Genetic studies of hereditary breast/ovarian cancer; localization and characterization of novel genes in the west Swedish population

PI:
Margareta Nordling
Introduction

Breast cancer is the most common type of cancer among women worldwide, corresponding to 23% of all female cancer (Parkin et al., 2005). Approximately one in ten women will be affected with breast cancer in our country and only Northern America has a higher incidence rate (Parkin, et al. 2005). Breast cancer incidence is 4-5 folds higher in developing countries than in Asia or Africa. Studies have shown that the risk increase for women from low incidence countries as they move to the west which may indicate that life style such as body size, alcohol, exogenous hormones and exposure to radiation has an influence on breast cancer risk. Other main risk factors are high age and endogenous hormonal exposure such as early menstruation and late menopause. Family history of the disease contributes to elevate a woman’s individual risk. Multiple cases of breast cancer within a family are an indication of an inherited predisposition to breast cancer. The proportion of familial cancer is approximately 5-10% (McPherson et al., 2000) of all breast cancer cases. Population-based studies have shown that approximately 15-20% of the familial cases can be explained by a mutation in one of the high penetrant breast cancer susceptibility genes \textit{BRCA1} and \textit{BRCA2} (Peto et al., 1999; Balmain et al., 2003). Other known susceptibility genes cause a minor proportion of non-\textit{BRCA1}/\textit{2} families but the majority is affected by yet unknown genetic factors (Fig. 1).

Breast cancer is a genetic disease like any other type of cancer. The development of a breast tumor is likely to occur through a similar multi-step process like the one proposed for colon cancer involving several mutational events. Although the exact pathway is uncertain, several oncogenes are activated in sporadic breast tumors \textit{MYC}, \textit{INT2}, \textit{EMSI}, \textit{CCND1} and \textit{ERBB2} (Kenemans et al., 2004). Genes involved in hereditary breast cancer differ from those involved in sporadic breast cancer. Two major breast cancer susceptibility genes have been identified so far, \textit{BRCA1} (Miki et al., 1994) and \textit{BRCA2} (Wooster et al., 1994). Both genes are defined as high penetrant genes with an estimated risk for female carriers of developing breast cancer to be 65% for \textit{BRCA1} mutation.
carriers (before age 70) and 45% for \textit{BRCA2} mutation carriers (Antoniou et al., 2003). Both \textit{BRCA1} and \textit{BRCA2} also confer an increased risk for women to develop ovarian cancer, 39% and 11% respectively (Antoniou, et al. 2003).

The Cancer Genetic Counseling Clinic functions within the health care at Sahlgrenska University Hospital in Göteborg since 1995. Mutation analysis (\textit{BRCA1} and \textit{BRCA2}) is offered to patients with a pedigree that indicates dominant inheritance of breast and/or ovarian cancer syndrome. Until now more than 250 families have been analyzed for \textit{BRCA1/2}-mutations after genetic counseling. About one third of the analyses results in an identified mutation and a confirmed \textit{BRCA1/2}-diagnosis. For more than a decade, extensive studies have been performed in order to localize and identify other high penetrant genes predisposing for breast/ovarian cancer. Regrettably, these studies have not yet been able to present confirmed results that could lead to the identification of further high risk associated genes. However, a number of low penetrant genes have recently been reported such as \textit{CHEK2}, \textit{TGFBI}, \textit{FGFR2} and \textit{XRCC1} (Duell et al., 2001; de Jong et al., 2002; Vahteristo et al., 2002; Easton et al., 2007; Hunter et al., 2007). These genes and other yet unknown genes are thought to interact in a polygenic manner and thereby cause a susceptibility to breast cancer. The estimated risk of such low penetrant genes vary between study populations and the presence, or absence, of yet unknown genetic factors are most likely contributing to observed effects of identified low penetrant genes on breast cancer susceptibility.

