searching for more holy grail. *Hum.Mol.Genet. 10:715-20 (2001)*
38. **Tirkkonen M, Kainu T, Loman N, et al.** Somatic genetic alterations in BRCA2-associated and sporadic male breast cancer. *Genes Chrom.Cancer 24:56-61 (1999)*
39. **Janne PA, Li C, Zhao XJ, et al.** High-resolution single-nucleotide polymorphism array and clustering analysis of loss of heterozygosity in human lung cancer cell lines. *Oncogene 23:2716-26 (2004)*
40. **Knudson AG.** Hereditary cancer: Theme and variations. *J.Clin.Oncol. 15:3280-7 (1997)*

41. **Zhuang Z, Park WS, Pack S, et al.** Trisomy 7-harbouring non-random duplication of the mutant MET allele in hereditary papillary renal carcinomas. *Nature Genet. 20:66-9 (1998)*
42. **Perou CM, Sorlie T, Eisen MB, et al.** Molecular portraits of human breast tumours. *Nature 406:747-52 (2000)*
43. **Sorlie T, Perou CM, Tibshirani R, et al.** Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc.Natl.Acad.Sci.USA 98:10869-74 (2001)*
44. **Sorlie T, Tibshirani R, Parker J, et al.** Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc.Natl.Acad.Sci.USA 100:8418-23 (2003)*
45. **Bibikova M, Talantov D, Chudin E, et al.** Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. *Am.J.Pathol. 165:1799-807 (2004)*
46. **Lakhani SR, Manek S, Penault-Llorca F, et al.** Pathology of ovarian cancers in BRCA1 and BRCA2 carriers. *Clin.Cancer Res. 10:2473-81 (2004)*
47. **Lakhani SR, Van de Vijver MJ, Jacquemier J, et al.** The pathology of familial breast cancer: Predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J.Clin.Oncol. 20:2310-8 (2002)*
48. **Foulkes WD, Stefansson IM, Chappuis PO, et al.** Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *J.Natl.Cancer Inst. 95:1482-5 (2003)*
49. **Maguire P, Holmberg K, Kost-Alimova M, Imreh S, Skoog L, Lindblom A.** CGH analysis of familial non-BRCA1/BRCA2 breast tumors and mutation screening of a candidate locus on chromosome 17q11.2-12. *Int.J.Mol.Med 16:135-41 (2005)*
50. **Gronwald J, Jauch A, Cybulski C, et al.** Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. *Int.J.Cancer 114:230-6 (2005)*
51. **Antoniou AC, Pharoah PDP, McMullan G, et al.** A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br.J.Cancer 86:76-83 (2002)*
52. **Antoniou AC, Easton DF.** Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet.Epidemiol. 25:190-202 (2003)*
53. **Cui J, Antoniou AC, Dite GS, et al.** After BRCA1 and BRCA2-What Next? Multifactorial Segregation Analyses of Three-Generation, Population-Based Australian Families Affected by Female Breast Cancer. *Am.J.Hum.Genet. 68:420-31 (2001)*
54. **Houlston RS, Peto J.** The future of association studies of common cancers. *Hum.Genet. 112:434-5 (2003)*
55. **Meijers-Heijboer H, Van den Ouweland A, Klijn J, et al.** Low-penetrance susceptibility to breast cancer due to CHEK2*1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nature Genet. 31:55-9 (2002)*

4.2. BRCAX BREAST TUMORS COMPRISE A HETEROGENEOUS CLASS, DISTINCT FROM SPORADIC AND BRCA1 BREAST TUMORS BY ARRAY-CGH, AND MAY REFLECT MULTIPLE ETIOLOGIES

E.H. van Beers,[1] R.A. Oldenburg,[6] S.A. Joosse,[1] L.F.A. Wessels,[2,5] M.J. Ligtenberg,[7] N. Hoogerbrugge,[7] K. Kroeze-Jansema,[8] H. Meijers-Heijboer,[10] F.B. Hogervorst,[4] S. Verhoef,[4] P. Devilee,[8,9] and P.M. Nederlof[1,3,11]

Submitted

[1] *Division of Experimental Therapy,* [2] *Molecular Biology,* [3] *Molecular Pathology, and* [4] *Family Cancer Clinic, The Netherlands Cancer Institute, Amsterdam, The Netherlands.*

[5] *Faculty of Electrical Engineering, Mathematics and Computer Science, University of Technology, Delft, The Netherlands.*

[6] *Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands.*

[7] *Dept. of Pathology and Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.*

[8] *Dept. Of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands.*

[9] *Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands.*

[10] *Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands*

[11] *Corresponing author P. Nederlof, e-mail: p.nederlof@nki.nl*

ABSTRACT

Only about 25% of familial breast cancer is explained by mutations in *BRCA1* and *BRCA2*, fewer by moderate penetrance genes like *P53*, *PTEN*, *CHEK2*, *ATM* and *PALB2* and an unknown fraction by common variants of genes with low penetrance. Evidence suggests that additional dominant breast cancer genes exist and these are referred to as *BRCAX*. Clinical presentation of families with highly increased incidence of breast cancer that are *non-BRCA1/BRCA2*, suggests dominant inheritance of such high penetrance breast cancer genes. Because cancer genes often confer a specific clinical presentation (e.g. age of onset, sex-ratio, tissue spectrum) it seems

useful to initiate their discovery by such clinical criteria. An earlier study of *BRCAX* / *non-BRCA1/2* breast cancer families aimed to enrich for a common genetic defect by setting stringent inclusion criteria, failed to identify new breast cancer susceptibility loci. Motivated by results of *BRCA1* and *BRCA2* breast tumors that have characteristic array-CGH signatures (array-CGH 'phenotypes'), we study *BRCAX* breast tumor by array-CGH and show that *BRCAX* tumors are distinct from sporadic controls but are otherwise still heterogeneous. This provides a possible explanation for the lack of high LOD scores in these patients and would be consistent with more than one *BRCAX* sub-type and therefore more than one *BRCAX* gene. We propose approaches that can be employed to further sub-stratify *BRCAX* families based on array-CGH data.

INTRODUCTION

The majority of excess familial breast cancer risk is unexplained. We now know that the underlying genetics of breast cancer susceptibility is very complex. *BRCA1*[1] and *BRCA2*[2] were the first breast cancer genes identified due to the combination of high penetrance (carriers have substantially increased risk) with large affected families for confirmative co-segregation analyses. These two genes have been widely studied and their collective contribution to hereditary breast cancer incidence is now estimated at perhaps 20-25%.[3,4] The other 75% of hereditary cases are likely caused by a multitude of unknown risk factors and are attributable to recessive genes, combinations of genes or to common variants of genes conferring only slightly elevated risks.[5,6] Yet, analyses of *BRCA1* and *BRCA2* allele frequencies and penetrance, and especially patterns of familial clustering of breast cancer suggested the existence of additional, dominant high penetrance breast cancer genes, referred to as either '*BRCAX*' or '*BRCA3*'.[7] A linkage study on 149 *BRCAX* families, including those in the present study, failed to identify sufficiently high LOD scores to guide positional cloning of the gene(s) in these families[8] which led the authors to conclude that perhaps more than one risk conferring locus was involved. Because similar to most breast tumors, *BRCAX* breast tumors present with genomic instability,[9,10] we performed array-CGH of *BRCAX* breast tumors to catalogue possible distinct and recurrent CGH profiles. We compare *BRCAX* array-CGH profiles with those of *BRCA1*, *BRCA2* and sporadic tumors and describe possible similarities and particularities among *BRCAX* breast tumors and propose an approach for further analysis of *BRCAX* families.

MATERIALS AND METHODS

Patients and tumor specimens

Family selection, tumor collection and immunohistochemistry have been described in Oldenburg et al.[8] Control tumors were selected to have no evidence of familial risk for breast or ovarian cancer and were on average as young (45.5) as the *BRCAX* patients (52.5) years. All control samples were selected from the institute's archival tissue bank and are described in detail elsewhere.[11] Of 92 patients from 42 *BRCAX* families[8,12] we isolated high-quality DNA from 58 unique tumor samples from 27 different families to perform array-CGH as before.[13] Immunohistochemistry was performed as described, for *BRCAX*[8] and for sporadic controls.[14]

Array-CGH

DNA from each tumor and a reference of pooled DNA from seven healthy females was used in ~1 Mb genome-wide BAC array-CGH. Automated hybridizations were performed for 72 hrs, followed by automated washes and drying as described.[13] Arrays were scanned with an Agilent DNA microarray scanner. Signal intensities were determined with Imagene software and raw data processing involved only median pintip (c.q. sub-array) normalization. Array-CGH profile log2 ratios were used in unsupervised hierarchical clustering using Matlab software (v.7.0.1, The Mathworks, Natick MA, USA).

RESULTS

Array-CGH of BRCAX tumors

Of a previous collection of *BRCAX* invasive breast tumors encompassing 84 FFPE blocks,[8] we isolated genomic tumor DNA of sufficient quality from 58 unique cases from 27 different families to perform automated array CGH. Forty-eight invasive ductal sporadic breast carcinomas (IDC) with similar age of incidence were used for comparison.

Array-CGH copy number alterations (CNA's) were interpreted as gains when log2 ratios were > 0.2 and losses when log2 ratios were < -0.2. Counts of aberrations for each of the 58 *BRCAX* and 48 sporadic control breast tumors are shown in figure 1. The average number of BAC's reporting CNA in *BRCAX* tumors was 1063 (sd = 442) and in sporadic controls 816 (sd = 386). This difference was statistically different in a two-sided, unpaired t-test (P = 0.003). We conclude that on average *BRCAX* tumors have (1063/816) 130.3% aberrant clones compared with sporadic tumors.

BRCAX array-CGH aberrations in comparison with BRCA1 and sporadic controls

Figure 1 is a frequency plot of array-CGH gains and losses in *BRCAX* versus control breast tumors (figure 1a) versus *BRCA1* tumors (figure 1b) and versus *BRCA2* tumors (figure 1c). Continuous CGH data consisted of 3248 measurements (BAC probes) from chromosome 1p-tel to Xq-tel. All CGH data were segmented[15] before counting outlier frequencies as a fraction of each tumor class showing a gain (log2 ratios > 0.2) or loss (log2 ratios < -0.2) for all 3248 positions measured. Gain and loss frequencies (0-1) are plotted on the left y-axis. The p-values for the differences between the tumor classes were computed by two-sided Fisher exact testing and plotted in green (differential gains) or red (differential losses). Longer sticks correspond with smaller p-values and therefore more significant regions (figure 1).

The comparison between *BRCAX* and sporadic tumors shows multiple regions of differential gains and losses (figure 1a). Differential gains are prominent on chromosome 2q-ter, 6p, 8p, 11p 12p 14q, 17p, 17q, 19p, 19q, and along the entire Chromosome 22. Differential losses are prominent on 1p, 1q, 4q, 5q, 9q, 13q, 14q, 15q, 19cen, 21p and Xp. Despite these differences there is also abundant overall similarity between *BRCAX* and sporadic breast tumor array-CGH profiles. For example, frequent aberrations observed in both classes are gain of chromosome 1q and 8q. Significant differential gains and losses depend highly on which tumor classes are compared (figure 1a, b, and c). A comparison of differential recurrent aberrations between *BRCAX* and *BRCA1* (figure 1b) or *BRCAX* and *BRCA2* (figure 1c) resulted in different sets of significant regions. Figure 1b shows for instance regions that are known to be highly specific to *BRCA1* tumors such as3p, 3q and 5cen.[16] Aberrations in these three regions in *BRCA1* tumors (figure 1a, black line) were more frequent then in *BRCAX* tumors (blue line), and thus more characteristic for BRCA1 tumors. Other significant differential CGH results in figure 1 include chromosome 13p loss, which is more frequent in *BRCA2*, while a region on 12p towards the centromere appeared as gain in approximately 30% of *BRCAX* tumors, which was not found in controls, *BRCA1* or *BRCA2* tumors.

Array-CGH aberration banding (Pearson banding)

Another analysis that highlights the specific gains and losses of *BRCAX* tumor array-CGH profiles is shown in figure 2. We calculated Pearson correlation coefficients of the log2 ratios between all possible pairs of BAC's per tumor class per chromosome. These Pearson coefficients are plotted in the three top panels as heat-maps. The Pearson heat maps detect Pearson-stable regions ('bands') with great sensitivity (as 'oran-

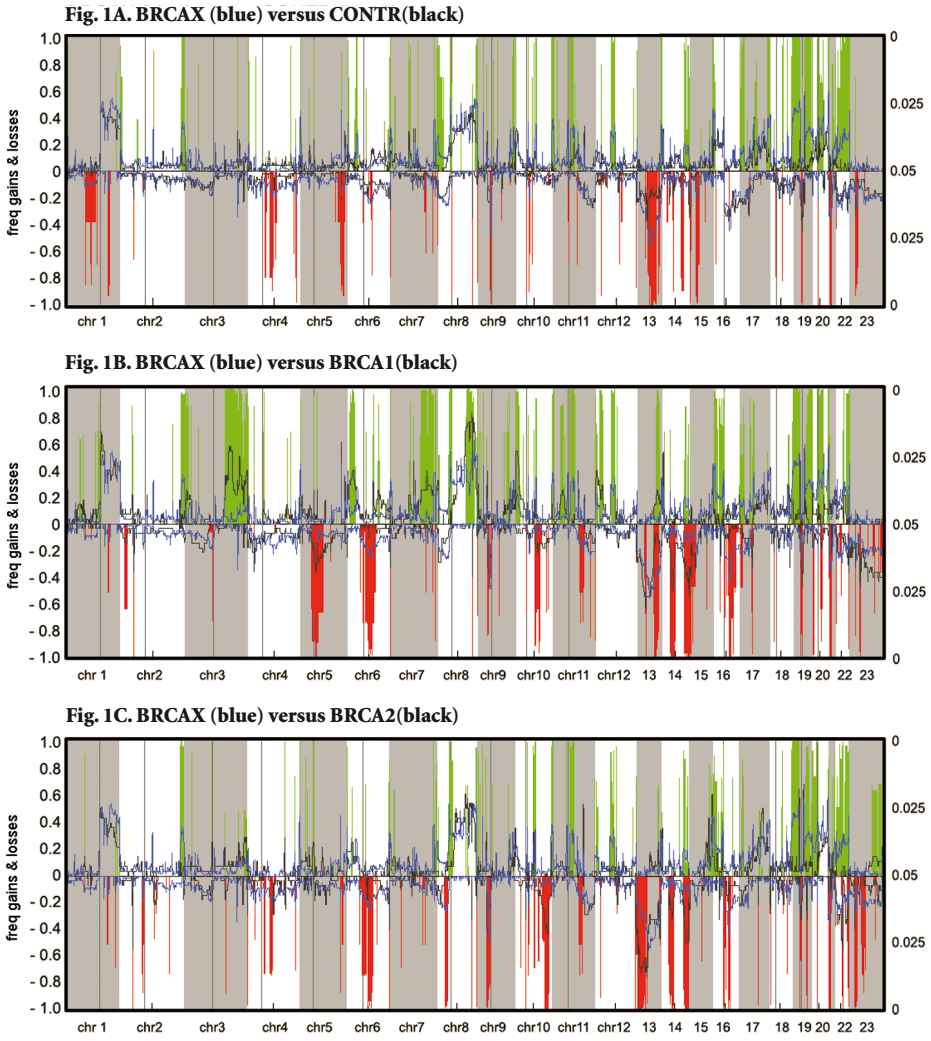
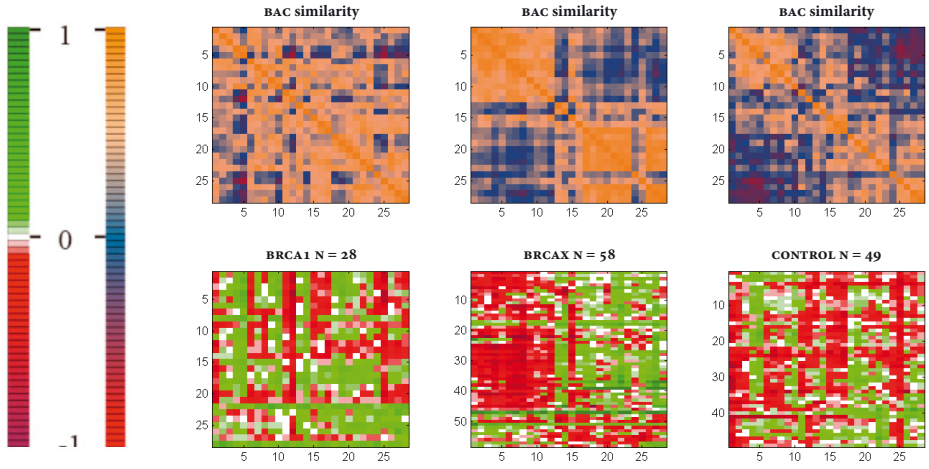


Figure 1. Significantly different CGH aberrations in *BRCAX* tumors (blue) and control tumors (black). The top panel shows Chromosomes 1 through 8, and the bottom panel Chromosome 9 through X. The x-axis represents all 3248 probes on Chromosomes 1 through X and vertical black lines indicate centromers. On the y-axis are the frequencies of aberrations (|log2ratio| > 0.2) in 58 *BRCAX* tumor CGH profiles (blue) and 48 sporadic control tumors (black). Vertical green bars correspond to between-class gain significance as determined with a two-sided Fisher exact test (P-value scale on right y-axis, ranging from P =0.05 at the X-intercept to 0.00 at the top and bottom, i.e. all depicted bars are significant at the 5% level). Similarly, red bars indicate significance for differential losses between *BRCAX* and sporadic tumors. (class X > Class X in figure).

ge' blocks). Below each Pearson heat map, we plotted the log2 ratio heat-maps for the same chromosomes (other chromosome figures are given as supplementary data).

Figure 2a-2e. Pearson correlations (top panels) and CGH log2 ratios (bottom panels) in whole chromosomes 13, 16 and 22. The top panels have all BAC clones for that particular chromosome on both the x and y-axis. The bottom panels show chromosome centromers if present (arrow) while individual samples are stacked and sorted along the y-axis. This vertical sorting in the lower panels samples is performed per class, based on sample-to-sample complete correlation clustering. Color scales were set to saturate at -1 and 1 for correlation (red-blue-orange) and log2 ratios (red-white-green). Horizontal axes not plotted to scale but depend on the number of features on each particular chromosome.

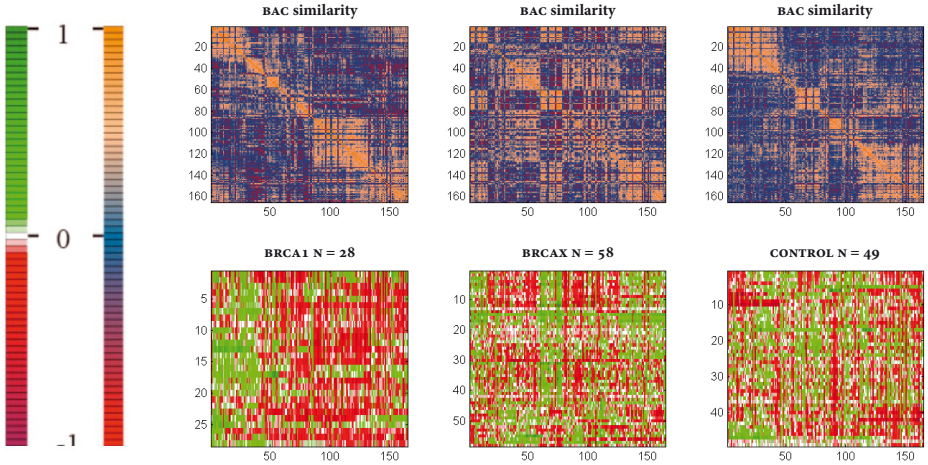
Fig. 2a Chromosome 21



Chromosome 21

Figure 2A shows results of 58 *BRCAX* tumors, flanked by 28 *BRCA1* and 49 sporadic control tumors. The majority of *BRCAX* tumors show loss of 21p and gain of 21q but the reverse pattern can be discerned also. Both other classes have quite different CGH profiles with more scattered patches of gains and losses. This is reflected in the three correlation panels with more distinct p and q 'blocks' of high correlation for *BRCAX* tumors compared with the two other classes.

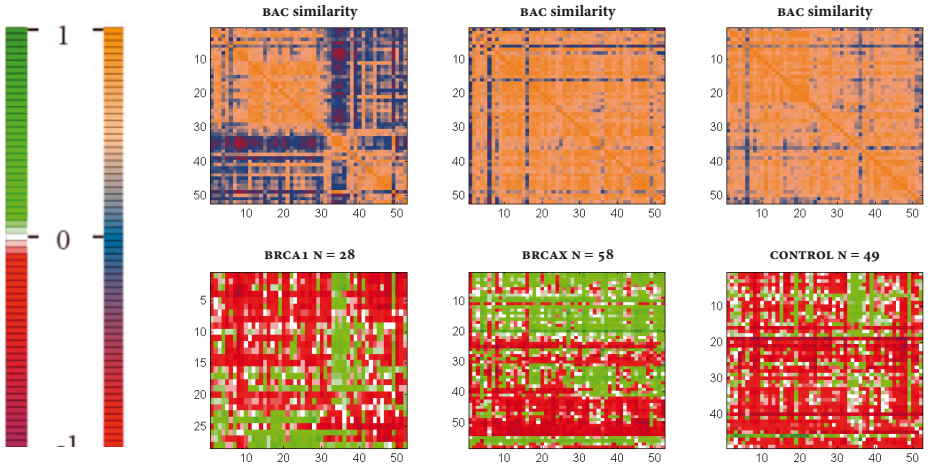
Fig. 2b Chromosome 12



Chromosome 12

Figure 2B shows chromosome 12. Here BRCA1 tumors seem the most homogeneous of the three classes plotted. While *BRCAX* and sporadic control CGH data (lower panels) might look somewhat similar and heterogeneous, the correlation panels indicate that *BRCAX* tumors have more off-diagonal high-correlation regions. This means similarity between dis-continuous segments of chromosome 12, for example BAC's 1~20 with 60~80. This could suggest intra chromosomal rearrangement joining these regions in *BRCAX* tumors. The *BRCAX* CGH log2 ratios further indicate that this class is heterogeneous, with at least several different types of profiles for this chromosome.

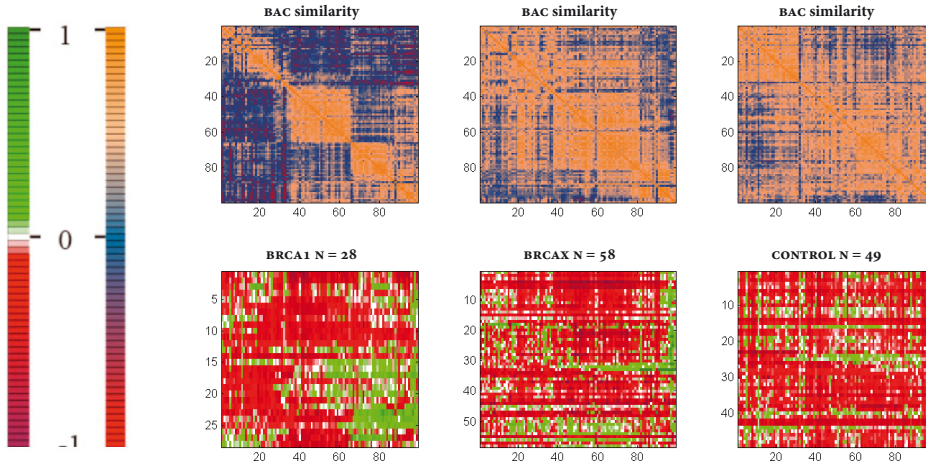
Fig. 2c Chromosome 22



Chromosome 22

Figure 2C shows the results for chromosome 22. It seems to suggest two types of *BRCAX* tumors, namely those with gain and those with loss of the entire chromosome, but also a few tumors with more complex rearrangements. This heterogeneity among *BRCAX* tumors could not be seen in Figure 1, which presents average data for whole tumor classes.

Fig. 2d Chromosome 13



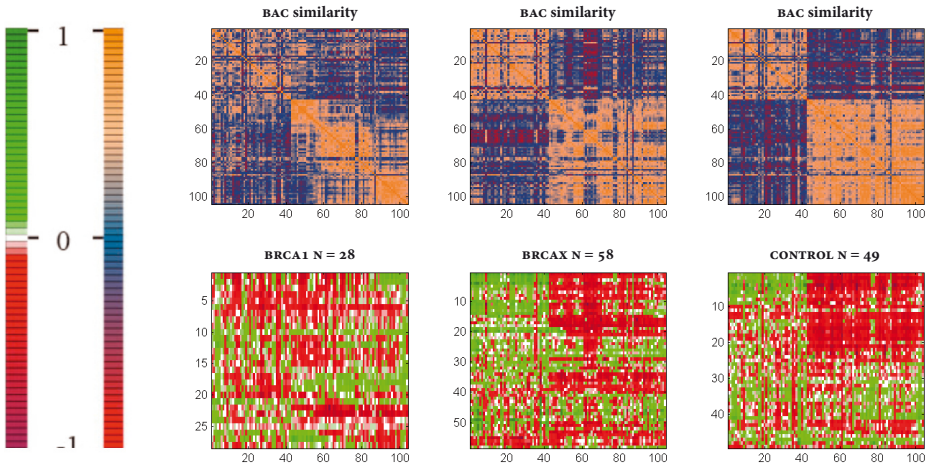
Chromosome 13

Figure 2D shows chromosome 13. Pearson correlations are high throughout this chromosome for *BRCAX* samples (top middle) but vary considerably across this chromosome in the *BRCA1* class (top left). We conclude that *BRCAX* tumors have fewer transitions from gain to loss or vice versa and thus have a more stable chromosome 13 compared with *BRCA1* tumors. Interestingly, there are a few *BRCAX* tumors with different chromosome 13 profiles.

Chromosome 16

Figure 2E shows chromosome 16. Aberrations of Chromosome 16p are more variable in *BRCA1* (top left, more blue) compared with either *BRCAX* (middle) or controls (right). The *BRCAX* class showed a unique recurrent loss (red in lower panel) between clone 61 and 70 that was inversely correlated to the log2 ratios (top middle, red correlations) of BAC clones 1 ~ 40 on the same chromosome. This means that a gain of 16p seems to co-occur with loss of another specific region of 16q only among *BRCAX*

Fig. 2e Chromosome 16

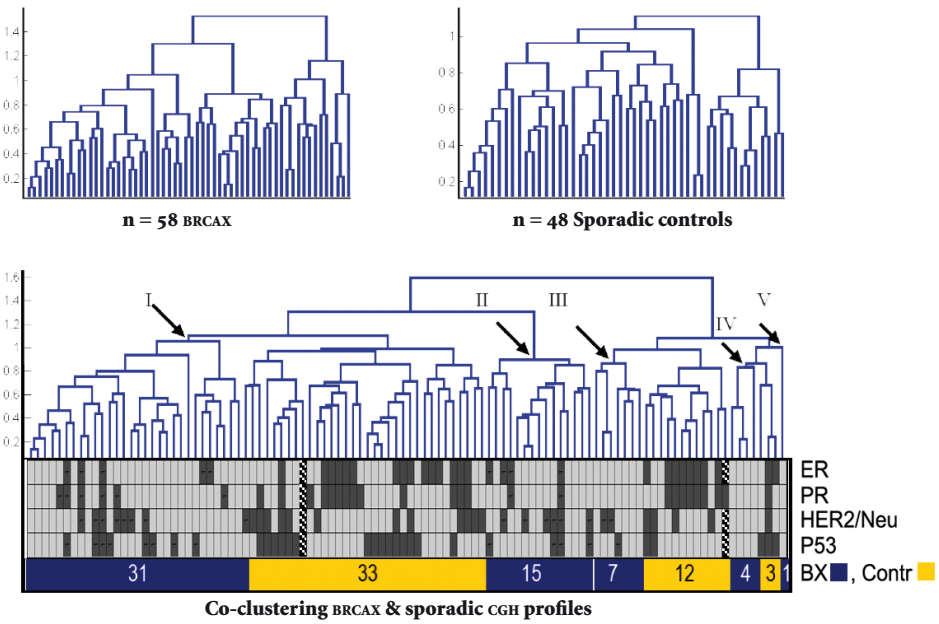


tumors. Overall *BRCAX* array-CGH profiles for this chromosome are quite similar to sporadic controls. This is consistent with Figure 1 that already indicated no significant differences for chromosome 16 between *BRCAX* and sporadic tumors.

Heterogeneity of BRCAX tumor array-CGH profiles

We have shown that CGH profiles of *BRCAX* tumors are different from those of sporadic controls and from *BRCA1* and *BRCA2* tumors. However this does not mean that *BRCAX* tumor profiles are homogeneous. Although *BRCAX* tumors were selected with stringent inclusion criteria,[8] we lack understanding of the genetic factor(s) that caused their apparent familial excess risk and the possibility remains that *BRCAX* families represent more then one risk factor that might associate with distinct *BRCAX* array-CGH sub-phenotypes. We therefore investigated the extent of heterogeneity among these 58 *BRCAX* tumors by unsupervised hierarchical clustering (complete linkage Pearson correlation of whole CGH profiles). Figure 3A uses all 3248 log2 ratios for all autosomes and chromosome X. Patients are never hybridized in sex mismatch since the normal reference pool is also female DNA. The hierarchical trees in Figure 3A and B show clustering of the 59 *BRCAX* and 49 sporadic tumor samples, respectively. These dendrograms indicated that the 'within class' heterogeneities are comparable between sporadic and *BRCAX* tumors and that there are no major branch points to suggest obvious distinct *BRCAX* array-CGH subtypes. Then we co-clustered all sporadic and *BRCAX* tumors in figure 3C together with their immunophenotypes, and found that individual tumors did not mix randomly. Both *BRCAX* and sporadic

Figure 3. Unsupervised Hierarchical Cluster of brcax and Control tumor cgh profiles.
(A) Unsupervised complete correlation clustering of 58 *BRCAX* tumor array-CGH profiles is performed in Matlab (The Mathworks, Natick MA, USA) using log2 ratios for 3248 probes from chromosome 1-X. Vertical distances represent the similarity distance calculated across all 3248 probes. (B) Similar to A, shows 48 sporadic breast tumors. (C) *BRCAX* tumors and sporadic tumors co-clustered. *BRCAX* are blue, sporadics are yellow. Immunohistochemical staining scores are given as no staining (grey), positive staining (black) or missing data (hatched). The bottom legend indicates the *BRCA1*-likelihood score in our *BRCA1* classifier[14]. Red = '*BRCA1*-like', grey = undecided, yellow = 'sporadic-like'.



tumors remained clustered in just eight sub-clusters, five of which (I-V) contained only *BRCAX* tumors and three clusters contained all sporadic cases plus one *BRCAX* tumor.

