Search for new breast cancer susceptibility genes
Oldenburg, R.A.

GENERAL INTRODUCTION

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

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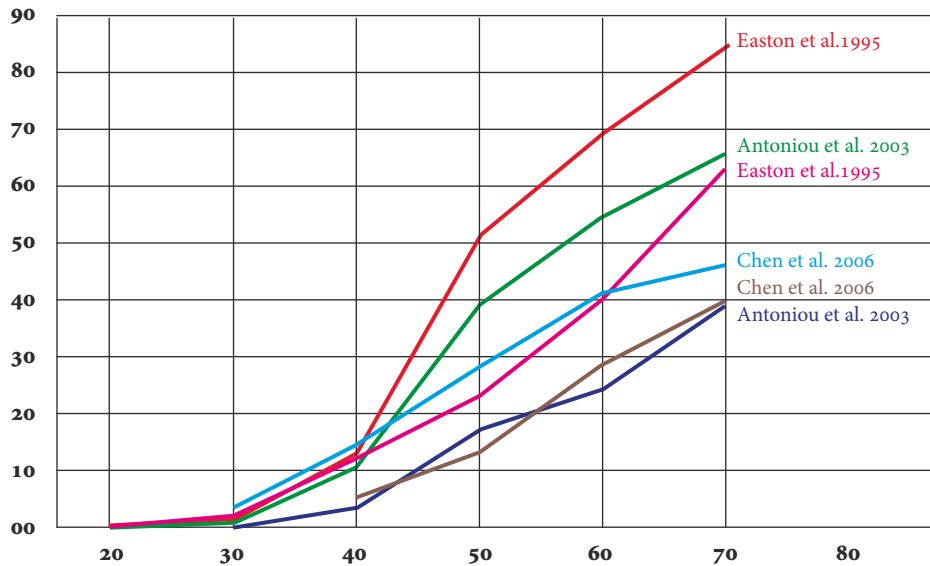
Fig. 1. Cumulative breast and ovarian cancer risk in BRCA1-mutation carriers as a function of age.
The red and pink line respectively represent family-based breast and ovarian cancer risk estimates (Easton et al.[274]).
The green / light blue and dark blue / brown lines respectively represent population-based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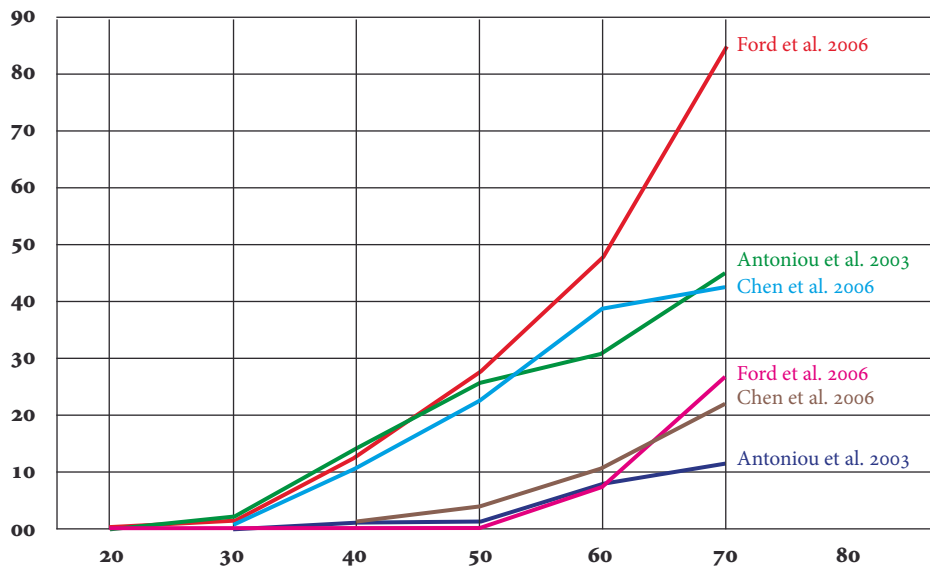

Fig. 2. Cumulative breast and ovarian cancer risk in BRCA2-mutation carriers as a function of age.
The red and pink line respectively represent family based breast and ovarian cancer risk estimates (Ford et al.[75]. The
green / light blue and dark blue / brown lines respectively represent population-based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

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 x 10^{-7}) and 0.74 (95% CI = 0.62-0.87, P_{trend} = 1.1 x 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
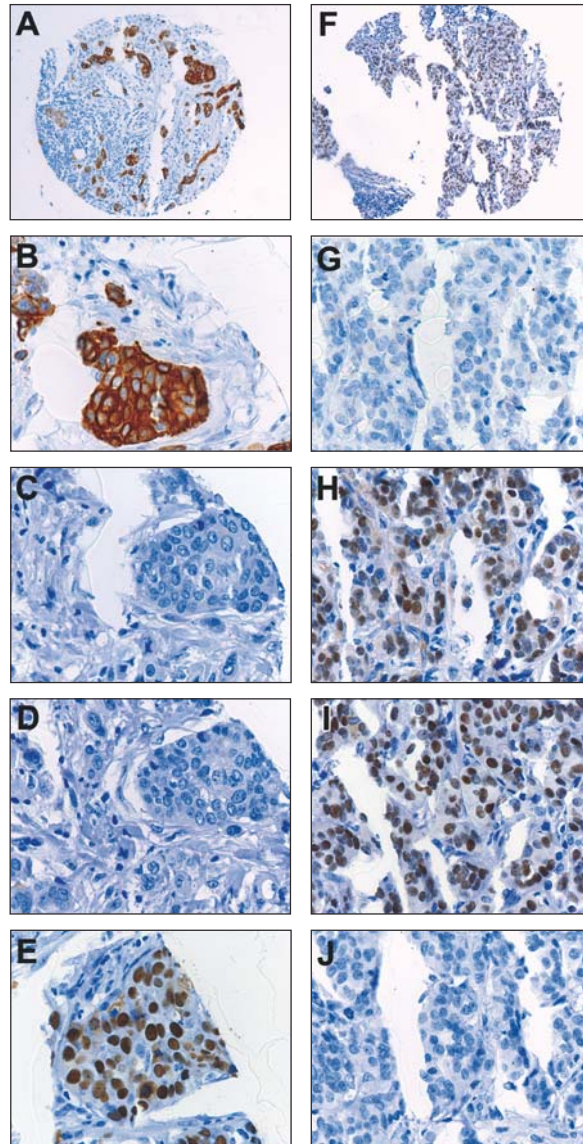
Fig. 3. Immunohistochemical staining results of a BRCA1- and BRCA2- related breast tumor on a tissue microarray. The samples A, B, C, D and E are from one BRCA1- (2315del5) tumor and the samples F, G, H, I and J of one BRCA2- (6648insA) tumor, both on the same paraffin tissue microarray block. A and 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).