Thousands of families worldwide suffer from hereditary breast cancer without a detectable mutation in any of the known genes. A clear dominant inheritance mode with multiple cases of breast cancer in separate branches of the family pedigree indicates a monogenic inheritance of a high risk gene rather than polygenic inheritance through multiple low risk alleles. The localization and identification of further breast cancer susceptibility genes is of great importance to concerned women in non-\textit{BRCA1/2}-families, but identification of such genes will also improve the comprehension of cancer development in a more general perspective.

**Aims of the project**

We aim to localize, identify and characterize genes that are involved in the development of breast cancer. Earlier studies in a patient-based family material have generated important findings and the foundation has been laid for clinically important results to be generated in the near future. Within the research project we attempt to reduce candidate chromosomal regions to improve the selection of genes that are suitable for mutation analysis, with the final outcome of isolation of the disease causing gene. For the individual woman in a family with confirmed hereditary breast cancer syndrome, it is of great value that the causative gene can be determined, and thereby give her a chance to receive pre-symptomatic diagnostics and accurate genetic counseling. Knowledge on whether or not she has inherited a faulty gene will aid her in making decisions on preventive/prophylactic measures. An increased knowledge of which genes that confer a susceptibility for cancer brings at the same time a better understanding of the molecular
genetic background of the development of a tumor. Knowledge that is expected to contribute to enhanced, directed therapies and preventive measures for women affected with, or at risk, of developing breast cancer.

Background

Characterization of the patient material

Among families seeking genetic counseling at the Cancer Genetic Clinic we observe the highest frequency of BRCA1/2-mutationer (75%, 18/24) in the group of families that has cases of both ovarian and breast cancer in the same woman (Bergman et al., 2005). A significantly smaller proportion of BRCA1/2 carriers are found within the group of families that are only affected with breast cancer diagnoses. In the west Swedish population that we have studied, a mutation in either BRCA1 or BRCA2 is observed in only 12% (10/81) in this group (Bergman et al., 2005). The majority of "breast cancer only" families seem to have inherited the syndrome through alterations in yet unknown genes.

In a previous report we established the mutation spectrum of the BRCA1- and BRCA2-genes in the west Swedish population. In the study we also performed an analysis of which clinical factors that are associated with the greatest chance of identifying a BRCA1/2-mutation. Little over 140 families with familiar breast/ovarian cancer syndrome were screened for mutations in BRCA1 and BRCA2 and as much as 39% of the tested women were positive for either BRCA1 or BRCA2 mutations. Two thirds of the BRCA1/2 families all carried the exact same mutation, the so called west Swedish founder mutation, BRCA1c.3171ins5. Studies of this mutation has shown that the mutation once occurred in a common ancestor and that it has been inherited through roughly 50 generations which has made it the most common mutation in Sweden today and most dominant in Western Sweden (Bergman et al., 2001). The estimated breast and ovarian cancer risk for carriers of the BRCA1c.3171ins5 does not deviate from the estimated risk for BRCA1/2 carriers in general (Einbeigi et al., 2001). Survival estimates 5 and 10 years after cancer diagnosis in carriers are in line with survival estimates for breast and ovarian cancer overall (Einbeigi, et al. 2001).

The highest probability of finding a BRCA1- or BRCA2-mutation is observed in families with cases of ovarian cancer (68%). If both diagnoses are seen in one woman the mutation detection frequency is even higher (75%). In a population study we tested women with both breast and ovarian cancer diagnoses without any regards to previous family history of such cancer. We tested 256 women for six frequently occurring BRCA1/2 mutations, see table 1(Einbeigi et al., 2006). The analyses showed that as many as 19% of the women were in fact mutation carriers which indicate that breast and ovarian cancer in the same woman is a good marker for presence of an inherited BRCA1/2-mutation regardless of family history of the disease.
Despite the relatively high mutation detection frequency that is achieved in western Sweden we can confirm that more than half of the affected families owe their predisposition to breast cancer to other yet unknown gene(s).