DISCUSSION

Selection of BRCAX tumors

Because the breast cancer gene(s) in *BRCAX* families is (are) unknown, *BRCAX* is solely defined by clinical criteria, including a negative test for known breast cancer genes like *BRCA1* and *2*. In our study, a nation-wide collection of such *BRCAX* tumors

from selected families[8] was analyzed by array-CGH. These families had ≥ 3 breast cancers below 60 years, and no ovarian or male breast cancer, and are more stringently selected compared with earlier *BRCAX* reports that have not excluded ovarian cancer,[10,17,18] that included samples with less then 70% tumor cells also,[10] or included 6 (of 18) families with just 2 cases of breast cancer.[19] The abovementioned differences may impact on risk factor stratification and therefore limits a comparison with our study. Approximately half of the samples (13 families) in this study were analyzed before in a linkage analysis of *BRCAX* families[20] and failed to map significant LOD scores leaving these families and their tumors largely uncharacterized at the genomic level. The current CGH analysis has revealed a high degree of heterogeneity among these 58 *BRCAX* samples, which could explain why the previous linkage analysis was unsuccessful.

BRCAX characteristic aberrations

The most significant differential chromosomal losses (< log2 ratio -0.2) between *BRCAX* and sporadic tumors were found on chromosomes 1, 4, 5, 13, 14, 15, 19, 21, X, and gains (> log2 ratio 0.2) on 2, 6, 8, 10, 11, 12, 14, 17, 19, 20, and 22. Notably, all these aberrations were more frequent in the *BRCAX* class than in the sporadic class. This is consistent with the results in Figure 1 showing that *BRCAX* tumors have ~30% more aberrations then sporadic breast cancers and also with results that indicate more genetic instability in hereditary compared with sporadic cases.[21] Chromosome 22 gain among *BRCAX* tumors seems unique since we have not observed this high frequency gain in either sporadic, *BRCA1* or *BRCA2* breast tumors. It will be of interest to unravel the role of chromosome 22 in *BRCAX* tumors.[22]

BRCAX heterogeneity

An elusive but crucial aspect of *BRCAX* families is whether or not different risk factors were co-selected by the clinical criteria used. We realize that this 'catch-22' will only end with identification of the risk factor(s) and therefore we hope to facilitate their identification by providing a possibly relevant stratification based on array-CGH profiles. At this time it remains impossible to predict whether IHC, array-CGH or any other method will provide such stratification.

Could BRCAX tumors be false negative BRCA1 tumors?

The possibility of false negative *BRCA1* diagnoses among the *BRCAX* tumors analyzed is low because these families, and individuals included in this study have tested ne-

gative for *BRCA1* and *BRCA2* in routine screening. Furthermore, a prognostic CGH study in HBOC families with the highest Evans' scores (i.e predicted to have *BRCA1* mutations)[14] found very few *BRCA1* mutations. Finally, one might argue that sensitivity of the *BRCA1* CGH classifier has not been estimated in elaborate studies, but it has detected one previously unclassified *BRCA1* variant M1775K.[23]

Future directions

Heterogeneity among tumors is difficult to quantify. Figure 3C serves as a hypothesis generator with respect to the question whether one or multiple *BRCAX* genes exist. It seems to indicate that certain tumor array-CGH profiles are more similar to each other compared with the rest but knowing which split(s) in the dendrogram might coincide with separate risk factors can only be tested by further analysis of multiple tumors from multiple *BRCAX* families and is currently in progress.

Due to the patient selection, we now have a dataset comprised of multiple tumors of multiple *BRCAX* families that will allow us to further analyze whether breast tumor profiles are more similar within families. This has proven difficult in this preliminary analyses perhaps due to the fact that not all genomic alterations recur with the same frequency, and because array-CGH profiles have unknown contributions from random (experimental and sampling noise) and a non-random (true) CNV's. Recent studies have estimated 26% 'phenocopies' in breast cancer in breast cancer families.[24] Therefore, it seems that equal weighing all log2 ratios, and more importantly including all *BRCAX* tumors without excluding phenocopies, will be inappropriate to define such 'family intrinsic profiles'. We hypothesize that further studies of *BRCAX* families based on CGH profile similarities, could contribute to the identification of the (perhaps multiple) *BRCAX* loci.

REFERENCE LIST

1. **Miki Y, Swensen J, Shattuck-Eidens D, et al.** A strong candidate for the breast and ovarian cancer susceptibility gene brca1. *Science. Oct 7;266(5182):66-71 (1994)*
2. **Wooster R, Neuhausen SL, Mangion J, et al.** Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science 265(5181):2088-90. (1994)*
3. **Antoniou AC, Pharoah PP, Smith P, and Easton DF.** The boadicea model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer. 91(8):1580-90. (2004)*
4. **Antoniou AC and Easton DF** Models of genetic susceptibility to breast cancer. *Oncogene 25(43):5898-905. (2006)*
5. **Cox A, Dunning AM, Garcia-Closas M, et al.** Breast Cancer Association Consortium. A common coding variant in casp8 is associated with breast cancer risk. *Nat Genet. 39(3):352-8. (2007)*

6. **Burwinkel B, Shanmugam KS, Hemminki K, et al.** Transcription factor 7-like 2 (TCF7L2) variant is associated with familial breast cancer risk: a case-control study. *BMC Cancer 6:268. (2006)*
7. **Hopper JL.** More breast cancer genes? *Breast Cancer Res. 3:154-7. (2001)*
8. **Oldenburg RA, Kroeze-Jansema K, Meijers-Heijboer H, et al.** Characterization of familial non-BRCA1/2 breast tumors by loss of heterozygosity and immunophenotyping. *Clin Cancer Res. 12(6):1693-700. (2006)*
9. **Jönsson G, Naylor TL, Vallon-Christersson J, et al.** Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. *Cancer Res. 65(17):7612-21. (2005)*
10. **Gronwald J, Jauch A, Cybulski C, et al.** Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. *Int J Cancer. 114(2):230-6. (2005)*
11 **Van Beers EH, van Welsem T, Wessels LF, Li Y, Oldenburg RA, et al.** Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. *Cancer Res. 65(3):822-7. (2005)*
12. **Thompson D, Szabo CI, Mangion J, Oldenburg RA, et al.** Evaluation of linkage of breast cancer to the putative brca3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium *Proc Natl Sci USA 99(2):827-31. (2002)*
13. **Joosse SA, van Beers EH, and Nederlof PM.** Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. *BMC Cancer 7;7:43. (2007).*
14. **Joosse SA, van Beers EH, Tielen IHG, et al.** Predicting BRCA1-association in non-mutated BRCA1/2 breast carcinomas using array-CGH. *(submitted)*
15. **Picard F, Robin S, Lavielle M, et al.** A statistical approach for array cgh data analysis. *BMC Bioinformatics 11;6:27. (2005)*
16. **Wessels LF, van Welsem T, Hart AA, et al.** Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for brca1 tumors. *Cancer Res. 62(23):7110-7. (2002)*
17. **Alvarez S, Diaz-Uriarte R, Osorio A, et al.** A predictor based on the somatic genomic changes of the brca1/brca2 breast cancer tumors identifies the non-brca1/brca2 tumors with brca1 promoter hypermethylation. *Clin Cancer Res. 11(3):1146-53. (2005)*
18. **Tirkkonen M, Kainu T, Loman N, et al.** Somatic genetic alterations in BRCA2-associated and sporadic male breast cancer. *Genes Chromosomes Cancer. 24(1):56-61. (1999)*
19. **Maguire P, Holmberg K, Kost-Alimova M, et al.** CGH analysis of familial non-BRCA1/BRCA2 breast tumors and mutation screening of a candidate locus on chromosome 17q11.2-12. *Int J Mol Med. 16(1):135-41. (2005).*
20 **Smith P, McGuffog L, Easton DF, et al.** A genome wide linkage search for breast cancer susceptibility genes. *Genes Chromosomes Cancer. 45(7):646-55. (2006)*
21. **Shen CY, Yu JC, Lo YL, et al.** Genome-wide search for loss of heterozygosity using laser capture microdissected tissue of breast carcinoma: an implication for mutator phenotype and breast cancer pathogenesis. *Cancer Res. 60:3884-92. (2000)*
22. **Benetkiewicz M, Piotrowski A, Díaz De Ståhl T, et al.** Chromosome 22 array-CGH profiling of breast cancer delimited minimal common regions of genomic imbalances and revealed frequent intra-tumoral genetic heterogeneity. *Int J Oncol. (4):935-45. (2006)*

23. **Tischkowitz M, Hamel N, Carvalho MA, et al.** Pathogenicity of the brca1 missense variant M1775K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach. *Eur J Hum Genet. Feb 20; [Epub ahead of print] (2008)*
24. **Couto E, Hemminki K.** Estimates of heritable and environmental components of familial breast cancer using family history information. *Br J Cancer. 96(11):1740-2. (2007)*

CHAPTER 5

LINKAGE SEARCHES

5.1. A GENOME WIDE LINKAGE SEARCH FOR BREAST CANCER SUSCEPTIBILITY GENES

Cambridge Group, Paula Smith,[1] Lesley McGuffog,[1] Douglas F. Easton,[1*] Australia Group, Graham J. Mann,[2] Gulietta M. Pupo,[2] Beth Newman,[3] Georgia Chenevix-Trench,[4] kConFab Investigators, IARC Group, Csilla Szabo,[5,6] Melissa Southey,[5] Hélène Renard,[5] Fabrice Odefrey,[5] Henry Lynch,[7] Dominique Stoppa-Lyonnet,[8] Fergus Couch,[6] John L. Hopper,[9] Graham G. Giles,[10] Margaret R. E. McCredie,[11] Saundra Buys,[12] Irene Andrulis,[13] Ruby Senie,[14] BCFS, BRCAX Collaborators Group, David E. Goldgar,[5] Netherlands Group, Rogier Oldenburg,[15] Karin Kroeze-Jansema,[16] Jaennelle Kraan,[16] Hanne Meijers-Heijboer,[15] Jan G. M. Klijn,[15] Christi van Asperen,[16] Inge van Leeuwen,[16] Hans F. A. Vasen,[16] Cees J. Cornelisse,[16] Peter Devilee,[16] United Kingdom Group, Linda Baskcomb,[17] Sheila Seal,[17] Rita Barfoot,[17] Jon Mangion,[17] Anita Hall,[17] Sarah Edkins,[17] Elizabeth Rapley,[17] Richard Wooster,[17] Jenny Chang-Claude,[18] Diana Eccles,[19] D. Gareth Evans,[20] P. Andrew Futreal,[21] Katherine L. Nathanson,[22] Barbara L. Weber,[22] the Breast Cancer Susceptibility Collaboration (UK), Nazneen Rahman,[17] Michael R. Stratton[17,21]

Genes Chromosomes Cancer. 2006 Jul;45(7):646-55.

[1] *CR-UK Genetic Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK*
[2] *Westmead Institute for Cancer Research, University of Sydney, Sydney, Australia*
[3] *Queensland University of Technology, Brisbane, Australia*
[4] *Queensland Institute for Medical Research, Brisbane, Australia*
[5] *Unit of Genetic Epidemiology, International Agency for Research, Lyon, France*
[6] *Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN*
[7] *Department of Preventive Medicine, Creighton University, Omaha, NE*
[8] *Unité de Génétique Oncologique, Institut Curie and Unité INSERM U.434, Paris, France*
[9] *Centre for Genetic Epidemiology, University of Melbourne, Melbourne, Australia*
[10] *Cancer Epidemiology Centre, Cancer Council of Victoria, Melbourne, Australia*

[11] *Department of Preventative and Social Medicine, University of Otago, Dunedin, New Zealand*
[12] *Department of Hematology-Oncology, University of Utah, Salt Lake City, UT*
[13] *Department of Laboratory Medicine, Mount Sinai Hospital, University of Toronto, Toronto, Canada*
[14] *Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY*
[15] *Erasmus University Medical Centre, Rotterdam, The Netherlands*
[16] *Leiden University Medical Centre, Leiden, The Netherlands*
[17] *Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, UK*
[18] *Division of Epidemiology, German Cancer Research Center DFKZ, Heidelberg, Germany*
[19] *Wessex Clinical Genetics Service, Princess Ann Hospital, Southampton, UK*
[20] *Regional Genetic Service, St. Mary's Hospital, Manchester, UK*
[21] *The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK*
[22] *University of Pennsylvania Cancer Centre, Philadelphia, PA*

ABSTRACT

Mutations in known breast cancer susceptibility genes account for a minority of the familial aggregation of the disease. To search for further breast cancer susceptibility genes, we performed a combined analysis of four genome-wide linkage screens, which included a total of 149 multiple case breast cancer families. All families included at least three cases of breast cancer diagnosed below age 60 years, at least one of whom had been tested and found not to carry a *BRCA1* or *BRCA2* mutation. Evidence for linkage was assessed using parametric linkage analysis, assuming both a dominant and a recessive mode of inheritance, and using nonparametric methods. The highest LOD score obtained in any analysis of the combined data was 1.80 under the dominant model, in a region on chromosome 4 close to marker D4S392. Three further LOD scores over 1 were identified in the parametric analyses and two in the nonparametric analyses. A maximum LOD score of 2.40 was found on chromosome arm 2p in families with four or more cases of breast cancer diagnosed below age 50 years. The number of linkage peaks did not differ from the number expected by chance. These results suggest regions that may harbor novel breast cancer susceptibility genes. They also indicate that no single gene is likely to account for a large fraction of the familial aggregation of breast cancer that is not due to mutations in *BRCA1* or *BRCA2*.

INTRODUCTION

Breast cancer aggregates in families, with the disease being approximately twice as common in the first-degree relatives of cases as in the general population (Collaborative Group on Hormonal Factors in Breast Cancer, [2001]). The higher risk to monozygotic twins of breast cancer cases than to dyzygotic twins of cases suggests that most of this familial clustering is likely to have a genetic basis (Peto and Mack, [2000]). However, although several important breast cancer susceptibility genes have now been identified, most of the familial aggregation of breast cancer remains unexplained.

In the 1990s, two important breast cancer susceptibility genes, *BRCA1* (MIM 113705) and *BRCA2* (MIM 600185), were identified by linkage studies in multiple case families (Miki et al. [1994]; Wooster et al. [1995]). Germline mutations in these genes confer high lifetime risks of breast cancer and ovarian cancer, together with smaller risks of some other cancer types (Antoniou et al. [2003]; Thompson and Easton, [2004]). Mutations in these genes are common in families with multiple cases of breast or ovarian cancer, and are present in most families with at least six or more cases (Ford et al. [1998]). Population-based studies have estimated that *BRCA1* and *BRCA2* mutations account for 15% of the excess familial risk of breast cancer (Peto et al. [1999]; Anglian Breast Study, [2000]; Dite et al. [2003]). Mutations in two other genes, *TP53* and *PTEN*, also confer high risks of breast cancer, but only in the context of rare syndromes. Mutations in the *ATM* and *CHEK2* genes confer more moderate (approximately twofold) risks of breast cancer (CHEK2 Case-Control Consortium, [2004]; Thompson et al. [2005]), although some mutations in *ATM* may confer higher risks. In total, the known susceptibility genes have been estimated to account for no more than 25% of the familial aggregation of breast cancer (Easton, [1999]), suggesting strongly that other susceptibility genes remain to be identified.

BRCA1 or *BRCA2* mutations are found in the majority of families with six or more cases of breast cancer cases consistent with dominant inheritance (Ford et al. [1998]). This suggests strongly that further susceptibility genes are likely to confer smaller risks than *BRCA1* and *BRCA2* mutations, but the number and characteristics of such genes remains unknown. One model, suggested by a recent segregation analysis (Antoniou et al. [2004]), proposes that there are a large number of such genes, each conferring only small risks of the disease. If true, such loci could not be identified through linkage studies. However, it is also possible that there are further loci conferring more substantial risks that could be detected by linkage. To evaluate this

possibility, we have conducted a genome-wide linkage analysis in multiple case breast cancer families that are unlikely to be segregating *BRCA1* or *BRCA2* mutations.

MATERIALS AND METHODS

As a basis for this linkage study, we sought to identify informative families with a low probability that they contained mutations in *BRCA1* or *BRCA2*. Families were collected independently by four groups, principally through family cancer clinics or epidemiological studies of breast cancer. All families were of Caucasian ancestry. The recruitment of the families used in the study took place over the last 15 years, but all families were regularly updated with regard to their cancer status. All groups obtained appropriate Institutional Review Board approvals. Specific sources of recruitment were as follows:

Australia: Families were identified through the Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer (kConFaB), which is a national multidisciplinary consortium for research on familial breast cancer (GJ Mann, unpublished). Several families were initially ascertained through the Australian Breast Cancer Family Registry (ABCFS); these kindreds were recruited as part of a population-based case-control-family study and all were recruited via a diagnosis of breast cancer in the proband under the age of 40 years (Hopper et al. [1999]).
IARC: Families were ascertained by a collaborative group of investigators from the USA, Canada, Australia, and France. Netherlands: The Dutch families were ascertained through the Clinical Genetic Centers in Leiden and Rotterdam, and through the Netherlands Foundation for the Detection of Hereditary Tumors (STOET).
United Kingdom: All but 17 of the families were ascertained through clinical genetics centers in the United Kingdom. Two families were initially ascertained in the Netherlands, six from centers in the USA, and nine from Heidelberg, Germany.

Initially, all families had to satisfy the following criteria: (1) at least three women diagnosed with breast cancer below age 60 years, all of whom were related such that they could share a single allele identically by descent, (2) no case of ovarian cancer or male breast cancer in a blood relative (since these phenotypes are strongly predictive of the presence of *BRCA1* or *BRCA2* mutation), and (3) DNA samples available for genotyping from at least three women affected with breast cancer, or from children of affected women such that the genotypes of at least three affected women might be inferred (in the latter case, at least two children of an affected women needed to be

TABLE 1

Summary of families by group

Group	Number of families							Number of genotyped individuals
		Number of cases of breast cancer				Cases of breast cancer diagnosed below age 50 years		
	Total	3	4	5	6+	<4	4+	
Australia	21	4	6	6	5	17	4	127
IARC	26	7	5	8	6	23	3	122
Netherlands	22	3	6	5	8	15	7	79
U.K.	80	25	24	18	13	70	10	395
Total	149	39	41	37	32	125	24	723

available). In addition, to minimize the probability that the family segregated a *BRCA1* or *BRCA2* mutation, DNA from at least one affected individual was screened for mutations across both genes, by a method that examined the entire coding sequence and splice junctions. Whenever possible, for families with five or more cases of breast cancer, a second affected individual was screened. Subsequently, we collected detailed information on the method of mutation screening for each family, as well as genotype data on at least three microsatellite markers flanking the *BRCA1* and *BRCA2* loci. Families with insufficient mutation screening (14 families) or linkage data (a further 6 families) were not included in further analyses. Finally, we estimated the residual probability that the index-affected individual carried a *BRCA1* and or *BRCA2* mutation, based on the assumed mutation detection sensitivity, the family history and linkage data at *BRCA1* and *BRCA2* (see statistical methods). Thirteen families, where this probability exceeded 15%, were excluded from all analyses presented here. Characteristics of the 149 families included in the analysis are summarized in Table 1.

Genotyping

To evaluate linkage to *BRCA1* and *BRCA2*, the following markers were used in various combinations in the four family sets: D17S800, D17S855, D17S951, D17S1322, D17S250 (for *BRCA1*); D13S260, D13S171, D13S1700, D13S267 (for *BRCA2*). At least three markers were analyzed at each locus in each family.

The entire coding sequences of *BRCA1* and *BRCA2* in each family were screened for mutations using several methods at the different centers. These include conformation sensitive gel electrophoresis, single strand conformational analysis, protein truncation test, DNA sequencing, and denaturing gradient gel electrophoresis. All of the Netherlands and United Kingdom and three of the IARC families were additionally screened for large deletions and insertions using deletion junction-PCR, multiplex ligation probe amplification (MLPA), or Southern analysis.

For the genome-wide linkage search, the Applied Biosystems Linkage Mapping Set MD10 was analyzed on ABI 3700 DNA sequencers, either on contract at the Australian Genome Research Facility (Australian families) or at the Wellcome Trust Sanger Institute (IARC), Netherlands and United Kingdom families). Genotypes were called automatically using Genotyper or Genemapper software and were then checked manually by at least one individual. Additional markers were used to investigate potential regions of interest in subgroups of the family set.

Statistical Analysis

To compute the residual probability that the index case carried a *BRCA1* and *BRCA2* mutation, we first used the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) model (Antoniou et al. [2004]) to calculate carrier probabilities based on the pedigree and the mutation testing that had been performed. This model allows for the effects of *BRCA1* and *BRCA2* and the combined effects of other genes in a polygenic component, and is implemented in MENDEL (Lange et al. [1988]). For this purpose, the sensitivity of mutation screening was assumed to be 70% for *BRCA1* and 80% for *BRCA2* (mutation sensitivities estimated from linked families in the Breast Cancer Linkage Consortium dataset; D.Easton, unpublished data). For samples that had been fully screened for large-scale rearrangements by MLPA, the *BRCA1* sensitivity was assumed to be 80%. The carrier probabilities were then adjusted to allow for linkage data at the *BRCA1* and *BRCA2* loci. Multipoint LOD scores were computed using Fastlink (Cottingham et al. [1993]), based on at least three markers tightly linked to each locus. The residual *BRCA1* carrier probability was then given by:

$$\frac{p_1 10^{LOD_1}}{(p_1 10^{LOD_1} + p_2 10^{LOD_2} + 1 - p_1 - p_2)}$$

and similarly for *BRCA2*, where p1 and p2 are the *BRCA1* and *BRCA2* probabilities generated by BOADICEA and LOD1 and LOD2 are the multipoint LOD scores at the *BRCA1* and *BRCA2* loci.

To conduct a combined linkage analysis, we first constructed a single linkage map incorporating all markers typed at any center. This map was based on the sex-averaged linkage map generated by deCODE (Kong et al. [2002]). For markers that were not present on the deCODE map, we interpolated their position between flanking markers, either using estimates from other linkage maps or based on their physical position in the human genome sequence relative to flanking markers. Allele frequencies for each marker were estimated by averaging over all typed individuals, separately for each center.

Evidence for linkage was assessed using both parametric and nonparametric (allele sharing approach) analyses. For the parametric analysis, we first assumed a model in which susceptibility to breast cancer is conferred by a dominant susceptibility allele with population frequency 0.003 that confers a cumulative breast cancer risk of 80% by age 80, when compared with 8% in noncarriers. This model is based on that derived from the segregation analysis of (Claus et al. [1981]) and has been used in most previous breast cancer linkage analyses. As in previous analyses, risks were modeled in seven age-categories (<30, 30-39, 40-49, 50-59, 60-69, 70-79, 80+) and implemented by using 14 liability classes, with separate classes for affected and unaffected individuals (Easton et al. [1993]). Since this model would have reduced power to detect a recessive susceptibility allele, we also analyzed the data under a recessive model. Under this model, the risks to carriers and noncarriers were identical to those under the dominant model, but the allele frequency was assumed to be 0.08. All analyses were carried out in the program GENEHUNTER (Kruglyak et al. [1996]), except for one large family (EUR60) that could not be run because the number of individuals exceeded the limits of the program, and where pruning the family would have lost a significant amount of information. For this family, analyses were run in VITESSE (O'Connell and Weeks,[1995]) for autosomes and FASTLINK (Cottingham et al. [1993]) for the X chromosome. This family separates into two distantly related branches, and these were treated as two distinct families (EUR60a and EUR60b) in the analysis.

For the GENEHUNTER analyses, multipoint LOD scores were calculated for locations at 1 cM intervals along each chromosome, using all markers for that chromosome. For the VITESSE and FASTLINK analyses, multipoint LOD scores based on every pair of adjacent markers and the disease locus were calculated. The LOD score for each

TABLE 2

Maximum LOD scores by chromosome under the dominant, recessive, and nonparametric analyses

Chromosome	Position	Model	HLOD or NPL LOD score*	
All families				
2	17	Dominant	1.21	0.18
2	16	NPL	1.10	
4	79	Dominant	1.80	0.18
5	196	Recessive	1.04	0.41
14	44	NPL	1.56	
22	41	Dominant	1.15	0.06
4+ cases dx < 50				
2	17	Dominant	2.38	0.50
4	66	Dominant	1.57	0.28
10	89	Dominant	1.12	0.35
22	41	Dominant	1.43	0.12

* HLOD, heterogeneity LOD score under dominant or recessive model; NPL, nonparametric LOD score (see Materials and Methods).

family at each position was based on an average of the LOD scores from all analyses relevant to that position. The multipoint LOD scores for each family at each position were then used to generate heterogeneity LOD scores (HLODs) based on the standard admixture model under which a certain proportion of families are assumed to be segregating a susceptibility allele at that locus (Ott, [1983]).

The nonparametric (allele sharing) analyses were conducted using the program GENEHUNTER-plus (Kong and Cox,[1997]), using the all scoring function (Whittemore and Halpern,[1994]). Analyses were conducted separately for each of the four centers and the results files combined. Nonparametric LOD scores were then generated using the program ASM, using the exponential scoring option and equal weighting of families.

RESULTS

Table 2 summarizes all linkage peaks with LOD scores greater than 1 in the combined dataset, for the whole family set and for analyses restricted to families with four or

TABLE 3

Maximum HLOD scores by chromosome and group, under the dominant model

Chrom	Australia	IARC	Netherlands	UK	TOTAL
1	0.78 (134)	0.54 (7)	0.12 (108)	0.41 (20)	0.22 (9)
2	1.07 (17)	0.57 (7)	0.28 (253)	0.68 (27)	1.21 (17)
3	0.62 (38)	0.72 (101)	0.08 (30)	0.68 (107)	0.93 (102)
4	1.02 (35)	2.02 (73)	0.08 (189)	0.92 (79)	1.80 (79)
5	0.91 (170)	0.43 (42)	0.15 (5)	0.15 (104)	0.13 (170)
6	0.65 (89)	0.72 (171)	0	0.04 (189)	0.04 (94)
7	0.37 (103)	0.19 (46)	0.03 (57)	0.86 (170)	0.20 (171)
8	0.62 (39)	0.12 (50)	0	0.35 (13)	0.07 (42)
9	0.48 (121)	0	0.70 (81)	0.07 (64)	0.01 (113)
10	0.53 (89)	0	0.16 (138)	0.04 (20)	0.12 (102)
11	0.50 (0)	0.89 (26)	0.47 (95)	0.01 (22)	0.42 (0)
12	0.57 (161)	0.15 (119)	0.43 (150)	1.04 (3)	0.43 (150)
13	0.86 (67)	0.15 (99)	0	0.19 (108)	0.0 (99)
14	0.44 (44)	0.93 (37)	0.20 (44)	0.49 (116)	0.79 (44)
15	0.0	0.11 (26)	0	1.30 (9)	0.41 (21)
16	0.13 (0)	0.14 (14)	0	0.27 (25)	0.0 (3)
17	1.17 (93)	0.09 (103)	0.12 (41)	0.64 (115)	0.60 (103)
18	0.20 (49)	0.01 (89)	0.07 (16)	0.79 (18)	0.49 (17)
19	0.0	0.2 (54)	0.0 (32)	0.21 (86)	0
20	0.34 (58)	1.40 (73)	0	2.06 (8)	0.70 (2)
21	0.84 (51)	0.54 (45)	0	0.22 (15)	0.94 (51)
22	0.03 (0)	0.01 (11)	0.28 (15)	1.52 (40)	1.15 (41)
X	0.06 (188)	0.25 (18)	0.65 (135)	0.11 (123)	0.06 (122)

* Numbers in parentheses indicate position (cM).

more breast cancer cases diagnosed below age 50 years. Tables 3-5 give the highest LOD scores for each chromosome by group, for each of the three analyses. Figure 1 gives the maximum scores for all chromosomal locations for the three analyses in the combined dataset.

