Search for new breast cancer susceptibility genes

Oldenburg, R.A.

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.