**Linkage analysis**

In our most recent publication we present results from a family based linkage analysis where we identified chromosomal regions linked to inheritance of breast cancer syndrome (Bergman et al., 2007). Microarray-technology was used to genotype 10,000 markers evenly spread over the entire genome. Affected women and healthy relatives from 14 Swedish non-\textit{BRCA1/2}-families contributed to the study. Linkage data indicated several interesting regions as possible locations for novel cancer susceptibility genes. The strongest indication for linkage was observed at chromosome 10 with a LOD score of 2.34. Other regions with high scores were identified at chromosome 12 and 19. The data was analyzed both in total but also family wise. Genotype data from family #14 (Fig.2) alone indicated linkage regions at chromosome 10, 17 and 19. A haplotype analysis of members in this family confirmed that all affected members carried identical markers in the regions, i.e. exactly the same haplotype. Several cancer-related genes are located in the defined linkage regions.

Hereditary breast cancer syndrome is a genetically heterogeneous disease which means that two phenotypically alike syndromes may be caused by alterations in different genes probably located far apart in the genome. Linkage analysis in a heterogeneous material (multiple unrelated families) may bring the unfortunate of potentially high LOD scores at a certain locus are being “masked” as some families are not linking to the region which always results in negative values. To avoid, or at least reduce, the factor of heterogeneity among families the data can be evaluated family wise. It is less likely that women within a family owe their disease to different genes. Normally one family does not bring enough genetic information to reach any LOD scores of particular interest, but if a family with multiple cases separated by sufficient meioses it may be more genetically informative than tens or even hundreds of families with genotypes from only a few cases.

The pedigree of family#14 in Figure 2 shows a severely affected family with as many as 16 cases of breast cancer, no ovarian cancer. The family is on its own genetically highly
informative and did in deed generate high LOD scores in a few regions, as mentioned earlier. Since the linkage analysis was performed we have received knowledge of novel cases of breast cancer in this family and we intend to add genotype data from these women in order to strengthen the evidence of linkage.

![Figure 2. Pedigree of family #14. Dotted circles represent breast cancer cases. Individuals marked with outer rings are new individuals that will be added to linkage and CGH (comparative genome hybridization)-analyses.](image)

**Candidate gene analysis**

Proteins that are expressed by the genes *SAFB1* (19p13) and *SAFB2* (19p13) are like Brca1/2 proteins large multi-functional proteins. An important function for both Safb1/2 and Brca1/2 is that they act as repressors of the estrogen receptor, which in turn plays a central role in the regulation of cell growth and gene expression. Studies have shown that the *SAFB1/2*-genes are down regulated in a significant proportion of breast tumors (Oesterreich 2003; Townson et al., 2003) which imply a possible involvement in cancer development. In order to investigate the potential association of these genes to inherited cancer susceptibility in our material, we performed a mutation screening of the *SAFB1* and *SAFB2* genes in germline DNA from 31 affected women. Three novel polymorphisms in coding sequence were identified, but as the changes do not lead to amino acid changes they are not likely to be disease causing mutations. The analyses have so far been restricted to analysis of coding sequence by DNA-sequencing. Further analyses are needed to exclude any involvement of the genes, such as analysis of whole gene or exon deletions or the presence of epigenetic alterations.

In the identified candidate regions there are a number of low penetrant genes located. Variants (alleles) of these genes have been shown to be associated with a minor risk increase regarding breast cancer development. To investigate whether these alleles have been inherited in affected individuals and thereby caused the linkage to the chromosomal regions, we analyzed the concerned alleles. Breast cancer associated low risk alleles in *XRCC1* (19q), *ERCC2* (19q), *TGFB1* (19q) and *CYP17* (10q) were genotyped and the
nature of the risk allele distribution in the investigated family contradicts the possibility of the observed linkage being caused by inherited low risk alleles (Table 2).

Tabell 2. Presence of low risk alleles in affected and healthy individuals in a large breast cancer family. Red fields indicate risk alleles in monozygous form, orange indicates homozygous risk allele. The two healthy women (green) were both over 80 years old.