In the parametric analysis under the dominant model, the highest HLOD was 1.80 on chromosome arm 4q, close to D4S392. Positive scores were obtained at this location

TABLE 4

Maximum HLOD scores by chromosome and group, under the recessive model

Chrom	Australia	IARC	Netherlands	UK	Total
1	0.18 (4)	0.70 (68)	0.29 (16)	0.64 (228)	0.37 (4)
2	0.67 (0)	0.62 (171)	0.13 (54)	0.66 (124)	0.64 (172)
3	0.17 (150)	0.0 (95)	0.09 (30)	0.51 (5)	0.07 (30)
4	0.57 (80)	0.31 (51)	0.12 (132)	0.30 (108)	0.15 (115)
5	0.70 (170)	0.26 (47)	0.44 (11)	0.88 (200)	1.04 (196)
6	0.10 (35)	0.63 (120)	0.02 (120)	0.27 (138)	0.37 (121)
7	0.25 (181)	0.14 (23)	0.95 (18)	1.03 (166)	0.25 (171)
8	0.40 (23)	0.61 (97)	0.02 (97)	0.96 (27)	0.95 (27)
9	0.37 (18)	0.02 (113)	1.36 (80)	0.19 (159)	0.06 (33)
10	0.34 (102)	0.22 (1)	0.02 (111)	0.71 (102)	0.76 (102)
11	0.27 (116)	0.19 (104)	0.03 (89)	0.27 (133)	0.23 (116)
12	0.86 (24)	0.31 (115)	1.29 (170)	0.56 (36)	0.14 (61)
13	0.16 (69)	0.16 (121)	0.0	0.12 (22)	0.0 (113)
14	0.33 (44)	0.90 (36)	0.0	0.10 (36)	0.46 (36)
15	0.02 (108)	0.29 (0)	0.28 (23)	0.21 (0)	0.08 (0)
16	0.41 (122)	1.16 (92)	0.0	0.37 (46)	0.66 (50)
17	0.38 (86)	0.20 (103)	0.15 (94)	0.74 (30)	0.55 (103)
18	0.63 (115)	0.0 (108)	0.0	0.94 (1)	0.19 (18)
19	0.0	0.0	0.65 (33)	0.69 (97)	0.03 (37)
20	0.0	0.0	0.06 (62)	0.98 (7)	0.0
21	0.01 (33)	0.12 (43)	0.03 (55)	0.08 (56)	0.08 (55)
22	0.28 (8)	0.04 (4)	0.36 (15)	0.13 (56)	0.33 (21)
X	0.32 (188)	0.88 (117)	0.37 (45)	1.27 (188)	0.63 (120)

* Numbers in parentheses indicate position (cM).

in the Australian, IARC, and United Kingdom series, but not in the Dutch dataset. Two other HLODs over 1 were found, on 2p (1.20, close to marker D2S2211) and on chromosome 22 (1.15, between D22S278 and D22S283). The latter result is predominantly due to a single family, EUR60, which includes 18 breast cancer cases and is the most informative family in the dataset. One branch of this family (EUR60b) generates a LOD score of 2.62. Seven women with breast cancer in this family, all belon-

TABLE 5

Maximum nonparametric scores by chromosome and group

Chrom	Australia	IARC	Netherlands	UK	Total
1	0.37 (89)	0.49 (75)	0.06 (171)	0.04 (271)	0.23 (83)
2	1.22 (17)	0.86 (4)	0.10 (129)	0.83 (28)	1.10 (16)
3	0.54 (112)	0.59 (100)	0.04 (30)	0.72 (106)	0.71 (102)
4	0.24 (35)	2.15 (73)	0.04 (131)	0.03 (170)	0.32 (73)
5	0.68 (169)	1.31 (47)	0.15 (9)	0.20 (205)	0.28 (205)
6	0.29 (90)	0.56 (172)	0.0	0.0	0.0
7	0.72 (104)	0.72 (46)	0.24 (22)	0.58 (122)	0.42 (122)
8	0.61 (39)	0.26 (43)	0.0	0.70 (0)	0.36 (21)
9	0.02 (45)	0.00 (34)	0.74 (81)	0.06 (66)	0.02 (31)
10	0.30 (138)	0.06 (111)	0.0	0.06 (103)	0.04 (102)
11	1.00 (117)	0.14 (152)	0.98 (95)	0.16 (130)	0.38 (133)
12	0.20 (161)	0.38 (117)	0.41 (150)	1.65 (31)	0.76 (150)
13	0.17 (76)	0.43 (106)	0.00 (99)	0.23 (109)	0.12 (109)
14	0.68 (98)	0.69 (44)	0.71 (37)	0.21 (37)	1.56 (44)
15	0.0	0.02 (15)	0.02 (23)	0.99 0	0.0
16	0.24 (50)	0.13 (51)	0.0	0.33 (101)	0.09 (85)
17	0.75 (93)	0.57 (55)	0.03 (86)	0.62 (29)	0.19 (5)
18	0.23 (121)	0.74 (18)	0.11 (99)	1.31 (14)	0.85 (15)
19	0.08 (24)	0.04 (63)	0.17 (32)	0.40 (86)	0.01 (16)
20	0.10 (60)	1.34 (74)	0.0	2.22 (8)	0.46 (0)
21	0.21 (34)	0.02 (43)	0.00 (55)	0.0	0.0
22	0.26 (3)	0.15 (3)	0.65 (15)	0.00 (50)	0.00 (0)
X	0.32 (141)	0.67 (116)	0.51 (28)	0.01 (120)	0.35 (137)

* Numbers in parentheses indicate position (cM).

ging to branch EUR60b, carry the *CHEK2 1100delC* variant (Meijers-Heijboer et al. [2002]). When both branches of this family were removed, the maximum HLOD on chromosome 22 reduced to 0.06.

When analyses were restricted to families with at least four cases of breast cancer diagnosed below age 50, the maximum HLOD on 2p rose to 2.38. HLODs over 1 in this

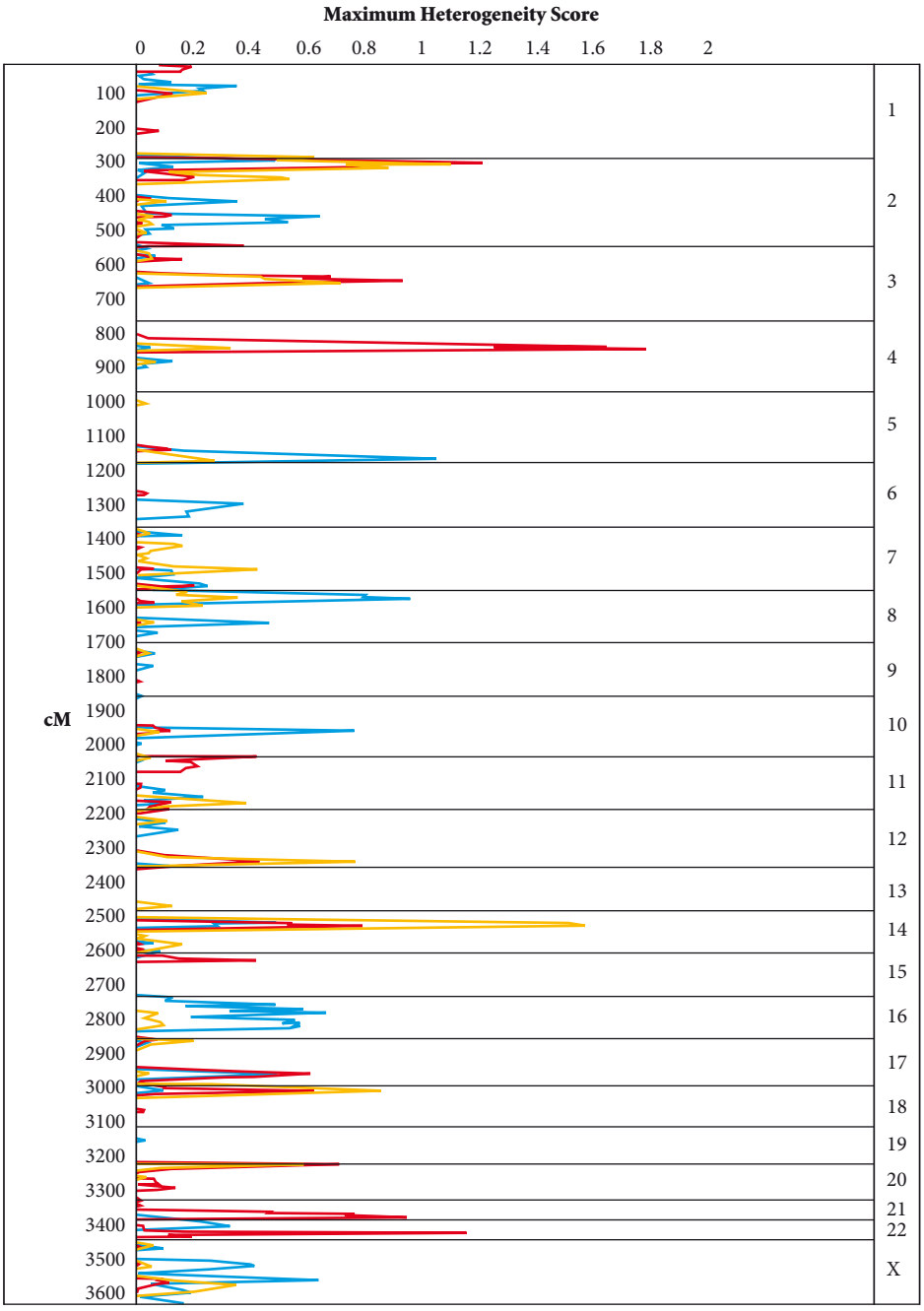


Figure 1.

Maximum HLOD's by location for the dominant model (red line), the recessive model (blue line), and maximum nonparametric LOD scores (yellow line).

TABLE 6

LOD scores greater than 1.5 in individual families, under the dominant model

Study center	Family	Chromosome	Position	LOD score
Australia	699003	2	67	1.67
IARC	2191	4	61	1.84
IARC	2191	20	70	1.80
IARC	MAYO151	3	95	1.52
IARC	MAYO151	11	43	1.59
Netherlands	RUL153	11	88	1.67
UK	EUR60a	15	24	1.50
UK	EUR60b	4	79	1.91
UK	EUR60b	22	41	2.62

subset were also found on chromosomes 4 and 22 close to the peaks in the overall analysis, and a further peak on chromosome 10 (HLOD 1.12) was also identified.
In addition to the aforementioned loci, an HLOD of 2.06 was found on 20p (at 8 cM) in the United Kingdom family set. There was, however, no evidence of linkage in the families from the other groups. LOD scores greater than 1.5 in individual families are summarized in Table 6. Of the eight scores, three contribute to the linkage peaks on chromosomes 2 and 4 found in the overall dataset. In addition two families showed linkage on chromosome 11. These peaks were however separated by over 40 cM and there was no evidence of linkage to this region in the overall analysis.
In analyses under a recessive model, only one locus reached a HLOD of greater than 1 (1.04 on 5q). In the nonparametric analysis, the highest peak was on chromosome 14 (LOD 1.56 at position 43). The only other LOD over 1 was on chromosome 2 (LOD 1.10, position 16), almost coincident with the peak in the analysis under the dominant model.

DISCUSSION

The analyses of 149 families reported here represented by far the largest genome-wide linkage screen for breast cancer susceptibility loci. The only other report since the identification of *BRCA1* and *BRCA2* was that by (Huusko et al. [2004]), who studied 14 *BRCA1/2* negative breast cancer families from Finland. Other reports have

examined specific loci on chromosome arms 6q, 8p, and 13q (Zuppan et al. [1991]; Kerangueven et al. [1995]; Seitz et al. [1997]; Kainu et al. [2000]; Rahman et al. [2000]; Thompson et al. [2002]).

The rationale for the genome-wide linkage searches is that there exist further breast cancer genes in which alleles confer high risks. The pattern of familial risks indicates that such alleles are likely to be dominant, and we therefore considered the parametric analysis assuming a dominant model to be the primary analysis. To provide some protection against model misspecification, we also conducted analyses under a recessive model and using an allele sharing approach. These approaches, however, identified no further strong linkage signals.

Under the dominant model, we found three regions with HLODs in excess of 1, but none with HLODs over 2. Of these linkage peaks, one on chromosome 22 is explained entirely by a single family (EUR60). This family is the most informative in the study, containing 18 breast cancer cases. Seven cases of breast cancer have been shown to carry the *CHEK2* variant 1100delC (Miejers-Heijboer et al. [2002]). Since *CHEK2* is located on chromosome 22, one might hypothesize that the linkage signal is a reflection of the segregation of this variant. However, the breast cancer risk conferred by *CHEK2* 1100delC is only twofold, and this would not be expected to generate strong linkage evidence. Furthermore, the LOD score in the larger branch of EUR60 at *CHEK2* itself is only 0.3. Thus, it remains unclear whether the linkage signal on chromosome 22 reflects the effect of *CHEK2 1100delC* together with chance segregation, or whether there is an additional susceptibility allele segregating in this family. If the latter is true, given the lack of any linkage evidence from other families, susceptibility alleles at this other locus must be rare.

The strongest linkage signal in our set was found on the short arm of chromosome 4. This score was also, in part, due to EUR60 (LOD score 1.91 in the larger branch), although some evidence of linkage remained when EUR60 was excluded. The third linkage peak was on 2p (HLOD 1.2). This evidence increased (HLOD 2.4) when analyses were restricted to families with at least four cases of breast cancer diagnosed below age 50 years.

Huusko et al. ([2004]) reported evidence for linkage to markers on 2q32 in 14 Finnish breast cancer families, with a maximum LOD score of 3.20 close to D2S2262. We found no evidence of linkage in this region (maximum HLOD under the dominant model 0.0, = 0.0; nonparametric LOD = 0.05). Huusko et al. ([2004]) found one other LOD score over 1 under a dominant model, at D9S283 (1.12). Again, we found no evidence for linkage in this region. Similarly, Zuppan et al. ([1991]) found

evidence of linkage to the estrogen receptor gene on 6q in two families. In our study, we found no evidence of linkage to this region (HLOD = 0 for both the dominant and recessive models).

Theoretical calculations indicate that, for a fully informative marker map, the expected number of regions with LOD scores of greater than 1 and 1.5 will be 5 and 2, respectively (Lander and Kruglyak,[1995]). These predictions are not strictly comparable to our analyses, since our marker sets are not fully informative. Nevertheless, they indicate that the number of linkage peaks is not clearly in excess of the number that might be expected by chance and, therefore, that the observed peaks may reflect the play of chance rather than true susceptibility loci.

Under the admixture model, the estimated proportion of families linked to the loci are 0.18, 0.18, and 0.06 for chromosomes 2, 4, and 22, respectively. Such estimates can be misleading, since they are highly dependent on the genetic model that is assumed, and the true model is unknown. However, they indicate that, even if one or more of these linkage peaks is ultimately shown to harbor a true susceptibility locus, its contribution to the familial aggregation of breast cancer is likely to be modest. Moreover, under the assumed parametric dominant model, 87% of the genome achieved an HLOD of -1 or lower if the proportion of linked families was assumed to be 0.3, and 66% of the genome achieved an HLOD <-2, indicating that such a locus was unlikely to have been missed elsewhere in the genome.

The failure to detect strong linkage signals might reflect extensive locus heterogeneity, whereby the disease is only linked to a particular locus in a small proportion of families. Under this scenario, greater power might be achievable by considering subsets of families from more homogeneous populations where genetic heterogeneity might be reduced. We were able to examine this to a limited extent by performing separate analyses of the families in each of the four study sets. Since the Australian families were largely of British and Irish origin, these two groups might be considered comparable. The Dutch population exhibits distinct founder mutations for many diseases and this group is, to an extent, genetically distinct, while the IARC families originated from many sources and are genetically heterogeneous. In the event, no strong linkage signals were observed either in the Dutch set or in the combined United Kingdom/Australian set. In particular, the linkage peaks identified in family EUR60 were not supported by linkage evidence in other Dutch families. The linkage peak on chromosome 2 did, however, become somewhat stronger when the Dutch families were excluded.

The failure to detect strong evidence for linkage may also reflect disease heterogene-

ity. Recent studies have demonstrated that breast tumors can be categorized into groups on the basis of CGH profiles and expression patterns, and that these patterns differ between *BRCA1*, *BRCA2*, and non-*BRCA1/2* familial breast cancer (Hedenfalk et al. [2001],[2003]; Gronwald et al. [2005]; Macguire et al. [2005]). These observations raise the possibility that mutations in other breast cancer susceptibility genes are associated with distinct tumor profiles. If so, incorporating tumor characteristics into the analyses could identify linkage signals that are not evident using breast cancer as a whole as the disease end point.

The positive signals found in this study indicate the most promising locations for further high-risk susceptibility genes, and would be worth following up in further families. Our results also indicate, however, that many genes are likely to be involved in breast cancer predisposition, with no gene accounting for a large fraction of the familial aggregation, and that alternative strategies will probably be necessary to identify them.

ACKNOWLEDGEMENTS

The authors would like to thank all of the families for their participation in this study. For the Australian study, the authors would like to thank the kConFab research nurses and staff for data collection, Heather Thorne, Lynda Williams, and Dani Surace for DNA preparation, Jan Groves for establishment of the LCL's, Eveline Niedermayr and Sandra Picken for supplying data, the staff of the Familial Cancer Clinics for their support of kConFab. B.A. Oostra helped out with the second phase genotyping of the Dutch study. The Breast Cancer Susceptibility Collaboration (United Kingdom) consists of the following contributors: A. Ardern-Jones, J. Berg, A. Brady, C. Brewer, G. Brice, B. Bullman, R. Cetnarsryj, C. Chapman, C. Chu, N. Coates, T. Cole, R. Davidson, A. Donaldson, H. Dorkins, F. Douglas, D. Eccles, R. Eeles, F. Elmslie, D.G. Evans, S. Goff, D. Goudie, J. Gray, L. Greenhalgh, H. Gregory, N. Haites, S.V. Hodgson, T. Homfray, R.S. Houlston, L. Izatt, L. Jeffers, V. Johnson-Roffey, F. Lalloo, M. Longmuir, J. Mackay, A. Magee, S. Mansour, Dr. Zosia Miedzybrodzka, J. Miller, P. Morrison, V. Murday, J. Paterson, M. Porteous, N. Rahman, K. Redman, M. Rogers, S. Rowe, A. Saggar, A. Schofield, L. Side, M. Steel. Anita Hall and Elizabeth Mackie supported the genotyping of the United Kingdom set. Families from the IARC group were contributed in part through the Breast Cancer Family Registry (BCFR), which is supported through collaborative agreements with The University of Melbourne, Cancer Care Ontario, the Huntsman Cancer Institute and Columbia University. DFE is a Principal Research Fellow of Cancer Research UK.

REFERENCE LIST

- **Anglian Breast Cancer Study Group.** Prevalence and penetrance of *BRCA1* and *BRCA2* mutations in a population- based series of breast cancer cases. *Br J Cancer 83: 1301-1308 (2000)*
- **Antoniou A, Pharoah PD, Narod S, et al.** Average risks of breast and ovarian cancer associated with *BRCA1* or *BRCA2* mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet 72: 1117-1130 (2003)*
- **Antoniou AC, Pharoah PDP, Smith P, Easton DF.** The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer 91: 1580-1590. (2004)*
- **The CHEK2 Breast Cancer Case-Control Consortium.** CHEK2*1100delC and susceptibility to breast cancer: A collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from ten studies. *Am J Hum Genet 74: 1175-1182 (2004)*
- **Claus EB, Risch N, Thompson WD.** Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet 48: 232-242. (1991)*
- **Collaborative Group on Hormonal Factors in Breast Cancer.** Familial breast cancer: Collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet 358: 1389-1399. (2001)*
- **Cottingham RW, Jr, Idury RM, Schaffer AA.** Faster sequential genetic linkage computations. *Am J Hum Genet 53: 252-263. (1993)*
- **Dite GS, Jenkins MA, Southey MC, et al.** Familial risks, early-onset breast cancer, and *BRCA1* and *BRCA2* germline mutations. *J Natl Cancer Inst 95: 448-457 (2003)*
- **Easton DF.** How many more breast cancer predisposition genes are there? *Breast Cancer Res 1: 14-17 (1999)*
- **Easton DF, Bishop DT, Ford D, Crockford GP.** 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)*
- **Ford D, Easton DF, Stratton M, et al. the Breast Cancer Linkage Consortium.** Genetic heterogeneity and Penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. *Am J Hum Genet 62: 334-345 (1998)*
- **Gronwald J, Jauch A, Cybulski C, et al.** Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. *Int J Cancer 20: 230-236 (2005)*
- **Hedenfalk I, Duggan D, Chen Y, et al.** Gene-expression profiles in hereditary breast cancer. *N Engl J Med 22: 539-548 (2001)*
- **Hedenfalk I, Ringner M, Ben-Dor A, et al.** Molecular classification of familial non-*BRCA1*/*BRCA2* breast cancer. *Proc Natl Acad Sci USA 100: 2532-2537 (2003)*
- **Hopper JL, Chenevix-Trench G, Jolley DJ, et al.** Design and analysis issues in a population-based, case-control-family study of the genetic epidemiology of breast cancer and the Co-operative Family Registry for Breast Cancer Studies (CFRBCS). *J Natl Cancer Inst Monogr 26: 95-100 (1999)*
- **Huusko P, Juo SH, Gillanders E, et al.** Genome-wide scanning for linkage in Finnish breast cancer families. *Eur J Hum Genet 12: 98-104 (2004)*
- **Kainu T, Juo SH, Desper R, et al.** Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proc Natl Acad Sci USA 97: 9603-9608 (2000)*
- **Kerangueven F, Essioux L, Dib A, et al.** Loss of heterozygosity and linkage analysis in breast carcinoma: indication for a putative third susceptibility gene on the short arm of chromosome 8. *Oncogene 10: 1023-1026 (1995)*

- **Kong A, Cox NJ.** Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet 61: 1179-1188 (1997)*
- **Kong A, Gudbjartsson D, Sainz J, et al.** A high-resolution recombination map of the human genome. *Nat Genet 31: 241-247 (2002)*
- **Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES.** Parametric and nonparametric linkage analysis: A unified multipoint approach. *Am J Hum Genet 58: 1347-1363 (1996)*
- **Lander E, Kruglyak L.** Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat Genet 11: 241-247 (1995)*
- **Lange K, Weeks D, Boehnke M.** Programs for Pedigree Analysis: MENDEL, FISHER, and dGENE. *Genet Epidemiol 5: 471-472 (1988)*
- **Maguire P, Holmberg K, Kost-Alimova M, et al.** CGH analysis of familial non-*BRCA1*/*BRCA2* breast tumors and mutation screening of a candidate locus on chromosome 17q11.2-12. *Int J Mol Med 16: 135-141 (2005)*
- **Meijers-Heijboer H, van den Ouweland A, Klijn J, et al.** Low-penetrance susceptibility to breast cancer due to CHK2 1100delC in non-carriers of *BRCA1* or *BRCA2* mutations. *Nat Genet 31: 55-59 (2002)*
- **Miki Y, Swensen J, Shattuck-Eidens D, et al.** A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science 266: 66-71 (1994)*
- **O'Connell JR, Weeks DE.** The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. *Nat Genet 11: 402-408 (1995)*
- **Ott J.** Linkage analysis and family classification under heterogeneity. *Ann Hum Genet 47: 311-320 (1983)*
- **Peto J, Mack TM.** High constant incidence in twins and other relatives of women with breast cancer. *Nat Genet 26: 411-414 (2000)*
- **Peto J, Collins N, Barfoot R, et al.** Prevalence of *BRCA1* and *BRCA2* gene mutations in patients with early- onset breast cancer. *J Natl Cancer Inst 91: 943-949 (1999)*
- **Rahman N, Teare MD, Seal S, et al.** Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. *Oncogene 19: 4170-4173 (2000)*
- **Seitz S, Rohde K, Bender E, et al.** Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: Linkage analysis in German breast cancer families. *Oncogene 14: 741-743 (1997)*
- **Thompson D, Easton D.** The genetic epidemiology of breast cancer genes. *J Mammary Gland Biol Neoplasia 9: 221-236 (2004)*
- **Thompson D, Duedal S, Kirner J, et al.** Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst 97: 813-822 (2005)*
- **Thompson D, Szabo CI, Mangion J, et al.** Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. *Proc Natl Acad Sci USA 99: 827-831 (2002)*
- **Whittemore AS, Halpern J.** A class of tests for linkage using affected pedigree members. *Biometrics 50: 118-127 (1994)*
- **Wooster R, Bignell G, Swift S, et al.** Identification of the breast cancer susceptibility gene *BRCA2*. *Nature 378: 789-792 (1995)*
- **Zuppan P, Hall JM, Lee MK, et al.** Possible linkage of the estrogen receptor gene to breast cancer in a family with late-onset disease. *Am J Hum Genet 48: 1065-1068 (1991)*

5.2. GENOME-WIDE LINKAGE SCAN IN DUTCH HEREDITARY NON-BRCA1/2 BREAST CANCER FAMILIES IDENTIFIES 9q21-22 AS A PUTATIVE BREAST CANCER SUSCEPTIBILITY LOCUS

Rogier A. Oldenburg,[1,2,] Karin H.G. Kroeze-Jansema,[1] Jeanine J. Houwing-Duistermaat,[6] Jean-Pierre Bayley,[1] Cheryl Dambrot,[1] Christi J. van Asperen,[1] Ans M.W. van den Ouweland,[2] Bert Bakker,[1] Erik H. van Beers,[7] Petra M. Nederlof,[8] Hans Vasen,[4] Nicoline Hoogerbrugge,[3] Cees J. Cornelisse,[5] Hanne Meijers-Heijboer[8] and Peter Devilee[1,5]

Submitted

[1] *Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden*
[2] *Dept. of Clinical Genetics, Erasmus Medical center, Rotterdam*
[3] *Dept. of Human Genetics, University Medical Center Nijmegen*
[4] *Netherlands Foundation for the Detection of Hereditary Tumors, Leiden*
[5] *Dept. of Pathology, Leiden University Medical Center, Leiden*
[6] *Dept. of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden*
[7] *Division of Experimental Therapy, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam*
[8] *Department of Pathology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands*

ABSTRACT

Breast cancer accounts for 20% of all female cancers. Many risk factors have been identified but a positive family history remains one of the most important risk factors, with first-degree relatives of patients having a 2-fold elevated risk. Known breast cancer susceptibility genes such as *BRCA1* and *BRCA2* explain only 20-25% of this risk, suggesting the existence of other breast cancer susceptibility genes.

Here we report the results of a genome-wide linkage scan in 55 high-risk Dutch breast cancer families with no mutations in *BRCA1* and *BRCA2*. In addition we performed CGH-analyses in 61 tumors of these families and 31 sporadic tumors.

Twenty-two of these cancer families were also included in the previous linkage study by the Breast Cancer Linkage Consortium.[1] Three regions were identified with parametric HLOD scores >1, and three with non-parametric LOD scores >1.5. Upon

further marker genotyping for the candidate loci, and the addition of another 30 families to the analysis, only the locus on chromosome 9 (9q21-22, marker D9S167) remained significant, with a non-parametric multipoint LOD score of 3.96 (parametric HLOD 0.56, α=0.18). With CGH-analyses we observed preferential copynumber loss at BAC RP11-276H19, containing D9S167 in familial tumors as compared to sporadic tumors (P<0.001). Five candidate genes were selected from the region around D9S167 and their coding regions subjected to direct sequence analysis in 16 probands. No clear pathogenic mutations were found in any of these genes.

INTRODUCTION

Breast cancer is the most commonly occurring cancer among women, accounting for 22% of all female cancers and the cumulative lifetime risk for a woman to develop breast cancer is approximately 1 in 10.[2] Many risk factors have been identified but a positive family history remains among the most important ones established for breast cancer, with first-degree relatives of patients having an approximately 2-fold elevated risk.[3] This risk increases with the number of affected relatives and is greater for women with relatives affected at a young age, bilateral disease or a history of benign breast disease.[4,5] It is currently estimated that approximately 20-25% of this risk is explained by known breast cancer susceptibility genes, mostly those conferring high risks, such as *BRCA1* and *BRCA2*.[6] This suggests that other susceptibility genes remain to be found, although it is not entirely clear which genetic model explains the remainder of familial risk best.[7-9] Depending on the population investigated, some of the risk could still be due to rare, moderately penetrant autosomal dominant effects, a common recessive effect, or a polygenic model. Recently the Breast Cancer Linkage Consortium (BCLC) published the results of a genome-wide linkage search for new breast cancer susceptibility genes in 149 high risk breast cancer families.[1] The highest LOD score obtained was 1.80 under the dominant model, for a region on chromosome 4. A maximum heterogeneity-LOD (HLOD) score of 2.40 was found on chromosome arm 2p in a subset of families with four or more cases of breast cancer diagnosed below age 50 years.[1] Other studies scanning for linkage were also unable to detect significant lod scores, but were much smaller in terms of number of families included.[10,11]

The failure to detect strong linkage signals might reflect extensive locus heterogeneity, whereby multiple susceptibility loci each explain only a small proportion of families. Greater statistical linkage power might be achieved by considering subsets of

families from more homogeneous populations in which the number of such loci might be reduced. We have here performed a search for linkage in a set of 55 breast cancer families of Dutch origin that are unlikely to be segregating BRCA1 or BRCA2 mutations. The cumulative lifetime risk of developing breast cancer in the Netherlands is about 1 in 9 women, which ranks among the highest worldwide. Founder effects at several major breast cancer loci have been detected in the Dutch population,[12-14] as well as for many other disease genes. The assumption of reduced genetic heterogeneity for breast cancer susceptibility in the Netherlands is therefore not unrealistic. No significant lod scores were obtained in parametric analyses under a dominant or recessive model. Non-parametric (allele-sharing) analysis identified a locus on chromosome 9q21 with a multipoint NPL-score of 3.96 (marker D9S167), but no clearly pathogenic mutations were detected in 5 candidate genes flanking this marker in 16 probands from families putatively linked to chromosome 9.

MATERIALS AND METHODS

Family collection

The families were ascertained through the Clinical Genetic Centers in Leiden, Rotterdam, and Nijmegen, as well as through the Netherlands Foundation for the Detection of Hereditary Tumors (STOET). The families were eligible for inclusion if there were at least three cases diagnosed with breast cancer before the age of 60, no cases of ovarian cancer, and no cases of male breast cancer.[1] Polymorphic marker information had to be retrievable for at least three cases under 60, either by direct genotyping of blood samples, or by inferring from genotyped spouses and children. The resulting 55 families constituted our 'linkage search group' (208 genotyped breast cancer cases). Twenty-two of these families were also included in the genome-wide linkage search conducted by the BCLC.[1] Another 30 families (119 breast cancer cases), were designated 'linkage conformation group', because they were selected on the same cancer phenotype, but differed slightly from the search group in that they did not meet the genotype or age of onset criteria. Thus, there were 4 families with two genotyped cases diagnosed before the age 60, and one diagnosed at the age 60, 18 families with three or more cases under 60, of which only two were genotyped, 3 families with three or more cases under 60, of which only one was genotyped, and 5 families with two genotyped cases diagnosed before 60, one of whom is a bilateral case (with both primaries diagnosed before age 60). To meet the 'three cases' sampling criteria in these families, we also genotyped cases diagnosed above 60 if they had donated a blood sample (67 breast cancer cases).

Paraffin-embedded tumor samples and pathological reports or medical reports were retrieved where available. Blood samples were collected after obtaining written informed consent. The institutional ethical committees of all of the hospitals involved approved this study.

BRCA1 and BRCA2 mutation testing

In each family, the youngest breast cancer patient from whom a blood sample was available was tested for mutations in the *BRCA1* and *BRCA2* genes (and for many families the next youngest as well). The joint Clinical Genetic Centers applied a variety of methodologies. The largest central exons (exon 11 in *BRCA1* and *BRCA2,* exon 10 of *BRCA2*) were scanned by protein truncation tests.[15] The small exons were scanned for mutations by denaturating gradient gel electrophoreses (DGGE) or direct sequencing. All of the laboratories specifically assayed the presence of large founder deletions in *BRCA1* by deletion junction-PCR.[13] The entire coding sequences of *BRCA1* and *BRCA2* were investigated by conformation-sensitive gel electrophoresis (CSGE) in families that were incompletely scanned at the time of ascertainment.[16] Since 2002, each center offers full sequence analysis and DGGE covering the entire coding regions of both genes, and Multiple Ligation-dependent Probe Amplification (MLPA) to detect large deletions/duplications in *BRCA1*.[17]

Linkage analysis

For the genome-wide linkage search, the Applied Biosystems Linkage Mapping Set MD10, consisting of 416 microsatellite markers at ~10 cM average spacing, was analyzed on an ABI3700 DNA sequencer.[1] Additional markers were used to investigate the region of interest on chromosome 9. Genotypes were called automatically using Genemapper software and were then checked manually by two individuals. DNA from CEPH 1347-02 was typed as reference to ensure consistency of allele sizing. Allele frequencies for parametric linkage analyses were calculated based on one randomly chosen individual from each family. Multipoint linkage analyses were carried out using the program GENEHUNTER version 2.1-B.[18] We used a model in which susceptibility to breast cancer is conferred by a dominant allele with a reduced penetrance and a population frequency of 0.003.[19,20] The risk of breast cancer by age 80 was assumed to be 0.85 in carriers and 0.096 in non-carriers. Risks are modeled in seven age categories (<30, 30-39, 40-49, 50-59, 60-69, 70-79, and 80+) as described.[20] Under the recessive model, the risk to carriers and noncarriers were identical to those under the dominant model, but the disease allele frequency was assumed to

be 0.08. We used the multipoint LOD-scores for each family to compute heterogeneity LOD scores, using the standard admixture model, and hence estimated the proportion of families (α) linked to the putative '*BRCAX*' locus by maximizing the heterogeneity LOD score. Non-parametric linkage analyses were carried out by the program GENEHUNTER version 2.1-B and MERLIN version 0.9.12b.[21] Both the singlepoint and multipoint settings were used, as well as both the 'pairs' and 'all' setting.

SNP-genotyping

Four single nucleotide polymorphisms were initially selected from an approximately 65-kb region surrounding D9S167. More recently, however, these SNP's were repositioned 1.74 Mb distal of this marker by the human genome sequencing effort. We then selected 4 other SNP's, immediately adjacent to D9S167 and covering about 17 kb within the 48-kb linkage disequilibrium-block around marker D9S167 (www.hapmap.org). These were rs12335588 (hapmap position 82,996,423), rs10867942 (83,002,124), rs11139937 (83,011,568), and rs11139938 (83,011,664). Marker D9S167 is at hapmap position 83,013,562. Primers were designed in such a way that the polymorphism would create or destroy a restriction site.22 PCR-products were digested by the appropriate restriction enzyme, and analysed on a 2.5% agarose gel. Results were scored by two observers independently. Data from all 8 SNP's were used to reconstruct haplotypes around D9S167.

Chromosome 9 copy number analysis

We performed array-CGH-analysis of 61 paraffin-embedded tumor samples from 58 patients from 27 families, using a method described previously.[23,24] Similar material from 31 sporadic cases served as control. These arrays contain approximately 3,500 BAC clones, of which 13 derive from an 8-cM region of interest on chromosome 9. The BAC's were considered to report copy number gain if the ratio of tumor derived genomic DNA compared to normal DNA exceeded 0.2 on a ^{2}log-scale, and copy number loss if the signal was below −0.2. The full dataset describing gains and losses on all chromosomes in this patient material will be described elsewhere (Van Beers et al., manuscript in preparation).

Sequence analysis of candidate genes

All known genes in an 14-cM interval D9S175-D9S167-D9S283 were retrieved from Ensemble (release 42). The cellular functions of these genes – in as much as they were known – were retrieved from OMIM. A literature search was then performed by

TABLE 1

Maximum LOD scores in 55 breast cancer families by chromosome under the dominant, nonparametric and recessive model.

Chromosome	Position	Model	HLOD or NPL LOD score	Alpha
1	226	Dominant	1.40	0.12
4	64	NPL	1.26	
6	164	NPL	1.94	
6	164	Recessive	1.52	0.53
7	90	NPL	1.25	
9	30	NPL	2.22	
9	86	NPL	2.34	
9	88	Recessive	1.18	0.47
9	90	Dominant	1.24	0.43
15	114	Dominant	1.19	0.29
15	114	NPL	1.12	
21	22	NPL	1.72	

Generated with the Genehunter software package. Position (cM) based on deCode map.
HLOD, heterogeneity LOD score (dominant and recessive models);
NPL, nonparametic (allele sharing) LOD score.
Alpha is the proportion of linked families in the admixture model.

a computer-program dubbed 'Anni', which can find functional associations between large numbers of genes and other biomedical concepts (in this case 'breast cancer') from free-text literature.[25] For each gene, a profile of related concepts was constructed that summarizes the context in which the gene is mentioned in literature. In addition, all genes from the region were analyzed by software termed 'Prioritizer'.[26] On this basis, 5 genes (of the 14 annotated genes with a known function in an approx 5-cM region around D9S167) were selected for direct sequence analysis in a set of 16 DNA samples from breast cancer patients from 16 different families. These families were selected because analysis of genotype data with the program 'Haploview'[27] had indicated that all patients share a haplotype in this region. Candidate genes were analyzed by DNA sequence analysis on the ABI3730 Prism Genetic Analyzer (Applied Biosystems, Foster City, CA)and the Mutation Surveyor® software package.

TABLE 2

Haplotypes around D9S167

	All families		Complete sharing		Near complete sharing[3]		
D9S167 allele (bp)	Number of families[1]	Number of haplotypes[2]	Number of families	Number of haplotypes	Number of families	Number of haplotypes	Total nr. of haplotypes[4]
313	11	7	2	1	2	2	2
317	25	19	3	2	3	3	4
319	20	11	2	2	1	1	3
321	29	22	6	6	2	2	7
323	9	8	2	1	0	0	1
325	15	12	3	2	0	0	2
327	5	7	1	1	1	1	1
329	2	2	0	0	1	1	1
331	9	9	1	1	0	0	1
333	3	3	0	0	1	1	1
335	13	14	4	4	1	1	5
337	3	3	1	1	0	0	1
Totals			25	21	12	12	29

[1] Total number of families (sharing and non-sharing)

[2] Total number of different haplotypes in the complete set of families

[3] Families with > 80% sharing, excluding families with complete sharing

[4] Total number of different haplotypes in families with > 80% sharing

RESULTS

Genome-wide linkage scan

We performed a genome-wide linkage analysis with 416 microsatellite markers, with an average spacing of approximately 10 cM in the group of 55 linkage search families. The highest heterogeneity LOD-score generated by GENEHUNTER under a parametric dominant model was 1.40 on chromosome 1 at position 226 cM (Table 1, Figure 1). Two other regions with HLOD-scores greater than 1.0 were identified on chromosome 9 (HLOD=1.23 at position 90 cM) and 15 (HLOD=1.19 at position 114 cM). Under a recessive model HLOD-scores >1.0 were found on chromosome 6 (164 cM) and 9 (88 cM). With non-parametric linkage analysis (NPL) we identified seven regions with a NPL-score higher than 1 (chromosome 4, 6, 7, 9, 15 and 21). The highest NPL-score found was on chromosome 9 (NPL=2.34, 86 cM, P=0.015). A second, distinct region on chromosome 9 had an NPL-score of 2.23 (30 cM, P=0.019). This second region also showed a HLOD-score >1 under the dominant model. To

evaluate these linkage signals further we genotyped an additional 30 families (confirmation group) for the microsatellite markers on chromosome 1, 6, 9, 15, and 21 at which the peak LOD scores were observed. For all these loci the evidence for linkage decreased, except for the locus on chromosome 1, for which the HLOD increased to 1.46 (α=0.13, P=0.39). The locus at position 86 cM on chromosome 9 decreased only slightly (NPL= 1.98, P=0.028) and the HLOD was 0.56 (α=0.18).

We then also generated LOD-scores using the method of Kong and Cox with the MERLIN software package, because this method is less conservative when marker information is not complete (Figure 1). The multipoint LOD-score at position 86 cM on chromosome 9 was 3.06 ($P=9x10^{-5}$) at marker D9S167 in the 55 families, and increased to 3.96 ($P=10^{-5}$) when the other 30 families were added to the analysis. The single-point LOD-score over all 85 families for D9S167 was 4.63 ($P=10^{-6}$). To evaluate this region on chromosome 9 further we genotyped 4 additional microsatellite markers, i.e., D9S1843 and D9S1674 proximal of D9S167, and D9S1865 and D9S1812 distal of it, defining a 9.4 cM-region. With these additional markers the multipoint LOD-score calculated by MERLIN at D9S167 in the 85 families declined to 3.02, while those at D9S1843 and D9S287 were below 1.5 (Figure 2). This defined the linked region to be between the markers D9S1674 and D9S287 (~15.7 cM).

Haplotype analysis

To aid haplotyping around D9S167, we genotyped 8 SNP's, 4 of which immediately proximal of D9S167 in a ~48-kb LD-block, and 4 covering a 65-kb region about 1.7 Mb distal of D9S167. We analyzed haplotype-sharing in each family with the program 'Haploview'.[27] In 32 families all genotyped patients shared an allele at D9S167, but in 5 families this allele was on a different haplotype, indicating that the shared alleles were not identical by descent (IBD). In another four families not all patients were succesfully genotyped at D9S167, but in two of those the patients shared a haplotype from D9S1674 to D9S1812, suggesting allele-sharing at D9S167. Thus, in 25 families all patients shared an allele at D9S167 IBD (Table 2), in total comprising 10 different alleles (range: 1 – 6 families per allele) on 21 different haplotypes. In 50 families there was no sharing of an allele among genotyped patients (in 3 families the marker data did not allow phasing of the haplotypes). In 12 of these 50 families, more than 80% of the patients shared the same haplotype (4 out of 5 patients in 5 families, 5 out of 6 in 4 families, and 6 out of 7, 7 out of 8, 8 out of 10 in 1 family each). Again, 8 different D9S167-alleles were shared on a total of 12 different haplotypes (Table 2). Overall, the 12 different D9S167-alleles found to be shared either

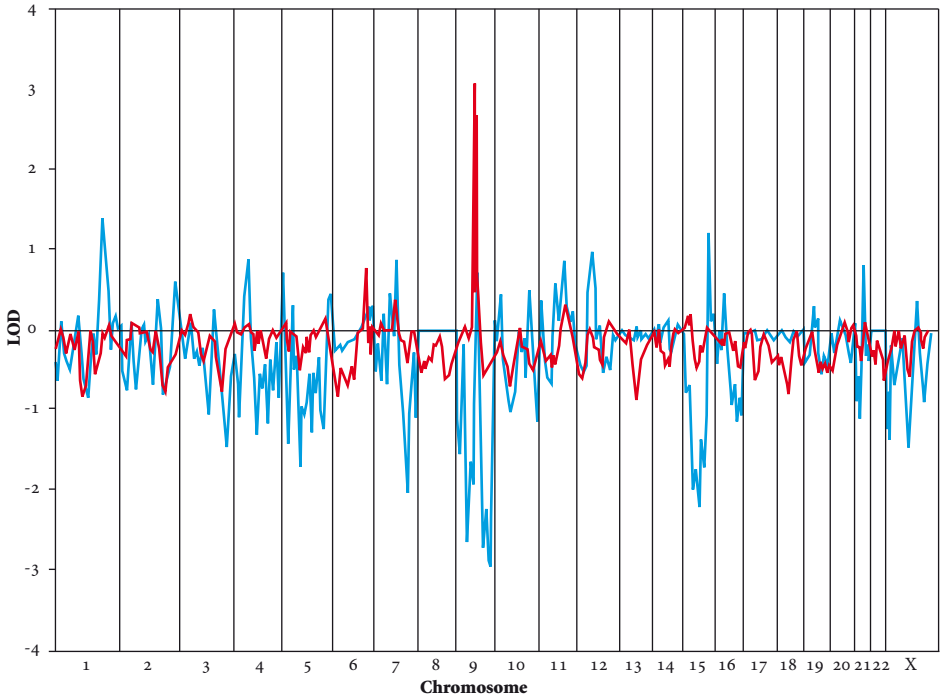


Figure 1. HLOD scores by chromosome for the dominant model (blue line), as computed by GENEHUNTER, and nonparametric LOD scores (red line), as computed by MERLIN, in 55 breast cancer families.

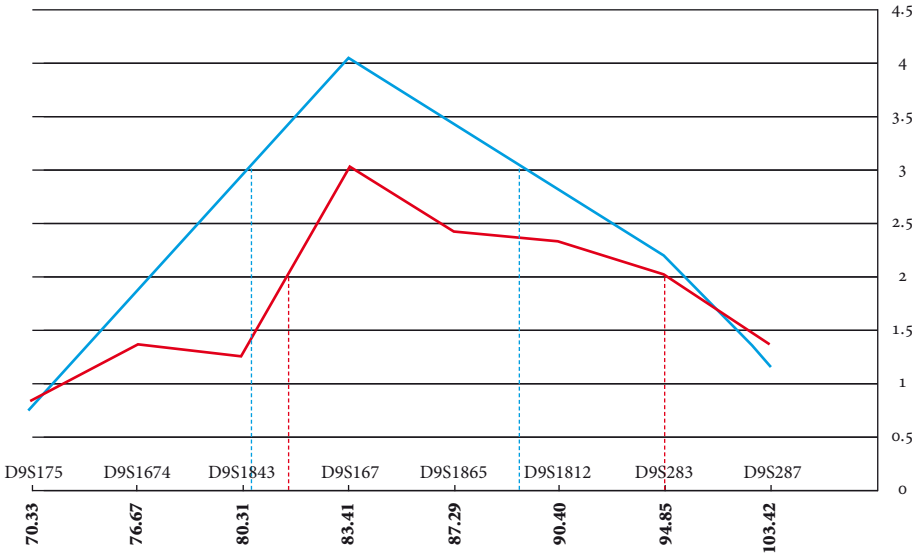


Fig. 2. Nonparametric LOD scores as computed by Merlin, in all families (blue line), and with additional markers (red line). The dashed lines represent the -1 LOD interval.

completely or almost completely, did so on 29 different haplotypes. Although, depending on the number of markers considered around D9S167, a suggestive core haplotype could sometimes be discerned between two or more haplotypes (data not shown). These results support the NPL scores for D9S167, but also indicate extensive allelic heterogeneity for this sharing, as well as genetic heterogeneity across families because not all families contribute to the NPL score.

Candidate gene analysis

We selected five genes from the region between markers D9S1843 and D9S283, on the basis of their presumed cellular function (see Materials & Methods). These included *UBQLN1*, *RASEF*, *DAPK1*, *TLE1*, and *GADD45γ*. The entire coding regions of these genes were sequenced in 16 patients from 16 families displaying complete haplotype sharing at D9S167. Nineteen variants were found in one or more patients (Table 3), 11 of which were known SNP's. For several variants we detected homozygotes for both alleles, making them unlikely candidates for susceptibility alleles. Of all the exonic variants found, there was only one missense change, in *TLE1* in one family. This variant did not co-segregate with disease. Three of the 6 intronic variants were known SNP's, and none were predicted to affect the nearest splice-site. The latter was also found for all the neutral exonic changes. We conclude that no clear disease-related changes were detected in this screen.

CGH-analysis

Copy number changes of the linked region on chromosome 9 were investigated in 61 tumors from 27 families by examing the intensity ratios of the 13 BAC clones representing this region on the array (Table 4). We were able to compare the results of 22 tumors from 10 families that displayed complete sharing of a 8-cM haplotype around D9S167 in all patients ('linked tumors'), with 39 tumors from 17 families without such haplotype sharing ('unlinked tumors'). A high proportion of linked tumors (55%) showed copy number loss at a BAC RP11-276H19 containing the *GAS1* gene and D9S167, and none showed gain. However, this was not statistically different from the unlinked tumors in which 31% showed copy loss at this BAC. We did observe a significant difference in the percentage of *BRCAX* tumors with loss of this BAC as compared to sporadic tumors (average ^{2}log ratio of −0.208 as compared with a ^{2}log ratio of 0.088 for control tumors). This difference has a two-tailed unpaired t-test p-value of 0.00039.

TABLE 3

Gene changes detected in sequence analysis of 5 candidate genes

Gene name	Gene Change	Exon / intron	Times found heterozygous[1]	Known SNP	Splice-site prediction[2]	Co-segregation[3]
DAPK1	c.393C>T, p.His131His	Exon 4	1	No	No change	NI
	TTCA(G/A)GAT, 143481A>AG, p.267Gln>Gln	Exon 9	1 (and 2 minor homozygotes)	No	No change	NI
	g.144573A>AG	Intron 9	7 (and 3 minor homozygotes)	rs3118846	No change	NI
	g.150139G>AG	Intron 14	4	rs2274607	No change	NI
	c.1608C>T, p.Asp537Asp	Exon 16	13	rs3818584	No change	NI
	c.1830G>A, p.Gly610Gly	Exon 18	2	No	No change	NI
	g.201463G>T	Intron 23	3	rs3128495	No change	NI
	c.3597C>T, p.Arg1200Arg	Exon 26	10	rs3118863	No change	NI
TLE1	c.983C>T, p.Thr328Met	Exon 12	1	No	No change	No
	c.1101A>G, p.Pro367Pro	Exon 13	1	No	No change	NI
GADD45G	c.102+65G>C	Intron 1	4	rs3138502	No change	NI
	c.157-18_19delCC;-18_19insCTAG	Intron 2	2	No	No change	No
RASEF	c.785C>T, p.Arg262Cys	Exon 5	10	rs4146960	No change	NI
	c.1202+57_61delGTAAA	Intron 9	6	No	No change	No
	c.1731 T>G, p.Asp577Asp	Exon 13	1	rs34303676	No change	NI
	c.2223+18A>T	Intron 17	1	No	No change	No
UBQLN1	c.1494C>T, p.Ser498Ser	Exon 10	5 (and 1 minor homozygote)	rs2781004	No change	No
	c.1590C>A, p.Leu530Leu	Exon 10	7 (and 1 minor homozygote)	rs7866234	No change	No
	c.1617+17G>A	Intron 10	8	rs9314722	No change	No

[1] Out of 16 patients tested

[2] Using splice-prediction programs NNSPLICE version 0.9 by Neural Network (www.fruitfly.org/seq_tools/splice.html), NetGene 2 Server version 2.42 by CBS Software Package Manager (www.cbs.dtu.dk/services/NetGene2) and Alex Dong Li's SpliceSiteFinder (http://violin.genet.sickkids.on.ca/~ali/splicesitefinder.html)

[3] NI = Not investigated

TABLE 4

Results of array-CGH analysis for 13 BACs in the region 9q21-q22

BAC	Mb	gene	9q-linked tumors[1]			9q-nonlinked tumors[2]		
			gain	loss	inc	gain	loss	inc
RP11-66D1	83	TLE1	5%	14%	81%	13%	23%	64%
RP11-432M2	84		9%	0%	91%	26%	3%	72%
RP11-541F16	84,7	RASEF	5%	10%	86%	0%	5%	95%
RP11-439A18	85,3	UBQLN1/HNRPK	32%	7%	62%	13%	13%	74%
RP11-59M22	86,3		18%	0%	82%	23%	5%	72%
RP11-172F7	87,1		64%	23%	13%	59%	18%	23%
RP11-280P22	87,9		41%	7%	52%	31%	5%	64%
RP11-276H19	88,8	GAS1	0%	55%	45%	8%	31%	62%
RP11-423O13	88,9		0%	18%	82%	0%	8%	92%
RP11-40C6	89,3		9%	10%	81%	15%	13%	72%
RP11-249H20	89,4	DAPK1	9%	15%	76%	13%	3%	85%
RP11-65B23	89,6	CCRK	18%	17%	65%	26%	10%	64%
RP11-8B23	91,7	GADD45	5%	29%	67%	0%	28%	72

[1] Group of 22 breast tumors from 10 families in which all genotyped patients shared a haplotype around D9S167;

[2] Group of 39 breast tumors from 17 families without sharing of a haplotype around D9S167.
gain, ratio tumor/normal > 0.2 on a 2log-scale; loss, ratio tumor/normal less than –0.2 on a 2log-scale; inc, inconclusive (ratios between –0.2 and 0.2)

DISCUSSION

The analysis reported here represents the largest single-center genome-wide linkage search for new susceptibility loci in non-*BRCA1/2* breast cancer families to date. The rationale for this study was that there exist further breast cancer genes which confer moderate to high risks (6,28). The patterns of familial clustering in the families that we selected for our study suggest that such alleles are likely to be dominant. Initial suggestive linkage peaks observed in a 'linkage search' group of 55 families were subsequently confirmed and confined to a locus on chromosome 9 in a total set of 85 families. Linkage evidence was most apparent using allele-sharing analyses with the Merlin package, with a single point non-parametric lodscore of 4.63 and a multipoint score of 3.96 at marker D9S167. These NPL scores represent the highest for any single locus in a linkage search after the identification of *BRCA1* and *BRCA2*, but they are dif-

ficult to compare with previous studies because these have mainly analysed marker data with the more conservative GENEHUNTER software. Under a parametric dominant model, however, the HLOD at D9S167 was 0.56 (α=0.18), indicating that even though the allele-sharing at D9S167 was significant in the total set of families, most families did not support linkage to this locus. Indeed, we noted complete allele-sharing among patients among 25 of the 85 families (29%) at D9S167, and suggestive incomplete sharing in 12 other families (14%). We observed extensive haplotype heterogeneity around shared alleles at D9S167. One explanation for this is that there is a gene (or genes) near D9S167 in which multiple rare variants confer substantially increased risks to breast cancer. The low HLOD score at this marker is probably due to the small number of families demonstrating complete haplotype sharing among the patients in conjunction with the fact that for most families and patients the (founder) parents were unavailable for genotyping. In the MERLIN analysis, the overall information content at D9S167 in the 85 families was 42%, which is in agreement with simulation studies with microsatellite maps of ~10 cM and incomplete parental genotypes.[29] Further genotyping of the region at much higher resolutions could therefore help to identify regions with more consistent allele-sharing.

Recently the Breast Cancer Linkage Consortium published the results of a genome wide linkage search for breast cancer susceptibility genes,[1] which included 149 multiple case non-*BRCA1/2* breast cancer families. The highest LOD score under the dominant model was 1.80, for a region on chromosome 4. Although several other suggestive LOD scores were reported, the number of linkage peaks did not differ from the number expected by chance and therefore these peaks probably reflect the play of chance rather than true susceptibility loci. In agreement with the BCLC-study we found no evidence for linkage to markers on 2q32,[10] 3p26,[30] 8p12-22,[31,32] 10q23.32-q25.3,[11] 11q23,[30] 13q21[33] and 22q13.1,[30,34] which were all previously suggested to harbour susceptibility loci. But we also did not find any evidence for linkage on the regions reported by the BCLC-publication. The inability to detect strong linkage signals may be a reflection of extensive locus heterogeneity.

The BCLC analysis[1] included 22 Dutch families that were also part of the 55 families investigated here. The NPL score near the D9S167 locus in that study was 0.74, and although this was the second highest score for the Dutch families, there was no evidence for allele-sharing at this locus in the other 127 families collected in that study, derived from Australia, United Kingdom, USA, Canada and France. Of the 22 Dutch families, 10 showed allele-sharing at D9S167, but 2 of these on different haplotypes. It is possible that our linkage study might have achieved greater statistical power

because the families derive from a more homogeneous population (i.e., with reduced genetic heterogeneity). The Dutch population exhibits distinct founder mutations for several known cancer susceptibility genes,[12,14,35,36] and therefore could be considered, to an extent, to be genetically distinct.[37] Although a strong founder effect at the 9q-locus seems less likely, given the diversity of shared haplotypes, such an effect has also been observed at *BRCA1* in the presence of extensive allelic heterogeneity.[12,13]

In a genome-wide scan for linkage in 14 Finnish breast cancer families,[10] the second highest HLOD peak was found at D9S283, just 5 cM distal of D9S167. D9S167 was also shown to be linked (with a multipoint parametric LOD score of 3.02) to ocular melanoma in three Danish families[38] with multiple cases of ocular malignant melanoma, cutaneous malignant melanoma and other malignancies, including breast cancer. In addition, D9S167 was in the center of a small chromosomal deletion in a case of acute myeloid leukemia.[39] These results suggest there is a gene in this region that can be linked to cancer susceptibility. Further evidence that the 9q21 region may be involved in a subset of the familial form of breast cancer comes from our observation that over 50% of *BRCAX* tumors putatively linked to 9q21 show copy-number loss at this locus, as opposed to 5-25% in sporadic breast tumors.[40,41] Others did not observe excess copy-number losses of 9q in familial non-*BRCA1/2* breast tumors using classical CGH,[33,42] although these patients were selected under different criteria than our cases. We previously reported[43] that ~30% of the same set of *BRCAX* tumors showed loss of heterozygosity (or allelic imbalance) at a marker for 9q34, which is not significantly higher than found in sporadic tumors.[44] Because that marker is a long distance away from D9S167, it is possible that some of the copy-number losses in the *BRCAX* tumors are tightly localized around 9q21.

The number of genes between D9S1843 and D9S283 presently annotated is 49. We performed sequence analysis of 5 of these to search for possible susceptibility alleles. No clear pathogenic changes were found in any of them. For all genes an apparent link with tumorigenesis could be made, such as a Ras GTPase motif in the *RASEF*-gene (closest to D9S167), transcription regulation (*TLE1*), or involvement in apoptosis (*DAPK1*) or stress response (*GADD45γ*).[45-47] However, a direct link with breast cancer has not yet been established for most of these candidates. In sporadic breast cancer, the expression of *RASEF* at mRNA-level is apparently not reduced.[38] *TLE1* has been suggested to play a role during epithelial differentiation[48] and tumor progression through inhibition of the *Wnt–CTNNB1* signaling pathway.[45] *DAPK1* and *GADD45G* are frequently targeted by inactivation through promotor hypermethylation in leukemias, lymphomas and a number of epithelial cancers.[47,49]

In conclusion, through linkage analysis we have identified a region on 9q21 which shows significant haplotype sharing among patients belonging to non-*BRCA1/2* families with at least three cases of breast cancer diagnosed before age 60. However, we observed extensive haplotype diversity at the shared locus, but have not yet identified sequence variants in candidate genes that could explain these results. There was some suggestion that the somatic genetic changes at this locus differ from that seen in sporadic breast tumors, which will have to be confirmed in larger series.

ACKNOWLEDGEMENTS

We want to thank Barbara Veldhuizen for the assistance with the Linkage analyses and Ben Oostra, Eline Slagboom and Mike Stratton for providing their genotype platforms. This research project was financialy supported by the Dutch Cancer Society (KWF, RUL1999-2021).

REFERENCE LIST

1. **Smith P, McGuffog L, Easton DF, et al.** Genome wide linkage search for breast cancer susceptibility genes. *Genes Chrom Cancer; 45:646-55 (2006)*
2. **Parkin DM, Pisani P, and Ferlay J.** Global cancer statistics. *CA Cancer J Clin; 49: 33-64 (1999)*
3. **Collaborative Group on Hormonal Factors in Breast Cancer Familial breast cancer** Collaborative reanalysis of individual data from 52 epidemiological studies including 58209 women with breast cancer and 101986 women without the disease. *Lancet 358:1389-99 (2001)*
4. **McPherson K, Steel CM, and Dixon JM.** ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *Br Med J; 321:624-8 (2000)*
5. **Thompson D and Easton D.** The genetic epidemiology of breast cancer genes. *Journal of Mammary Gland Biology and Neoplasia; 9:221-36 (2004)*
6. **Peto J, Collins N, Barfoot R, et al.** Prevalence of BRCA1 and BRCA2 gene mutations in patients with early- onset breast cancer. *J Natl Cancer Inst; 91:943-9 (1999)*
7. **Antoniou AC, Pharoah PDP, McMullan G, et al.** A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br J Cancer; 86: 76-83 (2002)*
8. **Antoniou AC and Easton DF.** Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol; 25:190-202 (2003)*
9. **Cui J, Antoniou AC, Dite GS, et al.** After BRCA1 and BRCA2-What Next? Multifactorial Segregation Analyses of Three-Generation, Population-Based Australian Families Affected by Female Breast Cancer. *Am J Hum Genet; 68:420-31 (2001)*
10. **Huusko P, Juo SH, Gillanders E, et al.** Genome-wide scanning for linkage in Finnish breast cancer families. *Eur J Hum Genet; 12:98-104 (2004)*
11. **Bergman A, Karlsson P, Berggren J, et al.** Genome-wide linkage scan for breast cancer susceptibility loci in Swedish hereditary non-BRCA1/2 families: suggestive linkage to 10q23.32-q25.3. *Genes Chrom Cancer; 46:302-9 (2007)*

12. **Peelen T, Van Vliet M, Petrij-Bosch A, 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; 60:1041-49 (1997)*
13. **Petrij-Bosch A, Peelen T, Van Vliet M, et al.** BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nature Genet; 17:341-5 (1997)*
14. **Broeks A, Urbanus JHM, De Knijff P, et al.** IVS10-6T > G, an ancient ATM germline mutation linked with breast cancer. *Hum Mutat; 21:521-8 (2003)*
15. **Hogervorst F, Cornelis R, Bout M, et al.** Rapid detection of BRCA1 mutations by the Protein Truncation Test. *Nature Genet; 10:208-12 (1995)*
16. **Korkko J, Annunen S, Pihlajamaa T, Prockop DJ, and Ala-Kokko L.** Conformation sensitive gel electrophoresis for simple and accurate detection of mutations: comparison with denaturing gradient gel electrophoresis and nucleotide sequencing. *Proc Natl Acad Sci USA; 95:1681-5 (1998)*
17. **Hogervorst FBL, Nederlof PM, Gille JJP, et al.** Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method. *Cancer Res; 63:1449-53 (2003)*
18. **Kruglyak L, Daly MJ, Reeve-Daly MP, and Lander ES.** Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet; 58:1347-63 (1996)*
19. **Claus EB, Risch N, and Douglas W.** Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet; 48:232-42 (1991)*
20. **Easton D, Bishop D, Ford D, Crockford G, and Breast Cancer Linkage Consortium** Genetic linkage analysis in familial breast and ovarian cancer: Results from 214 families. *Am J Hum Genet; 52:678-701 (1993)*
21. **Abecasis GR, Cherny SS, Cookson WO, and Cardon LR.** Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet; 30:97-101 (2002)*
22. **Ke X, Collins A, and Ye S.** PIRA PCR designer for restriction analysis of single nucleotide polymorphisms. *Bioinformatics; 17:838-9 (2001)*
23. **van Beers EH, Joosse SA, Ligtenberg MJ, et al.** A multiplex PCR predictor for aCGH success of FFPE samples. *Br J Cancer; 94:333-7 (2006)*
24. **Joosse SA, van Beers EH, and Nederlof PM.** Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. *Bmc Cancer; 7:43 (2007)*
25. **Jelier R, Jenster G, Dorssers LC, et al.** Text-derived concept profiles support assessment of DNA microarray data for acute myeloid leukemia and for androgen receptor stimulation. *BMC Bioinformatics; 8:e14 (2007)*
26. **Franke L, Bakel H, Fokkens L, de Jong ED, Egmont-Petersen M, and Wijmenga C.** Reconstruction of a functional human gene network, with an application for prioritizing positional candidate genes. *Am J Hum Genet; 78:1011-25 (2006)*
27. **Barrett JC, Fry B, Maller J, and Daly MJ.** Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics; 21:263-5 (2005)*
28. **Antoniou AC and Easton DF.** Models of genetic susceptibility to breast cancer. *Oncogene; 25:5898-905 (2006)*
29. **Evans DM and Cardon LR.** Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. *Am J Hum Genet; 75:687-92 (2004)*
30. **Hartikainen JM, Tuhkanen H, Kataja V, et al.** An autosome-wide scan for linkage disequilibrium-based association in sporadic breast cancer cases in eastern Finland: three candidate regions found. *Cancer Epidemiol Biomarkers Prev; 14:75-80 (2005)*