<table>
<thead>
<tr>
<th>Gen</th>
<th>rs-nummer</th>
<th>aa-change</th>
<th>riskallel/frekvens i kaukaspop</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>XRCC1</td>
<td>rs1799782</td>
<td>arg194trp</td>
<td>T /9.3%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs25487</td>
<td>arg399gin</td>
<td>A/46.8%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs743572</td>
<td>A1-T</td>
<td>C/23.3%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs1962073</td>
<td>A2-C</td>
<td>C/54.9%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs1799793</td>
<td>leu11pro</td>
<td>GG/34.5%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs13181</td>
<td>asp312asn</td>
<td>C/27.4%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>i ref angiven</td>
<td></td>
<td>C/A</td>
<td>frisk</td>
</tr>
<tr>
<td></td>
<td>rs1052559</td>
<td>lyr751gln</td>
<td></td>
<td>frisk</td>
</tr>
<tr>
<td>CYP17</td>
<td>rs2306669</td>
<td>R194W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/G</td>
<td>rs2075297</td>
<td>R399Q</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/C</td>
<td>rs1127346</td>
<td>5`UTR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs3775298</td>
<td>L10P</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs445620</td>
<td>C/T</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs1801273</td>
<td>C/T</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs2241761</td>
<td>G/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs921586</td>
<td>A/C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGFB1</td>
<td>rs1800469</td>
<td>arg194trp</td>
<td>T /9.3%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs1800470</td>
<td>arg399gin</td>
<td>A/46.8%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs1800471</td>
<td>A1-T</td>
<td>C/23.3%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs1800472</td>
<td>A2-C</td>
<td>C/54.9%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs1800473</td>
<td>leu11pro</td>
<td>GG/34.5%</td>
<td>sjuk</td>
</tr>
<tr>
<td></td>
<td>rs1800474</td>
<td>asp312asn</td>
<td>C/27.4%</td>
<td>sjuk</td>
</tr>
<tr>
<td>ERCC2</td>
<td>rs1799793</td>
<td>leu11pro</td>
<td></td>
<td>frisk</td>
</tr>
<tr>
<td></td>
<td>rs13181</td>
<td>asp312asn</td>
<td></td>
<td>frisk</td>
</tr>
<tr>
<td></td>
<td>i ref angiven</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs1052559</td>
<td>lyr751gln</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Research Outline

**Linkage analysis**

When DNA can be collected from suitable individuals in affected families it is possible to scan the genome (normal DNA) for aberrations in affected family members compared to the DNA of healthy individuals. Two loci (disease gene and marker) are linked when they are inherited together more often than could be attributed to pure chance as a result of their adjacent locations on the chromosome. The linkage analysis will be extended by adding genetic information from newly collected DNA samples from relatives of family #14 as well as samples from whole new families. As the new cases of family #14 are situated far away (counted in meioses) in the family pedigree, they are likely to bring highly useful genetic information to the linkage calculations. With the new genotype data we are aiming to yield sufficient information for significant linkage, which would be the ultimate proof of a novel susceptibility gene.

**Tumor profiling**

Breast tumors from affected women in family #14 (Fig.2) will be analyzed by Comparative Genome Hybridization (CGH). CGH is a method that compares tumor DNA to normal DNA in order to identify chromosomal regions of deletions or amplifications in tumor DNA. Traditional LOH (Loss of Heterozygosity) detects differences between normal and tumor DNA in the same individual but it is not quantitative. CGH on the other hand allows for the detection of alterations in copy number of defined chromosomal regions. Instead the CGH method has limitations in not being capable of determining parental origin as comparison of tumor DNA is made to a reference DNA. The deleted regions may indicate a gene involved in cancer development. Regions localized with CGH are sometimes numerous and most alterations found in tumor DNA are not inherited but are acquired at different stages through the tumor progression. However, we
intend to analyze tumors from several women from the same family which are likely to have a common genetic origin and there is a good chance that tumor DNA in these related women share a common genetic profile that will indicate the location of a susceptibility gene. Tumor profiling will be performed in collaboration with prof. Åke Borg at Lund University who has a lot of experience from studying more than 500 breast tumours, including \textit{BRCA1}, \textit{BRCA2} and other familiar or non-familiar tumors.