31. **Kerangueven F, Essioux L, Dib A, et al.** Loss of heterozygosity and linkage analysis in breast carcinoma: Indication for a putative third susceptibility gene on the short arm of chromosome 8. *Oncogene; 10:1023-26 (1995)*
32. **Seitz S, Rohde K, Bender E, et al.** Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: linkage analysis in German breast cancer families. *Oncogene; 14:741-3 (1997)*
33. **Kainu T, Juo SH, Desper R, et al.** Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proc Natl Acad Sci USA; 97:9603-8 (2000)*
34. **Hartikainen JM, Tuhkanen H, Kataja V, et al.** Refinement of the 22q12-q13 breast cancer-associated region: Evidence of TMPRSS6 as a candidate gene in an eastern Finnish population. *Clinical Cancer Research;12:1454-62 (2006)*
35. **Gruis NA, Van der Velden PA, Sandkuijl LA, et al.** Homozygotes for CDKN2 (p16) germline mutation in Dutch familial melanoma kindreds. *Nat Genet; 10:351-3 (1995)*
36. **Taschner PEM, Jansen JC, Baysal BE, et al.** Nearly all hereditary paragangliomas in the Netherlands are caused by two founder mutations in the SDHD gene. *Genes Chrom Cancer; 31:274-81 (2001)*
37. **Zeegers MP, van Poppel F, Vlietinck R, Spruijt L, and Ostrer H.** Founder mutations among the Dutch. *Eur J Hum Genet; 12:591-600 (2004)*
38. **Jonsson G, Bendahl PO, Sandberg T, et al.** Mapping of a novel ocular and cutaneous malignant melanoma susceptibility locus to chromosome 9q21.32. *J Natl Cancer Inst; 97:1377-82 (2005)*
39. **Sweetser DA, Peniket AJ, Haaland C, et al.** Delineation of the minimal commonly deleted segment and identification of candidate tumor-suppressor genes in del(9q) acute myeloid leukemia. *Genes Chromosomes Cancer; 44:279-91 (2005)*
40. **Rennstam K, Ahlstedt-Soini M, Baldetorp B, et al.** Patterns of chromosomal imbalances defines subgroups of breast cancer with distinct clinical features and prognosis. A study of 305 tumors by comparative genomic hybridization. *Cancer Res; 63:8861-8 (2003)*
41. **Chin K, DeVries S, Fridlyand J, et al.** Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell; 10:529-41 (2006)*
42. **Gronwald J, Jauch A, Cybulski C, et al.** Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. *Int J Cancer; 114:230-6 (2005)*
43. **Oldenburg RA, Kroeze-Jansema K, Meijers-Heijboer EJ, et al.** Characterization of familial non-BRCA1/2 breast tumors by loss of heterozygosity and immunophenotyping. *Clin Cancer Res; 12:1693-1700 (2006)*
44. **Wang Z, Lin M, Wei LJ, et al.** Loss of Heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. *Cancer Res; 64:64-71 (2004)*
45. **Levanon D, Goldstein RE, Bernstein Y, et al.** Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. *Proc Natl Acad Sci U.S.A; 95:11590-5 (1998)*
46. **Inbal B, Cohen O, Polak-Charcon S, et al.** DAP kinase links the control of apoptosis to metastasis. *Nature; 390:180-4 (1997)*
47. **Ying J, Srivastava G, Hsieh WS, et al.** The stress-responsive gene GADD45G is a functional tumor suppressor, with its response to environmental stresses frequently disrupted epigenetically in multiple tumors. *Clin Cancer Res; 11:6442-9 (2005)*
48. **Liu Y, Dehni G, Purcell KJ, et al.** Epithelial expression and chromosomal location of human TLE genes: implications for notch signaling and neoplasia. *Genomics; 31: 58-64 (1996)*
49. **Raval A, Tanner SM, Byrd JC, et al.** Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. *Cell; 129:879-90 (2007)*

CHAPTER 6

DOES THE BRCAX GENE EXIST? FUTURE OUTLOOK

Genetic research aimed at the identification of new breast cancer susceptibility genes is at an interesting crossroad. On the one hand, the existence of extended kindred's with many cases of (early-onset) breast cancer, in which a role for *BRCA1* and *BRCA2* has been excluded with high certainty, strongly suggests that there are still *BRCA1/2*-like genes to be found.[75] On the other hand, the absence of significant linkage signals in a set of 149 non-*BRCA1/2* breast cancer families indicates that if such a locus exists, it is unlikely to explain a major proportion of non-*BRCA1/2* families.[235] Are further 'classical' linkage studies therefore futile?

Before dismissing linkage analysis entirely, one should realize that the linkage results published to date do not permit a formal exclusion of the possibility that there are multiple genes causing breast cancer risks comparable to *BRCA1* or *BRCA2*, i.e., *BRCA3*, *BRCA4*, etc., but that their individual mutation frequencies are so low that each will explain no more than 10% of the families under study. The statistical power required to significantly resolve that kind of genetic heterogeneity was not achieved by any of the published linkage studies to date. Even in the largest study of 149 families, only 24 families had four or more cases of breast cancer diagnosed under 50 and 74 had four or more cases diagnosed before the age of 60 (ref. 235 and D. Easton, personal communication). For comparison's sake, previous studies addressing genetic heterogeneity analyzed more than 200 such families.[75,224] If four or five of those were in fact linked to a hypothetical *BRCA3* locus, one would have to be extremely fortunate with the informativity of the genotyped markers and patients to detect a significant linkage peak. For this reason, the Breast Cancer Linkage Consortium is now undertaking a study with the aim to obtain linkage data on at least 250 breast cancer families. Even though it might appear as if classical linkage approaches are running out of steam with regards to their potential to detect new breast cancer susceptibility loci, our claim on chromosome 9 proves that it's still possible to identify new breast cancer risk loci with classical linkage when genetic heterogeneity is reduced. In our case we selected families from the Dutch population, which is known to harbor many founder mutations for different diseases and therefore this group is, to an extent, genetically distinct. In addition, studies from different populations all show different susceptibility loci.[229,231,236,264] So, these loci might reflect population specific

effects. Our obtained linkage result suggests a region on chromosome 9 that may harbor a novel breast cancer susceptibility gene. However, if such gene exist in this region it's likely that it will account for only a limited fraction of the non-*BRCA1/2* families, especially in populations other than the Dutch.

Nowadays other new bioinformatics tools are being developed to improve linkage power. A promising option is to integrate chromosome segregation data with data obtained from functional genomic approaches such as large-scale, high-throughput molecular profiling technologies.[265]

Molecular profiling of global gene expression has already enabled the subclassification of breast cancer into prognostically relevant subgroups,[42,266] and has demonstrated to be capable of identifying cases who are very likely to be *BRCA1* gene carriers.[42,267] Similar results have been obtained by using array-CGH with tumor DNA.[255,256] On this basis, one could hypothesize that breast tumors with the same genetic etiology (because of a shared familial predisposition) will also be more likely to share a molecular signature. This could provide means to eliminate phenocopies from the linkage analysis, or to assign a liability to each patient as to how much her tumor resembles a typical *BRCAX*-related cancer. This will require sufficient resolution both in terms of number of probe sets, and the number of cases and controls to be analyzed, because initial results seem to suggest that *BRCAX*-related tumors resemble sporadic breast tumors in terms of somatic genetic and immunohistochemical make-up.[187,258,260] However, when comparing BRCAx-breast tumor array-CGH profiles with profiles of control samples we showed that there are multiple regions of differential gains and losses. Unfortunately, the dendrogram of BRCAx tumors indicated that there are no major branch points to suggest obvious distinct BRCAx CGH subtypes. Interesting though was the observation that when sporadic and BRCAx tumors were clustered together a non-random distribution occurred. Both BRCAx and sporadic tumors remained clustered in just eight sub clusters, five of which contained only BRCAx tumors and three clusters contained all sporadic cases plus one BRCAx tumor (van Beers/Oldenburg submitted, *Cancer Research*). Although, it was not possible to distinguish different groups within the BRCAX tumors at this point, it deserves further exploration.

Another intriguing possibility is to exploit gene expression patterns in normal cells of cases (such as lymphocytes or skin fibroblasts). This method relies on the fact that messenger RNA's with premature stop codons (i.e., nonsense mutations, which constitute a frequent class of mutations underlying inherited disease) are efficiently degraded by the conserved nonsense-mediated decay pathway. The number of genes

displaying consistently lower expression in normal cells of familial cases versus controls can be substantially reduced by comparing their genomic location to chromosomal segments shared among family-members identical-by-descent. An example of this approach is the recent identification of *AIP* as a susceptibility gene for pituitary adenoma in a Finnish founder population.[268]

Also, the discovery of micro-RNA's opens many new doors in cancer research (see the excellent review by G.A. Calin and C.M. Croce[269]). Micro-RNA's (miRNA) are a family of 19-25-nucleotide non-coding small RNA's that function as gene regulators and are involved in crucial biological processes, including development, apoptosis, proliferation and differentiation through pairing with target messenger RNA's (mRNA) of protein-coding genes. Perfect pairing of miRNAs to mRNAs leads to degradation of the mRNA, whereas less strict sequence complementarity results in translational repression. Recently, it has been shown that miRNAs can function either as tumor suppressors or oncogenes and the genomic abnormalities found to influence the activity of miRNAs are the same as those previously described for protein-coding genes, such as chromosomal rearrangements, genomic amplification or deletions and mutations. The role of miRNAs in tumor formation is strongly supported by the observation that the coding sequence of 50% of the known miRNAs, frequently situated in introns of coding genes, are located inside or close to fragile sites and minimal regions of LOH, minimal regions of amplification and common breakpoints associated with cancer. In addition, miRNA expression profiles showed that miRNAs are differentially expressed in normal and tumor samples (the expression in tumor tissue seems generally lower than in normal tissue) and that their expression fingerprints correlate with clinical and biological characteristics of tumors. In breast cancer the expression profile of a set of 15 miRNAs correctly predicted the nature of the breast cancer sample analyzed with 100% accuracy. Furthermore the expression of miRNAs was correlated with specific breast cancer pathological features such as estrogen-receptor status, tumor stage, vascular invasion, proliferation index and clinical features such as prognosis and response to therapy. Unfortunately, to date no attempt has been made to use miRNA expression profiles for subclassification of the heterogeneous group of familial breast cancer. It might be worthwhile to explore this possibility.

Obviously, the success of all of these approaches is dependent on the amount and quality of the information from the pedigrees under study and the availability of biological samples from the patients. The analysis of gene expression profiles of tumors still relies heavily on the availability of frozen tissue samples. Logistically, these

are particularly difficult to obtain from multiple affected family-members, which explains why so few of such studies have appeared in the published literature.

Notwithstanding these developments, attention is now shifting rapidly toward the whole genome association studies in population-based breast cancer cases and controls.[270] In a typical association study, the frequency of a genetic variant in affected individuals (cases) is compared to that in individuals without the disease (controls).[271,272] Allelic association is present when the distribution of genotypes differs in cases and controls. Most association studies are based on candidate genes that encode proteins thought to be involved in carcinogenesis, such as those involved in apoptosis, cell-cycle control, or DNA repair. Within candidate genes, variants for which a functional connotation can be inferred are preferably tested for association with disease. In this way, variants in *TGFβ1* and *CASP8* have been identified as breast cancer susceptibility alleles (see chapter 2, sections 3.2.2. and 3.2.3.).

Association studies are the only alternative to family-based linkage analyses for detecting alleles that confer low to moderate disease risks.[273,274] Even though risks are low, such alleles have the potential of explaining a substantial proportion of disease heredity, depending on their population frequency. There are several examples of common variants that contribute to common diseases,[275,276] but none of them could have been detected by linkage analysis. For example, the P12A variant in the *PPARG*-gene, which affects the risk of type 2 diabetes, would only be detected using linkage studies of over one million affected sib pairs.[277]

Mutations in the currently known high risk breast cancer genes are common in families with a large number of cases of breast and/or ovarian cancer,[75] but they have been estimated to explain at best 20-25% of the overall excess familial risk[278] and less than 5% of the total breast cancer incidence.[71] The proportion of breast cancer that can be attributed to genetic factors is not clear, but several studies have suggested it to be much larger than 5%. A large twin study has estimated that up to 30% of all breast cancer has a genetic basis,[279] while a study of the incidence of contralateral breast cancer has even suggested that the majority of all breast cancer occurs in a small minority of women who are susceptible for it.[280] It is unlikely that further *BRCA1/2*-like genes, if they are detected, will be capable of accounting for these attributable risks, because their allele frequencies are already predicted to be rare. More common alleles with moderate effects could do so, but it is not clear how many of such alleles exist and how much of the genetic predisposition to breast cancer can be attributed to them. Assuming relative risks in the order of 1.3 – 1.5, the remainder of excess risk could equally well be explained by a few hundred common variants

(with frequencies of >1%) or thousands of rare variants.[270] Such a polygenic model has in fact been supported by segregation analyses in non-*BRCA1/2* families (chapter 2, section 4.2). Under this model, many low to moderate risk cancer susceptibility genes cause breast cancer predisposition, together with environmental risk factors, in a multiplicative or additive way, with no single gene accounting for a large fraction of the familial aggregation. Individuals carrying few such alleles would be at lower or equal population risk compared to those carrying multiple such alleles. Conversely, there will also be alleles that protect against the development of cancer (such as *CASP8*, chapter 2, section 3.2.3).

The major problem haunting association studies is the lack of reproducibility by other, independent studies.[57,58] Thus, most studies are too small and probable report false positive results due to chance (type 1 error), which depends on the level of significance used. Unfortunately, the levels of significance appropriate in other contexts (P=0.05 or P=0.01) can be highly misleading in association studies.[270] By using more stringent levels of statistical significance this false positive rate can be reduced. Alternatively failure to confirm associations might be the result of heterogeneity in risk between populations due to for example interacting lifestyle and environmental factors. Also, strikingly little research has been performed on combinations of polymorphisms. It is still possible that polymorphisms not associated with breast cancer when studied separately, are associated with breast cancer when studied in combination with other polymorphisms. For example, recently a significant trend in risk with increasing numbers of variant alleles for 25 SNP's in *BRCA1*, *BRCA2*, *ATM*, *TP53* and *CHEK2* was observed[281] whereas common polymorphic variants in these genes separately are unlikely to increase breast cancer risk.[282]

In addition, a lack of association of a candidate SNP does not necessarily rule out the presence of another important variant in the same gene. For any given gene of interest, there might be tens or even hundreds of different sequence variants.

A large genome-wide association study, which would involve millions of SNP's with the use of stringent significance levels, would be ideal to identify common breast cancer susceptibility genes. However, the number of cases to be genotyped depends on the allele frequency of the variant and the disease risk conferred by it. For relative risks in the order of 1.5, allele frequencies in the range 10% – 40% will require at least a thousand cases and controls to be genotyped[270,273] in order to have 90% power to detect associations at a significance level of 10^{-4}. Much larger numbers are required for allele frequencies in the 1% – 10% range. This is presently not feasible at the going costs per genotype per sample. Fortunately, it is not necessary to genotype all pos-

sible SNP's to detect an association, because the alleles of SNP's that are physically close to each other tend to be correlated with each other (tag SNP mapping). This phenomenon is called linkage disequilibrium (LD).[283-285] The ability of one SNP to report on another depends on the strength of LD between them. The general consensus is that an $r^2 > 0.8$ (r^2 is a measure of correlation between a pair of variables[273]) is sufficient for a tag SNP to obtain a good coverage of untyped common SNP's. It has been estimated that 200,000 – 500,000 tagging SNP's will be needed to adequately tag all SNP's with a minor allele frequency of 5% or more.[286,287]

Accumulating the data necessary to choose such SNP's is one of the main goals of the human HapMap project.[287,288] The existence of LD can also be exploited to examine candidate genes by haplotype analysis, whereby the haplotypes are defined by a set of tagging SNP's. Any common variant in the gene that increases disease risk will then be detected as an increase of the particular haplotype on which this variant resides. This approach has been used to investigate a possible involvement of common variation in *CYP19* and *BRCA2* genes.[289,290]

In addition, it has been proposed that the power to detect associations may be increased by genotyping familial cases rather than sporadic population-based cases,[291,292] an effect which was indeed observed for the *CHEK2*1100delC* variant.[175,176] It seems therefore the most efficient to perform a whole-genome scan for association in a small sample of cases that are enriched for susceptibility. These could be familial cases or early-onset bilateral breast cancer cases, but one could also select cases enriched for other risk factors with a strong genetic component such as breast density[29] or cellular radio sensitivity.[293]

The recent publication by the Breast Cancer Association Consortium proved the success of this strategie.[294] In the first stage a panel of 266,722 SNP's (selected to tag known common variants across the entire genome) was genotyped in 408 breast cancer cases with a strong family history of breast cancer and 400 controls. In the second stage 12,711SNP's (approximately 5% of those typed in stage one) were selected on the basis of the significance of the difference in genotype frequency between cases and controls and genotyped in 3,990 invasive breast cancer cases and 3,916 controls. In the third stage 30 of the most significant SNP's were tested in 21,860 cases and 22,578 controls. This resulted in five novel loci strongly associated with breast cancer with a significance level ranging from 2×10^{-76} to 3×10^{-9} of which four contain plausible causative genes; *FGFR2, TNRC9, MAP3K1* and *LSP1*. The five SNP's that reached an overall P-value $<10^{-7}$ showed an increased breast cancer risk of the minor allele in a dose dependent manner, with higher risk of breast cancer in homozygous

than in heterozygous carriers. It is notable that none of the confirmed associations reached genome wide significance after stage 1 and only one reached this level after stage 2, emphasizing the critical importance of study size in genetic association studies.

This study has also demonstrated conclusively that some of the variation in breast cancer risk is due to common alleles, as these five identified susceptibility alleles are very common. So, a high proportion of the general population is carrier of at risk genotypes. For example 14% of the UK population are homozygous for the rare allele. However, the increased risks associated with these alleles are relatively small. On the basis of UK population rates, the estimated breast cancer risk by age 70 years for rare homozygote's at the SNP in *FGFR2* is 10.5%, compared to 6.7% in heterozygote's and 5.5% in common homozygote's. It is likely that there are still other common variants to be identified as *CASP8* D302H, which showed strong evidence of association in a previous large study[172] was missed, because it did not reach the threshold for testing in stage 2. Also the excess of association ($P<0.05$) after stage 2 is consistent with the existence of many such loci. In addition, because the coverage for SNP's with minor allele frequency's <10 % was low, many low frequency alleles have probable been missed. How much of the overall familial risk these alleles will be able to explain remains to be seen. It has been argued on the basis of evolutionary arguments that the role of rare alleles (i.e., frequencies <<1%) in causing late-onset disease such as cancer could be substantial.[295] Detecting this class of variants by current genetic approaches is impossible. Probable it will require genome-wide studies with more complete coverage (perhaps total genome sequencing) and using much larger number of (familial) cases and controls. Over the next decade, progress with the identification of common low risk variants will teach us how substantial this fraction is.

The proposed polygenetic model would not only be capable of explaining large but rare autosomal dominant-like familial clusters of (early-onset) breast cancer, but could also explain substantial proportions of the total breast cancer incidence.[296] It has been estimated that, should we be able to characterize all the relevant risk factors in all women of a given population, 50% of all breast cancer would occur in 12% of women with the highest risk profile.[296] From a health care perspective, the identification of these risk factors is therefore of great practical importance. Not only to define the cancer risk for women and their family members in order to make adequate decisions on surveillance and preventive strategies, but also for the development of gene targeted therapy. For example PARP [poly(ADP-ribose) polymerase] inhibitors may represent a novel way of selectively targeting BRCA2- or p53-deficient

breast cancer cells. Probably due to the additional inhibition of PARP activity an increase of unrepaired DNA damage occurs, causing a shift from DNA repair to apoptosis.[297-300]

Due to the low risk of the newly identified genes and SNP's and the existence of more low risk alleles to be identified it is too early to include these in predictive genetic testing at this stage. However, as further susceptibility alleles are identified over the next years, combinations of such alleles together with other breast cancer risk factors may become sufficiently predictive to be important clinically. And ideally, a chip with all risk alleles for predictive genetic testing is constructed in the near future to test women at risk.

CHAPTER 7

SUMMARY

CHAPTER 1

In chapter 1 the aims and outline of this thesis is described.

CHAPTER 2

In chapter 2 (introduction; based on a publication in *Critical Reviews in Oncology and Hematology, May 2007*) an outline of genetic aspects of breast cancer is given. However, before proceeding it is necessary for the layman to understand more about the mechanism causing cancer and thus about breast cancer. The human body is composed of cells. Each cell has a core (or nucleus) containing the major part of genetic information, DNA molecules, stored in 23 pairs of chromosomes. One set of chromosomes is derived from the father and one set of chromosomes is derived from the mother. Thus hereditary material is presented in duplicate, originating from both parents. DNA-molecules are constructed from a multiple of four building stones adenine (A), cytosine (C), guanine (G) and thymine (T). Selected regions of DNA, the genes, serve as template for synthesizing RNA-molecules, which in turn are utilised as a blueprint for creating proteins. The term protein stems from the Greek word *'proteios'* meaning 'from highest rank' reflecting the important role of proteins in different cellular functions, such as for instance transcription, transport, signaling and storage. All through our life cells divide themselves for replacement or multiplication (cellular proliferation), whereby genetic materials are copied and passed on. However, during this process changes may occur in the DNA (somatic mutations) supplying a new cell with a possible specific benefit. If through this mutation this cell is more capable of multiplying itself, such a cell will be inclined to dominate the organism. By comparison: any organism that shows hereditary variation in reproductive capacity will evolve by natural selection. Organisms that reproduce itself in a manner superior to the environment will come to dominate others. As tumors are distinguished by an unrestrained growth of cells, they do have through natural selection an advantage with respect to other cells. So humans actually have a natural inclination to change into tumors. However, tumors are incapable of having babies and care for them. Therefore strong genetic control mechanisms have developed over a trillion years of evolution, preventing a person, at least during his reproductive years, of changing into a tumor. Potential tumor cells are repaired and brought to heel or

forced into cell death (apoptosis). Experience nevertheless teaches us that tumors actually may develop during life. This is only possible when multiple defence mechanisms of the cell are halted. So in order to alter a cell into a tumor a number of successful mutations are required, especially in genes that enhance cell proliferation, also referred to as 'gatekeeper'-genes, through which a greater cell population does develop for the 'next' mutation, as well as in genes that affect the stability over the complete genome (on DNA or chromosome level), through which mutation frequency may increase, so-called 'caretaker' – genes, e.g. DNA-repair-genes.

Among women, breast cancer is the most frequently occurring type of cancer (22% of all female cancers). The number of patients with breast cancer annually increases worldwide with approximately 1 million. Cumulative lifetime risk for Dutch women is 9%. Several risk factors for breast cancer are known, of which positive family history for breast cancer is one of the most important. This indicates that hereditary factors play an important role in the development of breast cancer. First-degree family members (mother, sisters and daughters) of breast cancer patients run twice as high a risk for breast cancer. This risk increases with the number of breast cancer patients in the family, the age breast cancer manifests itself, the younger the patient the higher the risk, the occurrence of bilateral breast cancer and a history of benign breast disorders. At this moment approximately 10% of all breast cancers is accounted for by germline mutations, meaning: already present in the fetal cells at conception, in known breast cancer predisposition genes. These genes can roughly be divided into high-risk genes (*BRCA1, BRCA2, PTEN, TP53, LKB1/STK11* and *CDH1)* with a lifetime risk of over 4 times the average and in low to moderate increased-risk genes (*CHEK2*, *TGFβ1*, *CASP8, BARD1, BRIP1, PALB2* and *ATM*). High-risk genes are the principal cause of frequent occurrence of breast cancer within specific families and are mostly found through linkage studies where within families searches are made for loci on the genome shared among breast cancer patients, assuming a specific statistical model (hereditariness, allelfrequency and penetration). A 'Logarithm of Odds' (LOD-score) greater than 3 on a specific locus is interpreted as a significant finding and indicates that at that locus a possible breast cancer susceptibility gene may be discovered.

Low to moderate increased-risk genes however cannot be identified by linkage analysis because the genotype-phenotype relation is much weaker. The most common method of identifying these genes is the association study in which allelfrequency of specific variations in (candidate) genes is compared between a great number of breast cancer patients and a control group.

At this moment *BRCA1* and *BRCA2* account for the major part of families with more breast cancer patients, patients with cancer of the ovaries and/or male breast cancer patients, but to a lesser degree for families where female breast cancer is the only occurring form of cancer.

BRCA1 and *BRCA2* are both viewed as 'caretaker'-genes and play a significant role in spotting and repairing DNA-damage. The hereditary path of mutations in both genes takes place in a classical autosomal dominant way; meaning children from a person with a germline mutation in either gene have a 50% chance of inheriting this mutation. Functionally at the cellular level however, these mutations are recessive. In *BRCA1* and *BRCA2* associated tumors one mutant copy of the gene (allele) is inherited through the germline. Inactivation of the other allele is obtained on somatic level during life (in the epithelium of the mammary gland). Carriers of a mutation in other high risk cancer predisposition genes *TP53* (Li-Fraumeni syndrome), *PTEN* (Cowden syndrome), *CDH1* (HDGC-syndrome) and *LKB1/STK11* (Peutz-Jeghers syndrome) are also associated with a highly increased breast cancer risk, however germline mutations are very rare and are not found in breast cancer patients without other associated features of these disorders.

Clinical experience however teaches us that there still are many hereditary encumbered breast cancer families without a mutation in *BRCA1 or BRCA2*. The hypothesis is that there should exist other high-risk genes that may be identified through linkage research. The power of linkage research depends heavily on information rendering of the families to be screened and the number of still to be discovered predisposition genes (heterogeneity of the disorder). Alas, after the discovery of *BRCA1 and BRCA2* in the mid nineties, no new high-risk breast cancer predisposition gene was discovered through linkage research. One of its meanings could be that heterogeneity among families is greater than expected and the up-to-now completed research included too few families for reaching a significant LOD-score.

Our research aimed at attempting to identify new high-risk breast cancer predisposition genes through genome-wide linkage analysis. In collaboration with the Breast Cancer Linkage Consortium (BCLC) 150 Dutch, English, French and Australian BRCA1/2 negative families were selected with a minimum of 3 breast cancer patients diagnosed under the age of 60, without cancer of the ovaries or male breast cancer patients. Next to that we collected from 55 Dutch families as much paraffin imbedded tumour samples as possible, to endeavour reducing heterogeneity within the selected families. Research in *BRCA1* (and to a lesser extent also *BRCA2*) related tumours has demonstrated these tumours to distinguish themselves from sporadic

(viz.: non-hereditary) and *BRCA1/2*-negative tumours as regards to histopathology, array-CGH profile, micro-array profile and immunohistochemistry. This may possibly be the case with *BRCA3, 4* etc. (*BRCAX*).

CHAPTER 3

Chapter 3.1 (publication in *Journal of Medical Genetics. 2004*) describes one of the families we thus selected. This family carried apart from breast cancer an unexpectedly great number of other types of cancer, among them melanomas, lung cancer, intestinal cancer and oral squamous cell carcinoma. In this family a mutation was found in the *p16-gene* (*p16-Leiden* mutation), associated with an increased risk of melanomas. Seeing much breast cancer also occurred in this family and as other researchers already suggested that *p16* possibly played a role in the etiology of breast cancer, we examined the role of *p16* in the development of (breast) cancer within this family, supplemented with a survey of four additional breast tumours from *p16-Leiden* positive patients from four different families. We concluded there to be no clear connection between carriers of a *p16-Leiden* mutation and the development of breast cancer, seeing most (4 out of 5) breast cancer patients within the family quoted above did not carry the mutation and 3 out of four of the additionally selected breast tumours showed immunohistochemically no elimination of the *p16* gene. However we did find a connection between the development of lung cancer and oral squamous cell carcinoma and carriers of the *p16-Leiden* mutation.

During our search for new breast cancer predisposition genes the international research area in this field, didn't stand still. H. Meijers-Heijboer *et al.* identified the *CHEK2*1100delC* variation as a low-risk breast cancer predisposition gene (relative risk: 2.0).

Chapter 3.2 (publicized in *Cancer Research, in 2003*) describes the role of this variation within our selected families. Selection of breast cancer patients with a strong familiar burden clearly shows an increased occurrence of this variation as opposed to sporadic breast cancer patients. In 15 out of 71 families (21%) minimum one breast cancer patient with this variation was found. It was remarkable that within these families no apparent co-segregation of this variation with breast cancer was established. However, patients carrying this variation developed breast cancer at a younger age than patients without this variation. With this research we were also the first to demonstrate that *CHEK2*1100delC* carrier is coupled with an absent immunohistochemical staining in tumour cells. Our results support a model whereby an

increase from breast cancer risk possibly may be explained by an interaction between *CHEK2*1100delC* and a still to be identified new breast cancer predisposition gene or genes (oligogenetic/polygenetic model).

In the mean time a Scandinavian Group claimed a possible breast cancer predisposition gene to be discovered on the long arm of chromosome 13 (13q21).
Chapter 3.3 is a manuscript published in *Proc Natl Acad Sci U.S.A* in 2002, where we refute this. In this research the Breast Cancer Linkage Consortium demonstrated that in a group of 128 high-risk families there is no association between breast cancer and 13q21 (heterogeneity LOD score: -11).

CHAPTER 4

In this chapter we describe an attempt to decrease heterogeneity within our families through tumor features.

Chapter 4.1 is a manuscript published in *Clinical Cancer Research* in 2006. Recent studies demonstrated that *BRCA-1*-related tumors show a specific histopathological, immunohistochemical and genetic profile. This shows that it may be possible to decrease heterogeneity within our families, should several subgroups be identified within *BRCAX*-related tumors. To this aim 100 *BRCAX*-tumors were investigated and examined for 'Loss of heterozygosity (LOH)'. Here LOH-frequencies higher than 40% were found on 1q41, 4p16, 11q23.3, 16p13, 16q24, 17p12, 21q22, 22q11 and 22q13, with the highest frequency on 22q13 (59%). Except for areas on 22q, these loci had been found in sporadic breast tumors as well. It was possible to examine LOH in minimum 2 tumors from different patients in each of 28 families. Here we found markers on chromosome 2, 3, 6, 12, 13, 21 and 22 (however not on 22q13) on which LOH occurred significantly more frequently in tumors from patients belonging to the same families than one would expect based on total LOH-frequencies. Albeit, linkage analysis for markers on corresponding areas for chromosome 12, 21 and 22 returned no significant LOD-scores. Immunohistochemically *BRCAX* tumors were significantly more often positive for bcl2 than *BRCA1* tumors (p=0.000005) and than *BRCA2* tumors (p=0.00003). This actually was also the case for *CHEK2*1100delC* tumors. It was also noticeable that *CHEK2*1100delC* tumors were significantly more often negative for cytokeratin 19 staining compared to *BRCA1* (P=0.0008) and the remainder of *BRCAX* tumors (P=0.006). Alas cluster analysis for combined data (LOH and immunohistochemistry) did not return any useful sub groups for use in linkage analysis.

Chapter 4.2 is a manuscript submitted for publication describing results found using array-CGH in 58 *BRCAX* tumors compared to 48 sporadic tumors. *BRCAX* tumors generally show more significant copy number changes than sporadic tumors (P=0.003). *BRCAX* tumors show significantly more loss of genetic material on chromosome 1p, 1q, 4q, 5q, 9q, 13q, 14q, 15q, 19cen, 21p and Xp and an increase on chromosome 2q-ter, 6p, 8p, 11p, 12p, 14q, 17p, 17q, 19p, 19q and of more areas on chromosome 22 with regard to sporadic tumors. Increase on chromosome 22 appears to be specific for *BRCAX* tumors, as this is not found in either *BRCA1*, *BRCA2* or sporadic tumors. Using unsupervised hierarchical clustering an attempt was made in grouping 58 *BRCAX* tumors in more homogeneous sub groups for possible linkage analysis. Unfortunately no evident sub groups were found, however when *BRCAX* tumors together with sporadic tumors were clustered it was noticeable that no random fusion developed. *BRCAX* and sporadic tumors cluster separately.

CHAPTER 5

Chapter 5.1 was published in *Genes Chromosome and Cancer, 2006*. The manuscript describes results from the genome-wide linkage search performed by the Breast Cancer Linkage Consortium. The idea behind this research was that there still exist high-risk genes. In 149 high-risk families (22 originating from the Netherlands) a LOD-score of 1.80 was found under a dominant model on chromosome 4. A maximum 2.40 LOD-score on chromosome 2p was found, when only families with more than 4 breast cancer patients, diagnosed at less than 50 years of age were analysed. Neither were significant LOD-scores found under a recessive model and through nonparametric methods. The number of linkage peaks traced didn't differ from what could be expected based on coincidence. This research is by far the most extensive linkage research published up to now. Results suggest the heterogeneity among the families is high and possibly this may be solved by extending the set of families. At the moment the Breast Cancer Linkage Consortium is therefore trying to increase this number to 250 families or more. This may also mean that the marker set used is insufficiently informative.

Chapter 5.2 has recently been submitted for publication. The Dutch population is known for the fact that for many genetic disorders specific mutations occur that are less apparent in other populations. Therefore one could consider the Dutch population as being an unique genetic population. In order to evaluate the possibility that genetic heterogeneity among breast cancer families could be decreased through

selecting a more homogeneous population, we performed a linkage search among 85 Dutch families. 22 of these families were also included in the linkage search executed by the Breast Cancer Linkage Consortium. Assuming a dominant as well as a recessive model no significant LOD-scores were found. With nonparametric methods however on chromosome 9q21 a significant LOD-score was identified (for marker D9s167 the NPL-score being 3.96; P=0.00009). This suggests that at this locus a possible breast cancer predisposition gene is located. However, should this be the case only a small part of *BRCAX* families may possibly be accounted for. This will definitely be the case in non- Dutch populations.

CHAPTER 6

This chapter consists of a general discussion. Genetic research aimed at identification of breast cancer predisposition genes finds itself on interesting crossroads. On the one hand the existence of families with more (young) breast cancer patients without a mutation in *BRCA1* or *BRCA2* suggests that there must still be genes that cause a *BRCA1* or *BRCA2* comparable high breast cancer risk. On the other hand the absence of a significant linkage peak in a group of 149 high-risk families without *BRCA1* or *BRCA2* mutation made it clear that should such a gene exist, it can possibly only explain a small part of these families. There is a chance of the existence of more high-risk genes, but the individual contribution is too small to identify using current methods. This could be solved by extending the set of families or by grouping families in more homogeneous sub groups using tumor features (biomarkers) or by selecting families from a more homogeneous population. Using LOH, CGH and immunochemistry we made a first attempt at grouping families through biomarkers. Unfortunately this didn't lead to identification of a new gene. However, the first result obtained from CGH especially, indicates this needs further exploration.

Selecting families from a more homogeneous population also yielded an interesting result, namely the identification of chromosome 9q21 as a possible locus for a new breast cancer predisposition gene. Should this be the case the gene involved will mainly play a role in the Dutch population, as in other international linkage studies this locus did not occur.

Mutations in the currently known high risk breast cancer genes are common in families with a large number of cases of breast and/or ovarian cancer, but they have been estimated to explain at best 20-25% of the overall excess familial risk and less than 5% of the total breast cancer incidence. The contribution of genetic

factors in the etiology of breast cancer isn't quite clear. Several studies indicate that the possible role of genetic factors may be much higher than 5%. A large twin study has estimated that up to 30% of all breast cancer has a genetic basis, while a study on the incidence of bilateral breast cancer even suggested that the greatest part of breast cancer occurs in a small minority of women who are susceptible for it. It is unlikely that these attributable risks can completely be contributed to high-risk genes, as it was already suggested that should they indeed at all exist, mutations in these genes are very rare. Therefore the idea arose that frequently occurring low-risk variants and/or rare low-risk variants combined with each other may play a part. Such a polygenetic model is indeed supported by segregation analysis in non-*BRCA1/2* related families. In this model several combinations of more low-risk to moderate-risk cancer predisposition genes, together with environmental factors may explain families. Because such genes cannot be identified through genome-wide linkage analysis one sees at the moment a shifting taking place to genome-wide association studies. The problem of these studies is the great number of breast cancer patients and control patients required (in the order of 20,000 patients and an equal number of controls), this being very costly. As demonstrated for the *CHEK2*1100delC* variant, an enrichment for low-risk to moderate-risk variants occurs when high-risk families are selected. Therefore it appears to be very efficient to first perform a genome-wide study in a small group enriched with breast cancer predisposition. These may be familial cases but also for instance bilateral breast cancer patients or persons with other risk factors with a strong genetic component such as breast tissue density. Next, variants significantly associated with breast cancer may be typified in a great (multicentre) case-control study.

It may be clear that identification of these genes is of great importance seen from the perspective of health care. Not only for the assessment of (breast) cancer risk for women and their families and thus to attain adequate decisions regarding preventive strategies (check-up, preventive surgery and chemo prevention) but also for the development of therapies aimed at deviations of these genes.

CHAPTER 8

SAMENVATTING

HOOFDSTUK 1

In hoofdstuk 1 wordt het doel en de aanpak van dit onderzoek beschreven.

HOOFDSTUK 2

In Hoofdstuk 2 (introduction; gebaseerd op de publicatie in *Critical Reviews in Oncology and Hematology, mei 2007*) wordt een overzicht gegeven van de genetische aspecten van borstkanker. Alvorens hier verder op in te gaan is het voor de leek noodzakelijk om meer te begrijpen over het mechanisme van het ontstaan van kanker en dus ook borstkanker. Het menselijke lichaam is opgebouwd uit cellen. In de cel zit een celkern met daarin het overgrote deel van de genetische informatie (DNA moleculen), opgeslagen in 23 paar chromosomen. Het erfelijk materiaal is dus in tweevoud aanwezig, afkomstig van vader en afkomstig van moeder. DNA-moleculen zijn opgebouwd uit een veelvoud van vier bouwstenen (adenine (A), cytosine (C), guanine (G) en thymine (T). Geselecteerde regio's van het DNA (genen) dienen als sjablonen voor het synthetiseren van RNA-moleculen, die op hun beurt gebruikt worden als blauwdruk voor het vormen van eiwitten (proteïnen). De term proteïne komt van het Griekse woord '*proteios*' wat betekent 'van de hoogste rang' waarmee de belangrijke rol van eiwitten in de diverse cellulaire functies, zoals transcriptie, transport, signalering, en opslag wordt weerspiegeld. Gedurende ons leven delen cellen zich ter vervanging of ter vermenigvuldiging (cellulaire proliferatie), waarbij het genetische materiaal wordt gekopieerd en doorgegeven. Bij dit proces kunnen er echter veranderingen optreden in het DNA (somatische mutaties) die een bepaald voordeel kunnen geven aan de nieuwe cel. Indien door de verandering de cel beter in staat is zich voort te planten/te vermenigvuldigen zal deze de neiging vertonen om het organisme te gaan domineren. (ter vergelijking: een organisme dat zich beter kan voortplanten zal door natuurlijke selectie gaan domineren over de anderen). Daar tumoren worden gekenmerkt door een ongeremde groei (voortplanting) van cellen hebben ze dus door natuurlijke selectie een voordeel t.o.v. de andere cellen. Dus eigenlijk hebben mensen van nature de neiging om in tumoren te veranderen. Echter, tumoren zijn niet in staat om baby's te krijgen en deze te verzorgen. In de biljoen jaren van evolutie zijn op het niveau van het gehele organisme daarom sterke genetische controle mechanisme's ontwikkelt om te voorkomen dat een persoon (in elk geval in zijn reproductieve jaren) in een tumor verandert. Potentiële tumor cellen

worden of gerepareerd en terug in het gareel gebracht of gedwongen om zelfmoord (apoptose) te plegen. Desondanks leert de ervaring ons dat tumoren weldegelijk kunnen ontstaan. Hiervoor is het echter noodzakelijk dat meerdere van de verdedigingsmechanismen van de cel worden stil gelegd. Dus, om een cel te veranderen in een tumor cel zijn meerdere succesvolle mutaties nodig in met name genen die de celproliferatie bevorderen (ook wel 'gatekeepers'-genen genoemd), waardoor er een grotere populatie van cellen ontstaan voor de 'volgende' mutatie en genen die effect hebben op de stabiliteit van het gehele genome (op DNA of chromosomaal niveau), waardoor de mutatie frequentie kan toenemen (ook wel 'caretaker' -genen genoemd, bv DNA-repair-genen).

Onder vrouwen is borstkanker de meest frequent voorkomende vorm van kanker (22% van alle vormen van kanker bij vrouwen). Jaarlijks komen er wereldwijd ongeveer 1 miljoen borstkanker patiënten bij. Het cumulatieve lifetime risico op borstkanker voor Nederlandse vrouwen is 9%. Er zijn meerder risicofactoren voor borstkanker bekend waarvan een positieve familieanamnese voor borstkanker de belangrijkste is. Dit geeft aan dat erfelijke factoren een grote rol spelen bij het ontstaan van borstkanker. Eerstegraads familieleden (zussen, moeder, dochters) van borstkanker patiënten hebben ongeveer een tweemaal zo hoog risico op borstkanker. Dit risico neemt toe met het toenemen van het aantal borstkanker patiënten in de familie, de leeftijd waarop de borstkanker tot uiting komt (hoe jonger de patiënt des te hoger is het risico), het voorkomen van tweezijdig borstkanker en een voorgeschiedenis van goedaardige borstaandoeningen. Op dit moment wordt ongeveer 10% van alle borstkanker's verklaard door reeds vanaf de geboorte aanwezige mutaties (kiembaan mutaties) in bekende borstkanker predispositie genen. Deze genen kunnen grofweg verdeeld worden in hoogrisico genen (*BRCA1, BRCA2, PTEN, TP53, LKB1/STK11* en *CDH1*) met lifetime risico hoger dan 4 en laag tot matig verhoogd risico genen (*CHEK2*, *TGFβ1*, *CASP8, BARD1, BRIP1, PALB2* en *ATM*). De hoogrisico genen zijn de voornaamste oorzaak van frequent voorkomen van borstkanker patiënten binnen bepaalde families en zijn voornamelijk gevonden door middel van koppelingsonderzoek (Linkage analyse), waarbij binnen families gezocht wordt naar gebieden op het genoom die gedeeld worden tussen borstkanker patiënten, onder aanname van een bepaald statistisch model (overerving, allelfrequentie en penetrantie). Een 'Logaritme of Odds' (LOD-score; een maat voor waarschijnlijkheid) groter dan 3 op een bepaald gebied wordt gezien als een significante bevinding en geeft aan dat op dat gebied mogelijk een aan borstkanker gerelateerd gen te ontdekken is.

De laag tot matig verhoogd risico genen kunnen echter niet op deze manier geïdentificeerd worden doordat de genotype-fenotype relatie (relatie tussen afwijkingen in genen en het tot uiting komen van de ziekte) een stuk zwakker is. De meest gangbare methode om deze genen te identificeren is de associatie studie, waarbij de allelfrequentie van bepaalde varianten in (kandidaat) genen wordt vergeleken tussen een groot aantal borstkanker patiënten en een controle groep.

Op dit moment verklaren *BRCA1* en *BRCA2* het grootste gedeelte van families met meerdere borstkanker patiënten, eierstokkanker patiënten en/of mannelijk borstkanker patiënten, maar in veel mindere mate in families waarin vrouwelijk borstkanker de voornaamst voorkomende vorm van kanker is. *BRCA1* en *BRCA2* worden beiden gezien als 'caretaker'-genen en spelen een belangrijke rol bij het signaleren en repareren van DNA-schade. De erfgang van mutaties in beide genen geschiedt op een klassieke autosomaal dominante wijze (kinderen van een persoon met een kiembaan mutatie in een van beide genen hebben 50% kans deze te erven). Functioneel zijn deze mutaties echter recessief. In *BRCA1* en *BRCA2* geassocieerde tumoren is één mutant allel via de kiembaan geërfd. Inactivatie van het andere allel wordt op somatisch niveau (in het epitheel van de borstklier) verkregen gedurende het leven. Draagsters van een mutatie in een van de andere hoogrisico kanker predispositie genen *TP53* (Li-fraumeni syndroom), *PTEN* (Cowden syndroom), *CDH1* (HDGC-syndroom) en *LKB1/STK11* (Peutz-Jegher syndroom) worden ook geassocieerd met een sterk verhoogd borstkanker risico, echter kiembaan mutaties zijn zeer zeldzaam en worden niet gevonden bij borstkanker patiënten zonder de andere geassocieerde stigmata van deze aandoeningen.
De ervaring uit de kliniek leert ons echter, dat er nog veel sterk belaste borstkanker families zijn zonder een mutatie in *BRCA1* of *BRCA2*. De hypothese is dan ook dat er nog andere hoog risico genen zouden moeten bestaan, die mogelijk met linkage onderzoek geïdentificeerd kunnen worden. De kracht van linkage onderzoek hangt echter sterk af van de informativiteit van de te onderzoeken families en het aantal nog te ontdekken predispositie genen (de heterogeniteit van de aandoening). Helaas is na de ontdekking van *BRCA1* en *BRCA2* halverwege de jaren 90 geen nieuw hoogrisico borstkanker predispositie gen middels linkage onderzoek meer gevonden. Dit zou onder andere verklaard kunnen worden doordat de heterogeniteit onder de families groter is dan werd verwacht en de tot nu toe uitgevoerde onderzoeken te weinig families bevatte om tot een significante LOD-score te komen. Ons onderzoek had als doel om d.m.v. genoomwijd linkage analyse nieuwe hoogrisico borstkanker

predispositie genen te identificeren. In samenwerking met het Breast Cancer Linkage Consortium (BCLC) zijn hiervoor 150 Nederlandse, Engelse, Franse en Australische BRCA1/2 negatieve families geselecteerd met minimaal 3 borstkanker patiënten gediagnosticeerd onder 60 jaar, zonder eierstokkanker patiënten of mannelijke borstkanker patiënten. Daarnaast hebben wij uit 55 Nederlandse families zoveel mogelijk in paraffine ingebedde tumor samples verzameld, om te pogen de heterogeniteit binnen de geselecteerde families te verminderen. Onderzoek in *BRCA1* (en in mindere mate ook *BRCA2*) gerelateerde tumoren heeft namelijk uitgewezen dat deze tumoren zich onderscheiden van sporadische en BRCA1/2-negatieve tumoren qua histopathologie, array-CGH profiel, micro-array profiel en immunohistochemie. Wellicht is dit ook het geval voor *BRCA3*, *4*, etc (*BRCAx*).

HOOFDSTUK 3

Een van de families die wij voor dit doel geselecteerd hebben bevatte naast borstkanker een onverwacht groot aantal andere vormen van kanker, waaronder huidkanker (melanomen), longkanker, darmkanker en kanker in de mondholte (oral squamous cell carcinoma). In deze familie werd een mutatie in het *P16* gen gevonden (*P16-Leiden* mutatie), welke geassocieerd is met een verhoogd risico op melanomen. Aangezien er ook veel borstkanker in deze familie voorkomt en andere onderzoekers reeds gesuggereerd hadden dat het *P16* gen mogelijk een rol speelt in de etiologie van borstkanker, hebben we onderzoek verricht naar de rol van *P16* bij het ontstaan van borstkanker binnen deze familie, aangevuld met vier additionele borsttumoren van *P16-Leiden* positieve patiënten uit vier verschillende families. Met dit onderzoek zijn wij tot conclusie gekomen dat er geen duidelijk verband was tussen dragerschap van een *P16-Leiden* mutatie en het ontstaan van borstkanker, aangezien de meeste (4 van de 5) borstkanker patiënten binnen de hierboven beschreven familie de mutatie niet bij zich droegen en 3 van de vier additionele geselecteerde borsttumoren immunohistochemisch geen uitschakeling van het *P16* gen lieten zien. Wel werd er een verband tussen het ontstaan van longkanker, kanker in de mondholte (oral squamous cell carcinoma) en dragerschap van de *P16-Leiden* mutatie aangetoond. (Hoofdstuk 3.1*:* publicatie in de *Journal of Medical Genetics,* 2004).

Tijdens onze zoektocht naar nieuwe borstkanker predispositie genen stond het internationale onderzoeksveld niet stil op dit gebied. Zo identificeerde H. Meijers-Heijboer de *CHEK2*1100delC* variant als een laagrisico borstkanker predispositie gen. Hoofdstuk 3.2 *(*publicatie in *Cancer Research,* in 2003*)* beschrijft de rol van deze

variant binnen de door ons geselecteerd families. Het selecteren van borstkanker patiënten met een sterke familiaire belasting laat duidelijk een verrijking zien van deze variant ten opzichte van sporadische borstkanker patiënten. In 15 van 71 families (21%) werd minimaal één borstkanker patiënt met deze variant gevonden. Opvallend was echter dat er binnen de families geen duidelijke co-segregatie van de variant met borstkanker werd aangetoond. Echter, patiënten die de variant bij zich droegen ontwikkelde borstkanker op jongere leeftijd dan patiënten zonder deze variant. Ook toonde we met dit onderzoek als eerste aan dat *CHEK2*1100delC* dragerschap gepaard gaat met een afwezige immunohistochemische kleuring in de tumorcellen. Onze resultaten ondersteunen een model waarbij een verhoging van het borstkanker risico mogelijk verklaard kan worden door een interactie tussen *CHEK2*1100delC* en een nog te identificeren nieuw borstkanker predispositie gen of genen (polygenetisch model).

Een Scandinavische groep claimde ondertussen dat er mogelijk een borstkanker predispositie gen te ontdekken is op de lange arm van chromosoom 13 (13q21). Hoofdstuk 3.3 is een manuscript dat werd gepubliceerd in *Proc Natl Acad Sci U.S.A* in 2002 waarin we dit tegenspreken. In dit onderzoek toonde het Breast Cancer Linkage Consortium aan dat in een groep van 128 hoog-risico families er geen associatie is tussen borstkanker en 13q21 (Heterogeniteit LOD-score: -11).

HOOFDSTUK 4

In dit hoofdstuk beschrijven we een poging om d.m.v. tumorkarakteristieken de vermoedelijk aanwezige genetische heterogeniteit binnen onze families te verkleinen. Hoofdstuk 4.1 is een manuscript dat werd gepubliceerd in *Clinical Cancer Research* in 2006. Recente studies hebben aangetoond dat de *BRCA1*-gerelateerde tumoren een specifiek histopathologisch, immunohistochemisch en genetisch profiel vertonen. Dit geeft aan dat het wellicht mogelijk is om de heterogeniteit binnen onze families te verlagen, indien er verschillende subgroepen geïdentificeerd kunnen worden binnen de *BRCAX*-gerelateerde tumoren. Voor dit doel werden 100 *BRCAX*-tumoren immunohistochemisch onderzocht en onderzocht op 'Loss of heterozygosity (LOH)'. Hierbij werden LOH-frequenties groter dan 40% gevonden op 1q41, 4p16, 11q23.3, 16p13, 16q24, 17p12, 21q22, 22q11 en 22q13, met de hoogste frequentie op 22q13 (59%). Deze gebieden, behalve de gebieden op 22q waren reeds ook gevonden in sporadische borsttumoren. In 28 families was het mogelijk om LOH te onderzoeken in minimaal 2 tumoren van verschillende patiënten. Hierbij vonden we op chromo-

soom 2, 3, 6, 12, 13, 21 en 22 (echter niet op 22q13) markers waarop LOH significant frequenter optrad in tumoren van patiënten behorend tot dezelfde familie dan je zou verwachten op basis van de totale LOH-frequenties. Helaas leverde linkage analyse voor markers op de overeenkomende gebieden voor chromosoom 12, 21 en 22 geen significante LOD-scores op. Immunohistochemisch waren de *BRCAX* tumoren significant vaker positief voor bcl2 dan *BRCA1* (P=0.000005) en BRCA2 tumoren (P=0.00003), wat overigens ook het geval was voor de *CHEK2*1100delC* tumoren. Ook viel op dat de *CHEK2*1100delC* tumoren significant vaker negatief waren voor cytokeratine 19 kleuring t.o.v *BRCA1* (P=0.0008) en het restant van de *BRCAX* tumoren (P=0.006). Helaas leverde de cluster analyse voor de gecombineerde data (LOH en immunohistochemie) geen bruikbare subgroepen op die gebruikt konden worden voor linkage analyse.

Hoofdstuk 4.2 is een manuscript dat momenteel aangeboden is voor publicatie en beschrijft de m.b.v array-CGH gevonden resultaten in 58 *BRCAX* tumoren in vergelijking met 48 sporadische tumoren. The *BRCAX* tumoren laten in het algemeen significant meer kopienummer veranderingen zien dan sporadische tumoren (p=0.003). The *BRCAX* tumoren laten significant meer verlies van genetisch materiaal zien op chromosoom 1p, 1q, 4q, 5q, 9q, 13q, 14q, 15q, 19cen, 21p en Xp en toename op chromosoom 2q-ter, 6p, 8p, 11p, 12p, 14q, 17p, 17q, 19p, 19q en van meerdere gebieden op chromosoom 22 t.o.v. sporadisch tumoren. De toename op chromosoom 22 lijkt vrij uniek te zijn voor *BRCAX* tumoren daar dit niet wordt gevonden in zowel BRCA1, BRCA2 als sporadische tumoren. Met behulp van ongesuperviseerde hiërarchische clustering werd een poging gewaagd de 58 BRCAX tumoren te groeperen in meer homogene subgroepen voor eventuele linkage analyse. Helaas werden er geen duidelijke subgroepen gevonden, echter wanneer de BRCAX tumoren samen met de sporadische tumoren werden geclusterd viel op dat er geen willekeurige vermenging ontstond. De BRCAX en de sporadische tumoren bundelen zich apart van elkaar.

HOOFDSTUK 5

Hoofdstuk 5.1 werd in 2006 gepubliceerd in *Genes Chromosome and Cancer.* Het manuscript beschrijft de resultaten van de door het Breast Cancer Linkage Consortium uitgevoerde genoomwijde linkage onderzoek. De gedachte achter dit onderzoek is dat er nog hoogrisico genen bestaan. In 149 hoogrisico families (waarvan 22 uit Nederland afkomstig waren) werd onder een dominant model op chromosoom

4 een LOD-score van 1.80 gevonden. Een maximale LOD-score van 2.40 op chromosoom arm 2p werd gevonden, wanneer alleen de families met meer dan 4 borstkanker patiënten, gediagnosticeerd onder 50 jaar werden geanalyseerd. Ook onder een recessief model en middels nonparametrisch methodes werden geen significante LOD-scores gevonden. Het aantal gevonden linkage pieken was niet anders dan je op basis van toeval zou mogen verwachten. Dit onderzoek representeert verreweg het uitgebreidste linkage onderzoek dat tot op heden is gepubliceerd. De resultaten suggereren dat de heterogeniteit onder de families hoog is en wellicht is dit op te lossen door de set families verder uit te breiden. Momenteel wordt hiertoe door de BCLC een poging ondernomen (250 families). Ook kan het betekenen dat de gebruikte markermap onvoldoende informatief is.

Hoofdstuk 5.2 bevat een manuscript dat momenteel is aangeboden voor publicatie. De Nederlandse populatie staat bekend om het feit dat er voor veel genetische aandoeningen specifieke mutaties voorkomen die in andere populaties minder op de voorgrond staan. Daarom zou je de Nederlandse populatie kunnen beschouwen als een unieke genetische populatie. Om de mogelijkheid te evalueren dat de genetische heterogeniteit onder borstkanker families te verkleinen is door een meer homogene populatie te selecteren hebben we linkage onderzoek verricht in 85 Nederlandse families. 22 van deze families waren ook geincludeerd in het linkage onderzoek verricht door het Breast Cancer Linkage Consortium. Onder aanname van een dominant en recessief model werden er geen significante LOD-scores gevonden. Echter met nonparametrische methoden werd op chromosoom 9q21 een significante LOD-score geïdentificeerd (voor marker D9s167 was de NPL-score 3.96 ;p=0.00009). Dit suggereert dat er op dit gebied mogelijk een borstkanker predispositie gen ligt. Indien dit daadwerkelijk het geval is zal waarschijnlijk maar een klein deel van de *BRCAX* families hierdoor verklaard worden, zeker in andere dan de Nederlandse populatie.

HOOFDSTUK 6

Dit Hoofdstuk bevat een samenvattende discussie. Genetisch onderzoek gericht op de identificatie van borstkanker predispositie genen staat momenteel op een interessant kruispunt. Enerzijds suggereert het bestaan van families met meerdere (jonge) borstkanker patiënten zonder een mutatie in *BRCA1* of *BRCA2* dat er nog genen moeten zijn die een, met *BRCA1* of *BRCA2* vergelijkbaar hoog borstkanker risico veroorzaken. Anderzijds heeft de afwezigheid van een significante linkage piek in een

groep van 149 hoog risico families zonder een *BRCA1* of *BRCA2* mutatie ons duidelijk gemaakt dat indien zo'n gen bestaat het waarschijnlijk maar een klein gedeelte van deze families kan verklaren.
Het is een mogelijkheid dat er nog meerdere hoogrisico genen bestaan maar dat de individuele bijdrage te klein is om met de huidige verrichte studies te identificeren. Dit zou opgelost kunnen worden door de set families uit te breiden of de families te groeperen in meer homogene subgroepen, dmv het gebruik van tumorkarakteristieken (biomarkers) of families uit een meer homogene populatie te selecteren. Met behulp van LOH, CGH en immunohistochemie hebben wij een eerste poging ondernomen om middels biomarkers de families te groeperen. Helaas heeft dit tot nu toe niet geleid tot de identificatie van een nieuw gen. Het eerste resultaat verkregen uit, met name CGH, geeft aan dat dit nog verdere exploratie verdient.
Het selecteren van families uit een meer homogene populatie leverde eveneens een interessant resultaat op, namelijk de identificatie van chromosoom 9q21 als een mogelijk locus voor een nieuw borstkanker predispositie gen. Echter indien dit het geval is zal het betreffende gen waarschijnlijk vooral een rol zal spelen in de Nederlandse populatie, daar in ander internationale linkage studies dit gebied niet naar voren is gekomen.

De schatting is dat de huidige borstkanker genen minder dan 5% van de totale borstkankerincidentie verklaren. Het aandeel van genetische factoren in de etiologie van borstkanker is niet geheel duidelijk. Meerdere studies geven aan dat dit waarschijnlijk veel hoger is dan 5%. Een studie betreffende de incidentie van contralateraal borstkanker suggereerde zelfs dat het grootste gedeelte van alle borstkanker optreedt in een kleine minderheid van vrouwen die een verhoogde gevoeligheid hiervoor hebben. Het is onwaarschijnlijk dat dit risico geheel is toe te schrijven aan hoogrisico genen, aangezien reeds voorspeld is dat, indien ze überhaupt bestaan, mutaties in deze genen zeer zeldzaam zijn. Het idee is daarom ontstaan dat meerdere frequent voorkomende varianten met een laag risico en of zeldzamere varianten met een matig risico in combinatie met elkaar een rol spelen.Zo'n poligenetisch model wordt inderdaad ondersteund door segregatie analyses in niet BRCA1/2 gerelateerde families. Onder dit model zullen verschillende combinaties van meerdere laag-tot matig risico kanker predispositie genen, samen met omgevingsfactoren mogelijk de families verklaren. Aangezien zulke genen niet middels linkage onderzoek geïdentificeerd kunnen worden is er momenteel veel aandacht voor genoomwijde associatie studies. Het probleem van deze studies is dat je zeer grote aantallen borstkanker

patiënten en controles nodig hebt (in de orde van grote van 20.000 patiënten en een even groot aantal controle personen), wat zeer kostbaar is. Zoals wij hebben aangetoond voor de *CHEK2*1100delC* variant, ontstaat er een verrijking van laag tot matig risico varianten door het selecteren van hoogrisico families. Het lijkt daarom zeer efficiënt te zijn om eerst in een kleine groep personen dat verrijkt is voor borstkanker predispositie een genoomwijde associatie studie te verrichten. Dit kunnen familiare gevallen zijn maar ook bijvoorbeeld bilaterale borstkanker patiënten of personen met andere risico factoren met een sterke genetische component zoals borstweefsel densiteit. Vervolgens kunnen de varianten die hierbij significant geassocieerd zijn met borstkanker getypeerd worden in een grote (multicenter) case-control studie. Inmiddels zijn de eerste laag tot matig risico genen op deze manier ontdekt.