\textbf{Candidate genes}

Further analyses will be performed to investigate the potential involvement of \textit{SAFB1/2}-genes in cancer development. The two genes will be targeted for a mutation screening for large genomic rearrangements, such as amplification or deletion of entire exons or the whole genet. Large alterations like these are not detectable by conventional DNA sequencing. Novel linkage data together with CGH/LOH results from analyses of the familial tumors are likely to reduce the previously identified candidate regions. A narrowing down of these regions will be useful when selecting candidate genes for mutation screening. The genes will be chosen on functional characteristics together with chromosomal localization and subjected to germline mutation analysis. In follow-up studies, selected genes will be targeted for functional analysis in \textit{in vitro} and \textit{in vivo} tests (transfection into normal and tumor cells and/or cell lines, transgenesis, germline deletion and more). Functional analyses will be performed in collaboration with Dr Afrouz Behboudi at Göteborg University, highly experienced in mammary tumor biology in the rat system.

\textbf{Material and methods}

\textit{Patients}

Fresh-frozen and/or archived paraffin-embedded tumor samples have been collected with the written consent from 16 affected women in family #14. Peripheral EDTA-blood has been sampled from five affected women who have recently signed up to participate in the study. Blood samples were also collected from close relatives to the affected women. DNA and RNA will be extracted from blood and tumor samples for use in genetic analyses. Blood samples from 16 individuals in three novel non-\textit{BRCA1/2}-families are collected and DNA will be extracted for linkage analysis. Informative non-\textit{BRCA1/2}-families are continuously being collected at the Cancer Genetic Counseling Clinic and will be added to the study.

\textit{Genetic Methodology}

Genotyping of DNA samples will be performed by SNP- (single nucleotide polymorphism) analysis using the Affymetrix system with oligonucleotide-based microarrays. Microarray technique makes it feasible to genotype thousands of SNPs on one array and a genome scan by SNP analysis is potentially more informative than traditional microsatellite analysis. The microarrays are covered with short stretches of nucleotides (oligo-probes) that are complementary in DNA sequence to 10,000 – 500,000 SNP markers spread over the entire human genome. DNA samples will be hybridized to the arrays and subsequent signal detection will be interpreted to determine the genotypes
in all SNP markers for the analyzed samples. A scanner for microarray analysis is available for use at the department of Clinical Genetics. The data will be processed and used in linkage calculations in collaboration with statistician Dr Staffan Nilsson at Chalmers. Extracted tumor DNA will be analyzed by comparative genomic hybridization techniques. High resolution analyses will be performed on the Affymetrix system and confirmative analyses will be performed on spotted BAC- (Bacterial Artificial Chromosome) arrays in collaboration with the microarray unit (former Swegene) at Lund University. In the candidate gene analyses, DNA sequencing will be used as the primary mutation analysis method. For subsequent analyses of large chromosomal rearrangement the MLPA- (multiplex ligation-dependent probe amplification) technique will be used. Genetic Analyzer instruments (Applied Biosystems) are available at the department of Clinical Genetics and at the Genomics core facility at Göteborg University.

**Ethical considerations**

All patients involved in the study have received written information about the project and after consideration; the patients themselves have contacted the research group and received further information at a personal meeting. The patients have been informed that in order to receive information on what ever findings will be made they need to contact the group again. In this way we do not risk to give unwanted knowledge to the families in the study. The patients will receive general information on how the project is proceeding upon request. The research project has been reviewed and approved by an ethical committee, see copy of approval.

**Scientific publication and educational intentions**

Results generated from the study will be authored for publishing in international scientific journals. Members of the group will continuously report any progress on national and international meetings.

**References**


Einbeigi Z, Bergman A, Meis-Kindblom JM and others. 2006. Occurrence of both breast and ovarian cancer in a woman is a marker for the BRCA gene mutations: a population-based study from Western Sweden. Fam Cancer.