Het moge duidelijk zijn dat de identificatie van zulke genen uit het perspectief van de gezondheidszorg van groot belang is. Niet alleen om het (borst)kanker risico voor vrouwen en hun familieleden te kunnen inschatten en zo tot adequate beslissingen te komen qua preventieve strategieën (controle, preventieve ingrepen en chemopreventie). Maar ook voor de ontwikkeling van therapieën gericht op de onderliggende gen afwijkingen.

CHAPTER 9

PUBLICATION LIST

1

A genome wide linkage search for breast cancer susceptibility genes.
Genes Chromosomes Cancer. 2006 Jul;45(7):646-55
Smith P, McGuffog L, Easton DF, Mann GJ, Pupo GM, Newman B, Chenevix-Trench G; kConFab Investigators; Szabo C, Southey M, Renard H, Odefrey F, Lynch H, Stoppa-Lyonnet D, Couch F, Hopper JL, Giles GG, McCredie MR, Buys S, Andrulis I, Senie R; BCFS, BRCAX Collaborators Group; Goldgar DE, **Oldenburg R**, Kroeze-Jansema K, Kraan J, Meijers-Heijboer H, Klijn JG, van Asperen C, van Leeuwen I, Vasen HF, Cornelisse CJ, Devilee P, Baskcomb L, Seal S, Barfoot R, Mangion J, Hall A, Edkins S, Rapley E, Wooster R, Chang-Claude J, Eccles D, Evans DG, Futreal PA, Nathanson KL, Weber BL; Breast Cancer Susceptibility Collaboration (UK); Rahman N, Stratton MR.

2

Characterization of familial non-BRCA1/2 breast tumors by loss of heterozygosity and immunophenotyping.
Clin Cancer Res. 2006 Mar 15;12(6):1693-700.
Oldenburg RA, Kroeze-Jansema K, Meijers-Heijboer H, van Asperen CJ, Hoogerbrugge N, van Leeuwen I, Vasen HF, Cleton-Jansen AM, Kraan J, Houwing-Duistermaat JJ, Morreau H, Cornelisse CJ, Devilee P.

3

Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. *Cancer Res. 2005 Feb 1;65(3):822-7.*
van Beers EH, van Welsem T, Wessels LF, Li Y, **Oldenburg RA**, Devilee P, Cornelisse CJ, Verhoef S, Hogervorst FB, van't Veer LJ, Nederlof PM.

4

Extending the p16-Leiden tumour spectrum by respiratory tract tumours.
J Med Genet. 2004 Mar;41(3):e31.
Oldenburg RA, de Vos tot Nederveen Cappel WH, van Puijenbroek M, van den Ouweland A, Bakker E, Griffioen G, Devilee P, Cornelisse CJ, Meijers-Heijboer H, Vasen HF, Morreau H.

5

Are ATM mutations 7271T-->G and IVS10-6T-->G really high-risk breast cancer-susceptibility alleles?
Cancer Res. 2004 Feb 1;64(3):840-3.
Szabo CI, Schutte M, Broeks A, Houwing-Duistermaat JJ, Thorstenson YR, Durocher F, **Oldenburg RA**, Wasielewski M, Odefrey F, Thompson D, Floore AN, Kraan J, Klijn JG, van den Ouweland AM, Wagner TM, Devilee P, Simard J, van 't Veer LJ, Goldgar DE, Meijers-Heijboer H.

6

The CHEK2*1100delC variant acts as a breast cancer risk modifier in non-BRCA1/BRCA2 multiple-case families. *Cancer Res. 2003 Dec 1;63(23):8153-7.*
Oldenburg RA, Kroeze-Jansema K, Kraan J, Morreau H, Klijn JG, Hoogerbrugge N, Ligtenberg MJ, van Asperen CJ, Vasen HF, Meijers C, Meijers-Heijboer H, de Bock TH, Cornelisse CJ, Devilee P.

7

Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet. 2002 May;31(1):55-9.*

Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, **Oldenburg R**, Hollestelle A, Houben M, Crepin E, van Veghel-Plandsoen M, Elstrodt F, van Duijn C, Bartels C, Meijers C, Schutte M, McGuffog L, Thompson D, Easton D, Sodha N, Seal S, Barfoot R, Mangion J, Chang-Claude J, Eccles D, Eeles R, Evans DG, Houlston R, Murday V, Narod S, Peretz T, Peto J, Phelan C, Zhang HX, Szabo C, Devilee P, Goldgar D, Futreal PA, Nathanson KL, Weber B, Rahman N, Stratton MR; CHEK2-Breast Cancer Consortium.

8

Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium.

Proc Natl Acad Sci U S A. 2002 Jan 22;99(2):827-31.

Thompson D, Szabo CI, Mangion J, **Oldenburg RA**, Odefrey F, Seal S, Barfoot R, Kroeze-Jansema K, Teare D, Rahman N, Renard H, Mann G, Hopper JL, Buys SS, Andrulis IL, Senie R, Daly MB, West D, Ostrander EA, Offit K, Peretz T, Osorio A, Benitez J, Nathanson KL, Sinilnikova OM, Olah E, Bignon YJ, Ruiz P, Badzioch MD, Vasen HF, Futreal AP, Phelan CM, Narod SA, Lynch HT, Ponder BA, Eeles RA, Meijers-Heijboer H, Stoppa-Lyonnet D, Couch FJ, Eccles DM, Evans DG, Chang-Claude J, Lenoir G, Weber BL, Devilee P, Easton DF, Goldgar DE, Stratton MR; KConFab Consortium.

9

Genetic susceptibility for breast cancer: How many more genes to be found?

Crit Rev Oncol Hematol. 2007 May 9.

Oldenburg RA, Meijers-Heijboer H, Cornelisse CJ, Devilee P.

10

BRCAX breast tumors are distinct from sporadic and BRCA1 tumors by array-CGH but still heterogeneous supporting the possibility for multiple etiologies. *Submitted*

EH van Beers, **RA Oldenburg**, SA Joosse, MJ Ligtenberg, N Hoogerbrugge, K Kroeze-Jansema, H Meijers-Heijboer, FB Hogervorst, S Verhoef, P Devilee, PM Nederlof.

11

Genome-wide linkage scan in Dutch hereditary non-BRCA1/2 breast cancer families identifies 9q21-22 as a putative breast cancer susceptibility locus. *Submitted*

R.A. Oldenburg, K.H.G. Kroeze, J.J. Houwing-Duistermaat, C.J. van Asperen, A. van den Ouweland, E. Bakker, E.H. van Beers, P.M. Nederlof, H. Vasen, N. Hoogerbrugge, C.J. Cornelisse, H. Meijers-Heijboer, P. Devilee.

12

Genomewide high-density SNP linkage analysis of non-BRCA1/2 breast cancer families identifies various candidate regions and has greater power than microsatellite studies.

BMC Genomics. 2007 Aug 30;8:299.

Gonzalez-Neira A, Rosa-Rosa JM, Osorio A, Gonzalez E, Southey M, Sinilnikova O, Lynch H, **Oldenburg RA**, van Asperen CJ, Hoogerbrugge N, Pita G, Devilee P, Goldgar D, Benitez J.

CHAPTER 10

CURRICULUM VITAE

The author of this thesis was born on 6th of October 1968 in Lower Hutt, Wellington, New Zealand, as the third son of Hans Oldenburg en Nibs Bloem. He passed his VWO exam in 1987 at the 'Montessori School' in The Hague. After finishing his introductory exam Computer Science in 1990 he started his medical career at the University in Gent, Belgium. He obtained his MD degree from the Leiden University in 1997. He started working as a junior resident for the department of Clinical Genetics at the Erasmus University Rotterdam in 1998. In 1999 he was offered an 'AGIKO' position in which the training for clinical genetics is combined with the research project (group leader Prof. P. Devilee) involving familial breast cancer, of which this thesis is the end product. He has finished his training for clinical genetics in October 2007 and he has received a staff position at the Department of Clinical Genetics, Erasmus Medical Centre Rotterdam. His main focus will be oncogenetics. He lives with Mariëtte van Pelt since 1992. They have two daughters; Michelle (1999) and Cathelijne (2002).

CHAPTER 11

REFRENCE LIST CHAPTER 2 AND 6

1. **Breasted, J. H.** The Edwin Smith Surgical Papyrus 403-406 *University of Chicago Press, Chicago (1930)*
2. Tumours of the breast in World Health Organization: Tumours of the Breast and Female Genital Organs 9-110. *IARC Press, Lyon (2003)*
3. **Parkin, D. M., Pisani, P., & Ferlay, J.** Global cancer statistics. *CA Cancer J.Clin. 49, 33-64, 1 (1999)*
4. **Parkin, D. M.** International variation. *Oncogene 23, 6329-6340 (2004)*
5. **Ziegler, R. G. et al.** Migration patterns and breast cancer risk in Asian-American women. *J Natl.Cancer Inst. 85, 1819-1827 (1993)*
6. **Kliewer, E. V. & Smith, K. R.** Breast cancer mortality among immigrants in Australia and Canada. *J Natl.Cancer Inst. 87, 1154-1161 (1995)*
7. **Hill, T. D., Khamis, H. J., Tyczynski, J. E., & Berkel, H. J.** Comparison of male and female breast cancer incidence trends, tumor characteristics, and survival. *Ann.Epidemiol. 15, 773-780 (2005)*
8. **Hulka, B. S. & Moorman, P. G.** Breast cancer: hormones and other risk factors. *Maturitas. 38, 103-113 (2001)*
9. **Garfinkel, L., Boring, C. C., & Heath, C. W., Jr.** Changing trends. An overview of breast cancer incidence and mortality. *Cancer. 74, 222-227 (1994).*
10. **Bernstein, L., Teal, C. R., Joslyn, S., & Wilson, J.** Ethnicity-related variation in breast cancer risk factors. *Cancer. 97, 222-229 (2003)*
11. **Dumitrescu, R. G. & Cotarla, I.** Understanding breast cancer risk -- where do we stand in 2005? *J Cell Mol.Med. 9, 208-221 (2005)*
12. **Russo, J., Hu, Y. F., Yang, X., & Russo, I. H.** Developmental, cellular, and molecular basis of human breast cancer. *J.Natl.Cancer Inst.Monogr. 17-37 (2000).*
13. **Cavalieri, E., Frenkel, K., Liehr, J. G., Rogan, E., & Roy, D.** Estrogens as endogenous genotoxic agents--DNA adducts and mutations. *J.Natl.Cancer Inst.Monogr. 75-93 (2000).*
14. **Berkey, C. S., Frazier, A. L., Gardner, J. D., & Colditz, G. A.** Adolescence and breast carcinoma risk. *Cancer 85, 2400-2409 (1999)*
15. **Kelsey, J. L. & Horn-Ross, P. L.** Breast cancer: magnitude of the problem and descriptive epidemiology. *Epidemiol.Rev. 15, 7-16 (1993)*
16. **Collaborative Group on Hormonal Factors in Breast Cancer.** Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. *Lancet. 347, 1713-1727 (1996)*
17. **McPherson, K., Steel, C. M., & Dixon, J. M.** ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *Br.Med.J. 321, 624-628 (2000)*
18. **Gikas, P. D. & Mokbel, K.** Phytoestrogens and the risk of breast cancer: a review of the literature. *Int.J Fertil.Womens Med. 50, 250-258 (2005)*
19. **Huang, Z. et al.** Dual effects of weight and weight gain on breast cancer risk. *JAMA. 278, 1407-1411 (1997)*
20. **Lahmann, P. H. et al.** Body size and breast cancer risk: findings from the European Prospective Investigation into Cancer And Nutrition (EPIC). *Int.J Cancer. %20;111, 762-771 (2004).*
21. **Trentham-Dietz, A. et al.** Weight change and risk of postmenopausal breast cancer (United States). *Cancer Causes Control. 11, 533-542 (2000)*

22. **McTiernan, A. et al.** Adiposity and sex hormones in postmenopausal breast cancer survivors. *J Clin Oncol. 21, 1961-1966 (2003)*
23. **Biglia, N. et al.** Management of risk of breast carcinoma in postmenopausal women. *Endocr.Relat Cancer. 11, 69-83 (2004)*
24. **Lagerros, Y. T., Hsieh, S. F., & Hsieh, C. C.** Physical activity in adolescence and young adulthood and breast cancer risk: a quantitative review. *Eur.J Cancer Prev. 13, 5-12 (2004)*
25. **Hankinson, S. E., Colditz, G. A., & Willett, W. C.** Towards an integrated model for breast cancer etiology - The lifelong interplay of genes, lifestyle, and hormones. *Breast Cancer Res. 6, 213-218 (2004)*
26. **Byrne, C. et al.** Effects of mammographic density and benign breast disease on breast cancer risk (United States). *Cancer Causes Control. 12, 103-110 (2001)*
27. **Boyd, N. F. et al.** Heritability of mammographic density, a risk factor for breast cancer. *N.Engl.J.Med. 347, 886-894 (2002)*
28. **van Gils, C. H., Hendriks, J. H., Otten, J. D., Holland, R., & Verbeek, A. L.** Parity and mammographic breast density in relation to breast cancer risk: indication of interaction. *Eur.J Cancer Prev. 9, 105-111 (2000)*
29. **Stone, J. et al.** The heritability of mammographically dense and nondense breast tissue. *Cancer Epidemiol.Biomarkers Prev. 15, 612-617 (2006)*
30. **Tokunaga, M. et al.** Incidence of female breast cancer among atomic bomb survivors, Hiroshima and Nagasaki, 1950-1980. *Radiat.Res. 112, 243-272 (1987)*
31. **Butel, J. S.** Viral carcinogenesis: revelation of molecular mechanisms and etiology of human disease. *Carcinogenesis. 21, 405-426 (2000)*
32. **zur, H. H.** Cervical carcinoma and human papillomavirus: on the road to preventing a major human cancer. *J Natl.Cancer Inst. 93, 252-253 (2001)*
33. **Stuver, S. O., Boschi-Pinto, C., & Trichopoulos, D.** Infection with hepatitis B and C viruses, social class and cancer. *IARC Sci.Publ. 319-324 (1997)*
34. **Iscovich, J., Boffetta, P., Franceschi, S., Azizi, E., & Sarid, R.** Classic kaposi sarcoma: epidemiology and risk factors. *Cancer. 88, 500-517 (2000)*
35. **Bangham, C. R.** HTLV-1 infections. *J Clin Pathol. 53, 581-586 (2000)*
36. **Feller, W. F. & Chopra, H. C.** A small virus-like particle observed in human breast cancer by means of electron microscopy. *J Natl.Cancer Inst. 40, 1359-1373 (1968)*
37. **Segev, N., Hizi, A., Kirenberg, F., & Keydar, I.** Characterization of a protein, released by the T47D cell line, immunologically related to the major envelope protein of mouse mammary tumor virus. *Proc.Natl.Acad.Sci.U.S.A. 82, 1531-1535 (1985)*
38. **Dion, A. S.** Virus-like particles and macromolecules in human milk and breast tumors. *CRC Crit Rev.Clin Lab Sci. 11, 245-270 (1979)*
39. **Wang, Y. et al.** Detection of mammary tumor virus ENV gene-like sequences in human breast cancer. *Cancer Res. 55, 5173-5179 (1995)*
40. **Etkind, P., Du, J., Khan, A., Pillitteri, J., & Wiernik, P. H.** Mouse mammary tumor virus-like ENV gene sequences in human breast tumors and in a lymphoma of a breast cancer patient. *Clin Cancer Res. 6, 1273-1278 (2000)*
41. **Pogo, B. G. et al.** Sequences homologous to the mouse mammary tumor virus env gene in human breast carcinoma correlate with overexpression of laminin receptor. *Clin Cancer Res. 5, 2108-2111 (1999)*
42. **Van 't Veer, L. J. et al.** Gene expression profiling predicts clinical outcome of breast cancer. *Nature 415, 530-536 (2002)*

43. **Einav, U. et al.** Gene expression analysis reveals a strong signature of an interferon-induced pathway in childhood lymphoblastic leukemia as well as in breast and ovarian cancer. *Oncogene. 24, 6367-6375 (2005)*
44. **Yin, H. et al.** Transcription of human endogenous retroviral sequences related to mouse mammary tumor virus in human breast and placenta: similar pattern in most malignant and nonmalignant breast tissues. *AIDS Res.Hum.Retroviruses. 13, 507-516 (1997)*
45. **Witt, A. et al.** The mouse mammary tumor virus-like env gene sequence is not detectable in breast cancer tissue of Austrian patients. *Oncol Rep. 10, 1025-1029 (2003)*
46. **Mant, C., Gillett, C., D'Arrigo, C., & Cason, J.** Human murine mammary tumour virus-like agents are genetically distinct from endogenous retroviruses and are not detectable in breast cancer cell lines or biopsies. *Virology. 318, 393-404 (2004)*
47. **Titus-Ernstoff, L. et al.** Exposure to breast milk in infancy and adult breast cancer risk. *J Natl.Cancer Inst. 90, 921-924 (1998)*
48. **Stewart, T., Tsai, S. C., Grayson, H., Henderson, R., & Opelz, G.** Incidence of de-novo breast cancer in women chronically immunosuppressed after organ transplantation. *Lancet. 346, 796-798 (1995)*
49. **Frisch, M., Biggar, R. J., Engels, E. A., & Goedert, J. J.** Association of cancer with AIDS-related immunosuppression in adults. *JAMA. 285, 1736-1745 (2001)*
50. **Ross, S. R., Schofield, J. J., Farr, C. J., & Bucan, M.** Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. *Proc.Natl.Acad.Sci.U.S.A. 99, 12386-12390 (2002)*
51. **Steel, M., Thompson, A., & Clayton, J.** Genetic aspects of breast cancer. *Br.Med Bull. 47, 504-518 (1991)*
52. **Eisinger, F., Sobol, H., Serin, D., & Whorton, J. C.** Hereditary breast cancer, circa 1750. *Lancet. 351, 1366 (1998)*
53. **Collaborative Group on Hormonal Factors in Breast Cancer.** Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58209 women with breast cancer and 101986 women without the disease. *Lancet 358, 1389-1399 (2001)*
54. **Thompson, D. & Easton, D.** The genetic epidemiology of breast cancer genes. *Journal of Mammary Gland Biology and Neoplasia 9, 221-236 (2004)*
55. **Amundadottir, L. T. et al.** Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. *PLoS.Med. 1, e65 (2004)*
56. **Claus, E. B., Risch, N., & Thompson, W. D.** Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer. 73, 643-651 (1994)*
57. **Dunning, A. M. et al.** A systematic review of genetic polymorphisms and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev. 8, 843-854 (1999)*
58. **de Jong, M. M. et al.** Genes other than *BRCA1* and *BRCA2* involved in breast cancer susceptibility. *J.Med.Genet. 39, 225-242 (2002)*
59. **Tavtigian, S. V. et al.** The complete *BRCA2* gene and mutations in chromosome 13q-linked kindreds. *Nature Genet 12, 333-337 (1996)*
60. **Kinzler, K. W. & Vogelstein, B.** Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature 386, 761-763 (1997)*
61. **Breivik, J.** The evolutionary origin of genetic instability in cancer development. *Semin.Cancer Biol. 15, 51-60 (2005)*
62. **Scully, R. & Livingston, D.** In search of the tumour-suppressor functions of *BRCA1* and *BRCA2*. *Nature 408, 429-433 (2000)*

63. **Venkitaraman, A. R.** Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell 108, 171-182 (2002)*
64. **Venkitaraman, A. R.** Tracing the network connecting BRCA and Fanconi anaemia proteins. *Nature Reviews Cancer 4, 266-276 (2004)*
65. **Narod, S. A. & Foulkes, W. D.** BRCA1 and BRCA2: 1994 and beyond. *Nature Reviews Cancer 4, 665-676 (2004)*
66. **Howlett, N. G. et al.** Biallelic inactivation of BRCA2 in Fanconi anemia. *Science 297, 606-609 (2002)*
67. **Patel, K. J. et al.** Involvement of BRCA2 in DNA repair. *Mol.Cell 1, 347-357 (1998)*
68. **Seal, S. et al.** Evaluation of Fanconi anemia genes in familial breast cancer predisposition. *Cancer Res. 63, 8596-8599 (2003)*
69. **Thompson, E. et al.** A novel duplication polymorphism in the FANCA promoter and its association with breast and ovarian cancer. *BMC.Cancer. 5, 43 (2005)*
70. **Offit, K. et al.** Shared genetic susceptibility to breast cancer, brain tumors, and Fanconi anemia. *J.Natl.Cancer Inst. 95, 1548-1551 (2003)*
71. **Antoniou, A. C. et al.** A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br.J.Cancer 86, 76-83 (2002)*
72. **Roa, B. B., Boyd, A. A., Volcik, K., & Richards, C. S.** Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2. *Nature Genet 14, 185-187 (1996)*
73. **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)*
74. **Antoniou, A. et al.** Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: A combined analysis of 22 studies. *Am.J.Hum.Genet. 72, 1117-1130 (2003)*
75. **Ford, D. et al.** Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. *Am J Hum Genet 62, 676-689 (1998)*
76. **Chen, S. et al.** Characterization of BRCA1 and BRCA2 mutations in a large United States sample. *J.Clin.Oncol. 24, 863-871 (2006)*
77. **Wang, W. W. et al.** A single nucleotide polymorphism in the 5 ' untranslated region of RAD51 and risk of cancer among BRCA1/2 mutation carriers. *Cancer Epidemiol.Biomarkers Prev. 10, 955-960 (2001)*
78. **Kadouri, L. et al.** A single-nucleotide polymorphism in the RAD51 gene modifies breast cancer risk in BRCA2 carriers, but not in BRCA1 carriers or noncarriers. *Br.J.Cancer 90, 2002-2005 (2004)*
79. **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)*
80. **Rebbeck, T. R. et al.** Modification of BRCA1- and BRCA2-associated breast cancer risk by AIB1 genotype and reproductive history. *Cancer Res. 61, 5420-5424 (2001)*
81. **Spurdle, A. B. et al.** The androgen receptor CAG repeat polymorphism and modification of breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res. 7, R176-R183 (2005)*
82. **Kadouri, L. et al.** CAG and GGC repeat polymorphisms in the androgen receptor gene and breast cancer susceptibility in BRCA1/2 carriers and non-carriers. *Br J Cancer 85, 36-40 (2001)*
83. **Phelan, C. M. et al.** Ovarian cancer risk in *BRCA1* carriers is modified by the *HRAS1* variable number of tandem repeat (VNTR) locus. *Nature Genet 12, 309-311 (1996)*
84. **Runnebaum, I. B. et al.** Progesterone receptor variant increases ovarian cancer risk in BRCA1 and BRCA2 mutation carriers who were never exposed to oral contraceptives. *Pharmacogenetics 11, 635-638 (2001)*

85. **Thompson, D., Easton, D., & Breast Cancer Linkage Consortium.** Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. *Am J Hum Genet 68, 410-419 (2001)*
86. **Thompson, D., Easton, D., & Breast Cancer Linkage Consortium.** Variation in BRCA1 cancer risks by mutation position. *Cancer Epidemiol.Biomarkers Prev. 11, 329-336 (2002)*
87. **Van Asperen, C. J. et al.** Cancer risks in BRCA2 families Estimates for sites other than breast and ovary. *J Med Genet 42, 711-719 (2005)*
88. **Easton, D.** Breast cancer genes - What are the real risks? *Nature Genet 16, 210-211 (1997)*
89. **Breast Cancer Linkage Consortium.** Cancer Risks in BRCA2 Mutation Carriers. *J.Natl.Cancer Inst. 91, 1310-1316 (1999)*
90. **Thompson, D., Easton, D. F., & Breast Cancer Linkage Consortium.** Cancer incidence in BRCA1 mutation carriers. *J.Natl.Cancer Inst. 94, 1358-1365 (2002)*
91. **Pluquet, O. & Hainaut, P.** Genotoxic and non-genotoxic pathways of p53 induction. *Cancer Lett. 174, 1-15 (2001)*
92. **Malkin, D. et al.** Germ line P53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science 250, 1233-1238 (1990)*
93. **Frebourg, T. et al.** Germ-line p53 mutations in 15 families with Li-Fraumeni syndrome. *Am.J.Hum.Genet. 56, 608-615 (1995)*
94. **Birch, J. et al.** Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer Res. 54, 1298-1304 (1994)*
95. **Varley, J. M. et al.** Germ-line mutations of TP53 in Li-Fraumeni families: An extended study of 39 families. *Cancer Res. 57, 3245-3252 (1997)*
96. **Evans, D. G. et al.** Low rate of TP53 germline mutations in breast cancer/sarcoma families not fulfilling classical criteria for Li-Fraumeni syndrome. *J Med Genet. 39, 941-944 (2002)*
97. **Strong, L. C., Williams, W. R., & Tainsky, M. A.** The Li-Fraumeni syndrome: from clinical epidemiology to molecular genetics. *Am.J Epidemiol. 135, 190-199 (1992)*
98. **Chompret, A. et al.** P53 germline mutations in childhood cancers and cancer risk for carrier individuals. *Br J Cancer 82, 1932-1937 (2000)*
99. **Garber, J. E. et al.** Follow-up study of twenty-four families with Li-Fraumeni syndrome. *Cancer Res. 51, 6094-6097 (1991)*
100. **Cornelis, R. et al.** Evidence for a gene on 17p13.3, distal to TP53, as a target for allele loss in breast tumors without p53 mutations. *Cancer Res. 54, 4200-4206 (1994)*
101. **Nayak, B. K. & Das, B. R.** Mutation and methylation status of p53 gene promoter in human breast tumours. *Tumour.Biol. 20, 341-346 (1999)*
102. **Rapakko, K. et al.** Germline TP53 alterations in Finnish breast cancer families are rare and occur at conserved mutation-prone sites. *Br.J.Cancer 84, 116-119 (2001)*
103. **Borresen, A. L. et al.** Screening for germ line TP53 mutations in breast cancer patients. *Cancer Res. 52, 3234-3236 (1992).*
104. **Nelen, M. R. et al.** Localization of the gene for Cowden disease to chromosome 10q22-23. *Nature Genet 13, 114-116 (1996)*
105. **Liaw, D. et al.** Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nature Genet 16[1], 64-67. (1997)*
106. **Lindor, N. M. & Greene, M. H.** The concise handbook of family cancer syndromes. Mayo Familial Cancer Program. *J Natl.Cancer Inst. 90, 1039-1071 (1998)*
107. **Mallory, S. B.** Cowden syndrome (multiple hamartoma syndrome). *Dermatol.Clin. 13, 27-31 (1995)*

108. **Schweitzer, S., Hogge, J. P., Grimes, M., Bear, H. D., & de Paredes, E. S.** Cowden disease: a cutaneous marker for increased risk of breast cancer. *AJR Am.J Roentgenol. 172, 349-351 (1999)*
109. **Marsh, D. J. et al.** Mutation spectrum and genotype-phenotype analyses in Cowden disease and Bannayan-Zonana syndrome, two hamartoma syndromes with germline PTEN mutation. *Hum.Mol.Genet. 7, 507-515 (1998)*
110. **Eng, C.** Genetics of Cowden syndrome: through the looking glass of oncology. *Int.J Oncol. 12, 701-710 (1998)*
111. **Starink, T. M. et al**. The Cowden syndrome: a clinical and genetic study in 21 patients. *Clin Genet. 29, 222-233 (1986)*
112. **Fackenthal, J. D. et al.** Male breast cancer in Cowden syndrome patients with germline PTEN mutations. *J.Med.Genet. 38, 159-164 (2001)*
113. **Chen, J. D., Lindblom, P., & Lindblom, A.** A study of the PTEN/MMAC1 gene in 136 breast cancer families. *Hum Genet 102, 124-125 (1998)*
114. **Carroll, B. T., Couch, F. J., Rebbeck, T. R., & Weber, B. L.** Polymorphisms in PTEN in breast cancer families. *J Med Genet. 36, 94-96 (1999)*
115. **Rhei, E. et al.** Mutation analysis of the putative tumor suppressor gene PTEN/MMAC1 in primary breast carcinomas. *Cancer Res. 57, 3657-3659 (1997)*
116. **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)*
117. **Feilotter, H. E. et al.** Analysis of the 10q23 chromosomal region and the PTEN gene in human sporadic breast carcinoma. *Br J Cancer 79, 718-723 (1999)*
118. **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)*
119. **Kallioniemi, O. P. et al.** Comparative genomic hybridization: a rapid new method for detecting and mapping DNA amplification in tumors. *Semin.Cancer Biol. 4, 41-46 (1993)*
120. **Jenne, D. E. et al.** Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nature Genet 18, 38-43 (1998)*
121. **Wada, K., Tanaka, M., Yamaguchi, K., & Wada, K.** Carcinoma and polyps of the gallbladder associated with Peutz-Jeghers syndrome. *Dig.Dis.Sci. 32, 943-946 (1987)*
122. **Westerman, A. M. et al.** Peutz-Jeghers syndrome: 78-year follow-up of the original family. *Lancet. 353, 1211-1215 (1999)*
123. **Giardiello, F. M. et al.** Increased risk of cancer in the Peutz-Jeghers syndrome. *N.Engl.J Med. 316, 1511-1514 (1987)*
124. **Spigelman, A. D., Murday, V., & Phillips, R. K.** Cancer and the Peutz-Jeghers syndrome. *Gut. 30, 1588-1590 (1989)*
125. **Boardman, L. A. et al.** Increased risk for cancer in patients with the Peutz-Jeghers syndrome. *Ann.Intern.Med. 128, 896-899 (1998)*
126. **Lim, W. et al.** Further observations on LKB1/STK11 status and cancer risk in Peutz-Jeghers syndrome. *Br.J Cancer. 89, 308-313 (2003)*
127. **Giardiello, F. M. et al.** Very high risk of cancer in familial Peutz-Jeghers syndrome. *Gastroenterology. 119, 1447-1453 (2000)*
128. **Shen, Z., Wen, X. F., Lan, F., Shen, Z. Z., & Shao, Z. M.** The tumor suppressor gene LKB1 is associated with prognosis in human breast carcinoma. *Clin Cancer Res. 8, 2085-2090 (2002)*
129. **Bignell, G. R. et al.** Low frequency of somatic mutations in the LKB1/Peutz- Jeghers syndrome gene in sporadic breast cancer. *Cancer Res. 58, 1384-1386 (1998)*

130. **Guilford, P. et al.** E-cadherin germline mutations in familial gastric cancer. *Nature 392, 402-405 (1998)*
131. **Pharoah, P. D., Guilford, P., & Caldas, C.** Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. *Gastroenterology. 121, 1348-1353 (2001)*
132. **Berx, G. et al.** E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. *Oncogene 13, 1919-1925 (1996)*
133. **Chan, J. K. & Wong, C. S.** Loss of E-cadherin is the fundamental defect in diffuse-type gastric carcinoma and infiltrating lobular carcinoma of the breast. *Adv.Anat.Pathol. 8, 165-172 (2001)*
134. **Graziano, F., Humar, B., & Guilford, P.** The role of the E-cadherin gene (CDH1) in diffuse gastric cancer susceptibility: from the laboratory to clinical practice. *Ann.Oncol. 14, 1705-1713 (2003)*
135. **Mastracci, T. L., Tjan, S., Bane, A. L., O'Malley, F. P., & Andrulis, I. L.** E-cadherin alterations in atypical lobular hyperplasia and lobular carcinoma in situ of the breast. *Mod.Pathol. 18, 741-751 (2005)*
136. **Sarrio, D. et al.** Epigenetic and genetic alterations of APC and CDH1 genes in lobular breast cancer: relationships with abnormal E-cadherin and catenin expression and microsatellite instability. *Int.J Cancer. %20;106, 208-215 (2003)*
137. **Suriano, G. et al.** Characterization of a recurrent germ line mutation of the E-cadherin gene: implications for genetic testing and clinical management. *Clin Cancer Res. 11, 5401-5409 (2005)*
138. **Keller, G. et al.** Diffuse type gastric and lobular breast carcinoma in a familial gastric cancer patient with an E-cadherin germline mutation. *Am.J.Pathol. 155, 337-342 (1999)*
139. **Jonsson, B. A., Bergh, A., Stattin, P., Emmanuelsson, M., & Gronberg, H.** Germline mutations in E-cadherin do not explain association of hereditary prostate cancer, gastric cancer and breast cancer. *Int.J Cancer. %20;98, 838-843 (2002)*
140. **Salahshor, S. et al.** Low frequency of E-cadherin alterations in familial breast cancer. *Breast Cancer Res. 3, 199-207 (2001)*
141. **Bakkenist, C. J. & Kastan, M. B.** DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature. 421, 499-506 (2003)*
142. **Pippard, E. C., Hall, A. J., Barker, D. J., & Bridges, B. A.** Cancer in homozygotes and heterozygotes of ataxia-telangiectasia and xeroderma pigmentosum in Britain. *Cancer Res. 48, 2929-2932 (1988)*
143. **Swift, M., Chase, C., & Morrell, D.** Cancer predisposition of Ataxia-Telangiectasia heterozygotes. *Cancer Genet.Cytogenet. 46, 21-27 (1990)*
144. **Easton, D. F.** Cancer risks in A-T heterozygotes. *Int J Radiat Biol 66, S177-S182 (1994)*
145. **Hall, J.** The Ataxia-telangiectasia mutated gene and breast cancer: gene expression profiles and sequence variants. *Cancer Lett. 227, 105-114 (2005)*
146. **Thompson, D. et al.** Cancer risks and mortality in heterozygous ATM mutation carriers. *J.Natl.Cancer Inst. 97, 813-822 (2005)*
147. **Renwick, A. et al.** ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet 38, 873-875 (2006)*
148. **Vorechovsky, I. et al.** The ATM gene and susceptibility to breast cancer: analysis of 38 breast tumors reveals no evidence for mutation. *Cancer Res. 56, 2726-2732 (1996)*
149. **FitzGerald, M. G. et al.** Heterozygous ATM mutations do not contribute to early onset of breast cancer. *Nature Genet. 15, 307-310 (1997)*
150. **Broeks, A. et al.** ATM-heterozygous germline mutations contribute to breast cancer-susceptibility. *Am J Hum Genet 66, 494-500 (2000)*
151. **Chenevix-Trench, G. et al.** Dominant Negative ATM Mutations in Breast Cancer Families. *J.Natl.Cancer Inst. 94, 205-215 (2002)*

152. **Szabo, C. I. et al.** Are ATM mutations 7271T-->G and IVS10-6T-->G really high-risk breast cancer-susceptibility alleles? *Cancer Res. 64, 840-843 (2004)*
153. **Thompson, D. et al.** Two ATM variants and breast cancer risk. *Hum.Mutat. 25, 594-595 (2005)*
154. **Gatti, R. A., Tward, A., & Concannon, P.** Cancer risk in ATM heterozygotes: a model of phenotypic and mechanistic differences between missense and truncating mutations. *Mol.Genet Metab. 68, 419-423 (1999)*
155. **Meyn, M. S.** Ataxia-telangiectasia, cancer and the pathobiology of the ATM gene. *Clin Genet. 55, 289-304 (1999)*
156. **Bretsky, P. et al.** The relationship between twenty missense ATM variants and breast cancer risk: The multiethnic cohort. *Cancer Epidemiol.Biomarkers Prev. 12, 733-738 (2003)*
157. **Tamimi, R. M. et al.** Common ataxia telangiectasia mutated haplotypes and risk of breast cancer: a nested case-control study. *Breast Cancer Res. 6, R416-R422 (2004)*
158. **Bernstein, J. L. et al.** Population-based estimates of breast cancer risks associated with ATM gene variants c.7271T>G and c.1066-6T>G (IVS10-6T>G) from the Breast Cancer Family Registry. *Hum.Mutat. 27, 1122-1128 (2006)*
159. **Rosfjord, E. C. & Dickson, R. B.** Growth factors, apoptosis, and survival of mammary epithelial cells. *J Mammary.Gland.Biol.Neoplasia. 4, 229-237 (1999)*
160. **Janda, E. et al.** Ras and TGF[beta] cooperatively regulate epithelial cell plasticity and metastasis: dissection of Ras signaling pathways. *J Cell Biol. 156, 299-313 (2002)*
161. **Lehmann, K. et al.** Raf induces TGFbeta production while blocking its apoptotic but not invasive responses: a mechanism leading to increased malignancy in epithelial cells. *Genes Dev. 14, 2610-2622 (2000)*
162. **Oft, M., Akhurst, R. J., & Balmain, A.** Metastasis is driven by sequential elevation of H-ras and Smad2 levels. *Nat.Cell Biol. 4, 487-494 (2002)*
163. **Schulze, A., Lehmann, K., Jefferies, H. B., McMahon, M., & Downward, J.** Analysis of the transcriptional program induced by Raf in epithelial cells. *Genes Dev. 15, 981-994 (2001)*
164. **Goumans, M. J. et al.** Balancing the activation state of the endothelium via two distinct TGF-beta type I receptors. *EMBO J. 21, 1743-1753 (2002)*
165. **Derynck, R., Akhurst, R. J., & Balmain, A.** TGF-beta signaling in tumor suppression and cancer progression. *Nat.Genet. 29, 117-129 (2001)*
166. **Lucke, C. D. et al.** Inhibiting mutations in the transforming growth factor beta type 2 receptor in recurrent human breast cancer. *Cancer Res. 61, 482-485 (2001)*
167. **Chen, T., Carter, D., Garrigue-Antar, L., & Reiss, M.** Transforming growth factor beta type I receptor kinase mutant associated with metastatic breast cancer. *Cancer Res. 58, 4805-4810 (1998)*
168. **Xie, W. et al.** Alterations of Smad signaling in human breast carcinoma are associated with poor outcome: a tissue microarray study. *Cancer Res. 62, 497-505 (2002)*
169. **Dunning, A. M. et al.** A transforming growth factorbeta1 signal peptide variant increases secretion in vitro and is associated with increased incidence of invasive breast cancer. *Cancer Res. 63, 2610-2615 (2003)*
170. **Chun, H. J. et al.** Pleiotropic defects in lymphocyte activation caused by caspase-8 mutations lead to human immunodeficiency. *Nature. 419, 395-399 (2002)*
171. **MacPherson, G. et al.** Association of a common variant of the CASP8 gene with reduced risk of breast cancer. *J.Natl.Cancer Inst. 96, 1866-1869 (2004)*
172. **Cox, A. et al.** A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet. 39, 352-358 (2007)*

173. **Frank, B. et al.** Association of the CASP10 V410I variant with reduced familial breast cancer risk and interaction with the CASP8 D302H variant. *Carcinogenesis 27, 606-609 (2006)*
174. **Sodha, N. et al.** Increasing evidence that germline mutations in CHEK2 do not cause Li- Fraumeni syndrome. *Hum.Mutat. 20, 460-462 (2002)*
175. **Meijers-Heijboer, H. et al.** Low-penetrance susceptibility to breast cancer due to CHEK2*1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nature Genet. 31, 55-59 (2002)*
176. **Vahteristo, P. et al.** A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. *Am.J.Hum.Genet.* **71**, 432-438 (2002).
177. **Oldenburg, R. A. et al.** The CHEK2*1100delC variant acts as a breast cancer risk modifier in non-BRCA1/BRCA2 multiple-case families. *Cancer Res. 63, 8153-8157 (2003)*
178. **Meijers-Heijboer, H. et al.** The CHEK2 1100delC Mutation Identifies Families with a Hereditary Breast and Colorectal Cancer Phenotype. *Am.J.Hum.Genet. 72, 1308-1314 (2003)*
179. **The CHEK2 Breast Cancer Case-Control Consortium.** CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from ten studies. *Am.J.Hum.Genet. 74, 1175-1182 (2004)*
180. **Schutte, M. et al.** Variants in CHEK2 other than 1100delC do not make a major contribution to breast cancer susceptibility. *Am J Hum Genet 72, 1023-1028 (2003)*
181. **Dufault, M. R. et al.** Limited relevance of the CHEK2 gene in hereditary breast cancer. *Int.J.Cancer* **110**, 320-325 (2004).
182. **Bogdanova, N. et al.** Association of two mutations in the CHEK2 gene with breast cancer. | *Int.J.Cancer 116, 263-266 (2005)*
183. **Kilpivaara, O. et al.** CHEK2 variant 1157T may be associated with increased breast cancer risk. *Int.J.Cancer 111, 543-547 (2004)*
184. **de Bock, G. H. et al.** Tumour characteristics and prognosis of breast cancer patients carrying the germline CHEK2*1100delC variant. *J.Med.Genet. 41, 731-735 (2004)*
185. **Johnson, N. et al.** Interaction between CHEK2*1100delC and other low-penetrance breast-cancer susceptibility genes: a familial study. *Lancet 366, 1554-1557 (2005)*
186. **Huzarski, T. et al.** Pathology of breast cancer in women with constitutional CHEK2 mutations. *Breast Cancer Res Treatm 90, 187-189 (2005)*
187. **Oldenburg, R. A. et al.** Characterization of familial non-*BRCA1*/2 breast tumors by loss of heterozygosity and immunophenotyping. *Clin.Cancer Res. 12, 1693-1700 (2006)*
188. **Wu, L. J. C. et al.** Identification of a ring protein that can interact in vivo with the BRCA1 gene product. *Nature Genet 14, 430-440 (1996)*
189. **Meza, J. E., Brzovic, P. S., King, M. C., & Klevit, R. E.** Mapping the functional domains of BRCA1 - Interaction of the ring finger domains of BRCA1 and BARD1. *J.Biol.Chem. 274, 5659-5665 (1999)*
190. **Joukov, V., Chen, J., Fox, E. A., Green, J. B. A., & Livingston, D. M.** Functional communication between endogenous BRCA1 and its partner, BARD1, during Xenopus laevis development. *Proc.Natl.Acad.Sci.USA 98, 12078-12083 (2001)*
191. **Irminger, F., I et al.** Identification of BARD1 as mediator between proapoptotic stress and p53-dependent apoptosis. *Mol.Cell 8, 1255-1266 (2001)*
192. **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)*
193. **Ghimenti, C. et al.** Germline mutations of the BRCA1-associated ring domain (BARD1) gene in breast and breast/ovarian families negative for BRCA1 and BRCA2 alterations. *Genes Chrom.Cancer 33, 235-242 (2002)*

194. **Karppinen, S. M. et al.** Nordic collaborative study of the BARD1 Cys557Ser allele in 3956 patients with cancer: enrichment in familial BRCA1/BRCA2 mutation-negative breast cancer but not in other malignancies. *J.Med.Genet. 43, 856-862 (2006)*
195. **Mathew, C. G.** Fanconi anaemia genes and susceptibility to cancer. *Oncogene 25, 5875-5884 (2006)*
196. **Hirsch, B. et al.** Association of biallelic BRCA2/FANCD1 mutations with spontaneous chromosomal instability and solid tumors of childhood. *Blood 103, 2554-2559 (2004)*
197. **Reid, S. et al.** Biallelic BRCA2 mutations are associated with multiple malignancies in childhood including familial Wilms tumour. *J Med Genet 42, 147-151 (2005)*
198. **Lewis, A. G. et al.** Mutation analysis of FANCD2, BRIP1/BACH1, LMO4 and SFN in familial breast cancer. *Breast Cancer Res. 7, R1005-R1016 (2005)*
199. **Barroso, E. et al.** FANCD2 associated with sporadic breast cancer risk. *Carcinogenesis 27, 1930-1937 (2006)*
200. **Cantor, S. B. et al.** BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell 105, 149-160 (2001)*
201. **Cantor, S. et al.** The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. *Proceedings of the National Academy of Sciences of the United States of America 101, 2357-2362 (2004)*
202. **Peng, M., Litman, R., Jin, Z., Fong, G., & Cantor, S. B.** BACH1 is a DNA repair protein supporting BRCA1 damage response. *Oncogene 25, 2245-2253 (2006)*
203. **Luo, L. P. et al.** No mutations in the BACH1 gene in BRCA1 and BRCA2 negative breast-cancer families linked to 17q22. *Int.J.Cancer 98, 638-639 (2002)*
204. **Karppinen, S. M., Vuosku, J., Heikkinen, K., Allinen, M., & Winqvist, R.** No evidence of involvement of germline BACH1 mutations in Finnish breast and ovarian cancer families. *Eur.J.Cancer 39, 366-371 (2003)*
205. **Rutter, J. L. et al.** Mutational analysis of the BRCA1-interacting genes ZNF350/ZBRK1 and BRIP1/BACH1 among BRCA1 and BRCA2-negative probands from breast-ovarian cancer families and among early-onset breast cancer cases and reference individuals. *Hum Mutat 22, 121-128 (2003)*
206. **Vahteristo, P. et al.** BACH1 Ser919Pro variant and breast cancer risk. *BMC. Cancer 1-7 (2005)*
207. **Vahteristo, P. et al.** BARD1 variants Cys557Ser and Val507Met in breast cancer predisposition. *Eur J Hum Genet 14, 167-172 (2006)*
208. **Seal, S. et al.** Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet. 38, 1239-1241 (2006)*
209. **Xia, B. et al.** Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol.Cell 22, 719-729 (2006)*
210. **Reid, S. et al.** Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat.Genet 39, 162-164 (2007)*
211. **Rahman, N. et al.** PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet. 39, 165-167 (2007)*
212. **Shih, H. A. et al.** BRCA1 and BRCA2 mutation frequency in women evaluated in a breast cancer risk evaluation clinic. *J Clin Oncol 20, 994-999 (2002)*
213. **Verhoog, L. C. et al.** Large regional differences in the frequency of distinct BRCA1/BRCA2 mutations in 517 Dutch breast and/or ovarian cancer families. *Eur J Cancer 37, 2082-2090 (2001)*
214. **FitzGerald, M. G. et al.** Germline mutations in PTEN are an infrequent cause of genetic predisposition to breast cancer. *Oncogene 17, 727-731 (1998)*

215. **Claus, E. B., Risch, N., & Douglas, W.** Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am.J.Hum.Genet. 48, 232-242 (1991)*
216. **Go, R. C., King, M. C., Bailey-Wilson, J., Elston, R. C., & Lynch, H. T.** Genetic epidemiology of breast cancer and associated cancers in high-risk families. I. Segregation analysis. *J Natl.Cancer Inst. 71, 455-461 (1983)*
217. **Eccles, D., Marlow, A., Royle, G., Collins, A., & Morton, N. E.** Genetic epidemiology of early onset breast cancer. *J Med Genet. 31, 944-949 (1994)*
218. **Miki, Y. et al.** A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science 266, 66-71 (1994)*
219. **Wooster, R. et al.** Identification of the breast cancer susceptibility gene BRCA2. *Nature 378, 789-792 (1995)*
220. **Cui, J. & Hopper, J. L.** Distribution of family history of a disease as a function of mode of inheritance, genetic relative hazard, allele frequency and disease status of the proband, with application to female breast cancer. *J Epidemiol.Biostat. 6, 331-342 (2001)*
221. **Antoniou, A. C. et al.** Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. *Genet.Epidemiol. 21, 1-18 (2001)*
222. **Hall, J. et al.** Linkage of early-onset familial breast cancer to chromosome 17q21. *Science 250, 1684-1689 (1990)*
223. **Narod, S. et al.** An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. *Am J Hum Genet 56, 254-264 (1995)*
224. **Easton, D., Bishop, D., Ford, D., Crockford, G., & Breast Cancer Linkage Consortium.** Genetic linkage analysis in familial breast and ovarian cancer: Results from 214 families. *Am.J.Hum.Genet. 52, 678-701 (1993)*
225. **Stratton, M. et al.** Familial male breast cancer is not linked to the BRCA1 locus on chromosome 17q. *Nature Genet 7, 103-107 (1994)*
226. **Wooster, R. et al.** Localization of a breast cancer susceptibility gene BRCA2 to chromosome 13q12-13. *Science 265, 2088-2090 (1994)*
227. **Miller, B. J., Wang, D., Krahe, R., & Wright, F. A.** Pooled analysis of loss of heterozygosity in breast cancer: a genome scan provides comparative evidence for multiple tumor suppressors and identifies novel candidate regions. *Am J Hum Genet 73, 748-767 (2003)*
228. **Kainu, T. et al.** Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proc.Natl.Acad.Sci.USA 97, 9603-9608 (2000)*
229. **Kerangueven, F. et al.** Loss of heterozygosity and linkage analysis in breast carcinoma: Indication for a putative third susceptibility gene on the short arm of chromosome 8. *Oncogene 10, 1023-1026 (1995)*
230. **Eisinger, F. et al.** Germ line mutation at BRCA1 affects the histoprognostic grade in hereditary breast cancer. *Cancer Res. 56, 471-474 (1996)*
231. **Seitz, S. et al.** Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: linkage analysis in German breast cancer families. *Oncogene 14, 741-743 (1997)*
232. **Seitz, S. et al.** Deletion mapping and linkage analysis provide strong indication for the involvement of the human chromosome region 8p12-p22 in breast carcinogenesis. *Br J Cancer 76, 983-991 (1997)*
233. **Rahman, N. et al.** Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. *Oncogene 19, 4170-4173 (2000)*
234. **Thompson, D. et al.** Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. *Proc.Natl.Acad.Sci.USA 99, 827-831 (2002)*

235. **Smith, P. et al.** A genome wide linkage search for breast cancer susceptibility genes. *Genes Chrom.Cancer 45, 646-655 (2006)*
236. **Huusko, P. et al.** Genome-wide scanning for linkage in Finnish breast cancer families. *Eur.J.Hum.Genet. 12, 98-104 (2004)*
237. **Lakhani, S. R. et al.** Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J.Natl.Cancer Inst. 90, 1138-1145 (1998)*
238. **Lakhani, S. R. et al.** The pathology of familial breast cancer: histological features of cancers in families not attributable to mutations in BRCA1 or BRCA2. *Clin.Cancer Res. 6, 782-789 (2000)*
239. **Lakhani, S. R.** The pathology of familial breast cancer: Morphological aspects. *Breast Cancer Res. 1, 31-35 (1999)*
240. **Catteau, A., Harris, W. H., Xu, C. F., & Solomon, E.** Methylation of the BRCA1 promotor region in sporadic breast and ovarian cancer: correlation with disease characteristics. *Oncogene 18, 1957-1965 (1999)*
241. **Esteller, M. et al.** Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J.Natl.Cancer Inst. 92, 564-569 (2000)*
242. **Hamann, U. & Sinn, H. P.** Survival and tumor characteristics of German hereditary breast cancer patients. *Breast Cancer Res.Treat. 59, 185-192 (2000)*
243. **Stoppa-Lyonnet, D. et al.** Familial invasive breast cancers: worse outcome related to BRCA1 mutations. *J Clin Oncol. 18, 4053-4059 (2000)*
244. **Robson, M. et al.** BRCA-associated breast cancer in young women. *J Clin Oncol 16, 1642-1649 (1998)*
245. **Breast Cancer Linkage Consortium.** Pathology of familial breast cancer: Differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. *Lancet 349, 1505-1510 (1997)*
246. **Callahan, R. et al.** Genetic and molecular heterogeneity of breast cancer cells. *Clin.Chim.Acta 217, 63-73 (1993)*
247. **Lasko, D., Cavenee, W., & Nordenskjöld, M.** Loss of constitutional heterozygosity in human cancer. *Annu.Rev.Genet. 25, 281-314 (1991)*
248. **Newton, M. A., Gould, M. N., Reznikoff, C. A., & Haag, J. D.** On the statistical analysis of allelic-loss data. *Stat.Med. 17, 1425-1445 (1998)*
249. **Newton, M. A. & Lee, Y.** Inferring the location and effect of tumor suppressor genes by instability-selection modeling of allelic-loss data. *Biometrics. 56, 1088-1097 (2000).*
250. **Osorio, A. et al.** Loss of heterozygosity analysis at the BRCA loci in tumor samples from patients with familial breast cancer. *Int.J.Cancer 99, 305-309 (2002).*
251. **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).*
252. **Tirkkonen, M. et al.** Somatic genetic alterations in BRCA2-associated and sporadic male breast cancer. *Genes Chrom Cancer 24, 56-61 (1999)*
253. **Devilee, P., Cleton-Jansen, A. M., & Cornelisse, C. J.** Ever since Knudson. *Trends Genet 17, 569-573 (2001)*
254. **Zhuang, Z. et al.** Trisomy 7-harbouring non-random duplication of the mutant MET allele in hereditary papillary renal carcinomas. *Nature Genet. 20, 66-69 (1998)*
255. **Wessels, L. F. A. et al.** Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. *Cancer Res. 62, 7110-7117 (2002)*
256. **van Beers, E. H. et al.** Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. *Cancer Res. 65, 822-827 (2005)*

257. **Jonsson, G. et al.** Distinct Genomic Profiles in Hereditary Breast Tumors Identified by Array-Based Comparative Genomic Hybridization. *Cancer Res. 65, 7612-7621 (2005)*
258. **Gronwald, J. et al.** Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. *Int.J.Cancer 114, 230-236 (2005)*
259. **Lakhani, S. R. et al.** The pathology of familial breast cancer: Predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol 20, 2310-2318 (2002)*
260. **Palacios, J. et al.** Immunohistochemical characteristics defined by tissue microarray of hereditary breast cancer not attributable to BRCA1 or BRCA2 mutations: Differences from breast carcinomas arising in BRCA1 and BRCA2 mutation carriers. *Clinical Cancer Research 9, 3606-3614 (2003)*
261. **Sorlie, T. et al.** Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc.Natl.Acad.Sci.USA 100, 8418-8423 (2003)*
262. **Hedenfalk, I. et al.** Molecular classification of familial non-BRCA1/BRCA2 breast cancer. *Proc.Natl.Acad.Sci.USA 100, 2532-2537 (2003)*
263. **Abd El-Rehim, D. M. et al.** Expression of luminal and basal cytokeratins in human breast carcinoma. *J.Pathol. 203, 661-671 (2004)*
264. **Bergman, A. et al.** Genome-wide linkage scan for breast cancer susceptibility loci in Swedish hereditary non-BRCA1/2 families: suggestive linkage to 10q23.32-q25.3. *Genes Chrom.Cancer 46, 302-309 (2007)*
265. **Schadt, E. E.** Exploiting naturally occurring Dna variation and molecular profiling data to dissect disease and drug response traits. *Current Opinion in Biotechnology 16, 647-654 (2005)*
266. **Sorlie, T. et al.** Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc.Natl.Acad.Sci.USA 98, 10869-10874 (2001)*
267. **Hedenfalk, I. et al.** Gene-expression profiles in hereditary breast cancer. *N.Engl.J.Med. 344, 539-548 (2001)*
268. **Vierimaa, O. et al.** Pituitary adenoma predisposition caused by germline mutations in the AIP gene. *Science 312, 1228-1230 (2006)*
269. **Calin, G. A. & Croce, C. M.** MicroRNA signatures in human cancers. *Nat Rev.Cancer 6, 857-866 (2006)*
270. **Pharoah, P. D., Dunning, A. M., Ponder, B. A., & Easton, D. F.** Association studies for finding cancer-susceptibility genetic variants. *Nat.Rev.Cancer 4, 850-860 (2004)*
271. **Risch, N. J.** Searching for genetic determinants in the new millennium. *Nature. 405, 847-856 (2000)*
272. **Cardon, L. R. & Bell, J. I.** Association study designs for complex diseases. *Nat.Rev.Genet. 2, 91-99 (2001)*
273. **Wang, W. Y., Barratt, B. J., Clayton, D. G., & Todd, J. A.** Genome-wide association studies: theoretical and practical concerns. *Nat Rev.Genet. 6, 109-118 (2005)*
274. **Hirschhorn, J. N. & Daly, M. J.** Genome-wide association studies for common diseases and complex traits. *Nat Rev.Genet. 6, 95-108 (2005)*
275. **Rioux, J. D. et al.** Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat.Genet. 29, 223-228 (2001)*
276. **Stefansson, H. et al.** Neuregulin 1 and susceptibility to schizophrenia. *Am.J Hum.Genet. 71, 877-892 (2002)*
277. **Altshuler, D. et al.** The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat.Genet. 26, 76-80 (2000)*

278. **Peto, J. et al.** Prevalence of BRCA1 and BRCA2 gene mutations in patients with early- onset breast cancer. *J.Natl.Cancer Inst. 91, 943-949 (1999)*
279. **Lichtenstein, P. et al.** Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N.Engl.J.Med. 343, 78-85 (2000)*
280. **Peto, J. & Mack, T. M.** High constant incidence in twins and other relatives of women with breast cancer. *Nature Genet 26, 411-414 (2000)*
281. **Johnson, N. et al.** Counting potentially functional variants in BRCA1, BRCA2 and ATM predicts breast cancer susceptibility. *Hum.Mol.Genet. 16, 1051-1057 (2007)*
282. **Baynes, C. et al.** Common variants in the ATM, BRCA1, BRCA2, CHEK2 and TP53 cancer susceptibility genes are unlikely to increase breast cancer risk. *Breast Cancer Res. 9, R27 (2007)*
283. **Patil, N. et al.** Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science. 294, 1719-1723 (2001)*
284. **Johnson, G. C. et al.** Haplotype tagging for the identification of common disease genes. *Nat.Genet. 29, 233-237 (2001)*
285. **Gabriel, S. B. et al.** The structure of haplotype blocks in the human genome. *Science. 296, 2225-2229 (2002)*
286. **Kruglyak, L.** Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genet. 22, 139-144 (1999)*
287. **Hinds, D. A. et al.** Whole-genome patterns of common Dna variation in three human populations. *Science 307, 1072-1079 (2005)*
288. **Altshuler, D. et al.** A haplotype map of the human genome. *Nature 437, 1299-1320 (2005)*
289. **Haiman, C. A. et al.** A comprehensive haplotype analysis of CYP19 and breast cancer risk: the Multiethnic Cohort. *Hum.Mol.Genet. 12, 2679-2692 (2003)*
290. **Freedman, M. L. et al.** Common variation in BRCA2 and breast cancer risk: a haplotype-based analysis in the Multiethnic Cohort. *Hum.Mol.Genet. 13, 2431-2441 (2004)*
291. **Antoniou, A. C. & Easton, D. F.** Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet.Epidemiol. 25, 190-202 (2003)*
292. **Houlston, R. S. & Peto, J.** The future of association studies of common cancers. *Hum.Genet. 112, 434-435 (2003)*
293. **Roberts, S. A. et al.** Heritability of cellular radiosensitivity: a marker of low-penetrance predisposition genes in breast cancer? *Am J Hum Genet 65, 784-794 (1999)*
294. **Easton, D. F. et al.** Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature 447, 1087-1095 (2007)*
295. **Wright, A., Charlesworth, B., Rudan, I., Carothers, A., & Campbell, H. A** polygenic basis for late-onset disease. *Trends Genet. 19, 97-106 (2003)*
296. **Pharoah, P. D. P. et al.** Polygenic susceptibility to breast cancer and implications for prevention. *Nature Genet. 31, 33-36 (2002)*
297. **Munoz-Gamez, J. A. et al.** PARP inhibition sensitizes p53-deficient breast cancer cells to doxorubicin-induced apoptosis. *Biochem.J. 386, 119-125 (2005)*
298. **Hay, T. et al.** Efficient deletion of normal Brca2-deficient intestinal epithelium by poly(ADP-ribose) polymerase inhibition models potential prophylactic therapy. *Cancer Res. 65, 10145-10148 (2005)*
299. **Farmer, H. et al.** Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature 434, 917-921 (2005)*
300. **Bryant, H. E. et al.** Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature 434, 913-917 (2005)*

301. **Ford, D. et al.** Risks of cancer in BRCA1-mutation carriers. *Lancet 343, 692-695 (1994)*
302. **King, M. C., Go, R. C., Elston, R. C., Lynch, H. T., & Petrakis, N. L.** Allele increasing susceptibility to human breast cancer may be linked to the glutamate-pyruvate transaminase locus. *Science. 208, 406-408 (1980)*
303. **Skolnick, M. H., Thompson, E. A., Bishop, D. T., & Cannon, L. A.** Possible linkage of a breast cancer-susceptibility locus to the ABO locus: sensitivity of LOD scores to a single new recombinant observation. *Genet Epidemiol. 1, 363-373 (1984).*
304. **Zuppan, P., Hall, J., Lee, M., Ponglikitmongkol, M., & King, M.-C.** Possible linkage of the estrogen receptor gene to breast cancer in a family with late-onset disease. *Am.J.Hum.Genet. 48, 1065-1068 (1991)*
305. **Ellis, N. A. et al.** Localization of breast cancer susceptibility loci by genome-wide SNP linkage disequilibrium mapping. *Genet.Epidemiol. 30, 48-61 (2006)*
306. **Honrado, E. et al.** Immunohistochemical expression of DNA repair proteins in familial breast cancer differentiate BRCA2-associated tumors. *J Clin Oncol. %20;23, 7503-7511 (2005)*
307. **Eerola, H. et al.** Histopathological features of breast tumours in BRCA1, BRCA2 and mutation-negative breast cancer families. *Breast Cancer Res 7, R93-R100 (2004)*
308. **Easton, D., Ford, D., Bishop, D., & Breast Cancer Linkage Consortium.** Breast and ovarian cancer incidence in BRCA1-mutation carriers. *Am J Hum Genet 56, 265-271 (1995)*