From the Department of Molecular Medicine & Surgery
Karolinska Institutet, Stockholm, Sweden

Investigation of the genetic basis of familial non-BRCA1/2 breast cancer

Paula Maguire

Karolinska Institutet

Stockholm 2005

Published and printed by Repro Print AB
Box 21085, SE-100 31 Stockholm, Sweden

ISBN 91-7140-602-6

To my family

Níl aon tinteán mar do thinteán féin

ABSTRACT

Breast cancer is the most common female malignancy in the Western world and approximately 5-10% of all breast cancer cases present with some degree of family history. In the mid-nineties genetic linkage analyses successfully identified two breast cancer predisposing genes, *BRCA1* and *BRCA2*. Mutations in these genes are responsible for the majority of large early onset breast and breast-ovarian cancer families. However, a large proportion of breast cancer families do not have mutations in *BRCA1* or *BRCA2* and therefore the genetic component in these non-BRCA1/2 families remains to be elucidated.

HIN-1 is a putative tumour suppressor gene on chromosome 5q, which is silenced by methylation in the majority of sporadic breast cancers. Ten families exhibiting suggestive linkage to the region on 5q were investigated for germline mutations in the *HIN-1* gene. No sequence alterations were identified in the ten families, or in DNA from 15 BRCA1 tumours and 35 sporadic tumours. In contrast to sporadic tumours, the *HIN-1* promoter was completely unmethylated in BRCA1 tumours. HIN-1 is therefore unlikely to play a major role in breast cancer predisposition, however its altered expression may have consequences for breast cancer pathogenesis. (Paper I)

Comparative genomic hybridization of non-BRCA1/2 breast carcinomas revealed that loss of chromosome 17 and chromosome 6q were frequent events in high-risk families while gain of 8q was a frequent event in low-risk families. Loss of genetic material from chromosome 17 suggested the presence of a tumour suppressor gene. Investigation of ten genes within a candidate locus on chromosome 17q11.2-12 revealed no obvious pathogenic sequence alterations. However, the frequent observation of genetic alterations involving chromosome 17 in breast tumours suggests the presence of novel genes, which may be involved in breast carcinogenesis. (Paper II)

Re-evaluation of genetic linkage data based on global gene expression profliling of non-BRCA1/2 breast tumours identified chromosome 6 as a candidate locus for two non-BRCA1/2 families. Breast cancer in families 6006 and 6043 was linked to a 43.8 Mb region on chromosome 6, with suggestive LOD scores of 1.48 and 0.78 in the region for families 6006 and 6043 respectively. Detailed fine mapping revealed that these families shared a common four-marker haplotype 2-7-5-2 over a 2.8 Mb region on chromosome 6q14.1. The six genes within this region were investigated for the presence of a possible founder mutation in these two families. A number of shared sequence variants were identified in the two families, none of them were obviously pathogenic. (Paper III)

Three polymorphisms in the *estrogen receptor beta* (*ESR2*) were investigated for their association to breast cancer. A total of 723 breast cancer cases were genotyped, 323 sporadic cases and 400 familial cases. No statistically significant differences in genotype distributions were observed for any of the three polymorphisms individually. However haplotype analysis revealed an association between one common haplotype G-A-G and increased risk of sporadic breast cancer (OR=3.0 p=0.03). This result suggests a role for *ESR2* in breast cancer. (Paper IV)

Keywords: familial non-BRCA1/2 breast cancer, linkage analysis, CGH, association study

LIST OF PUBLICATIONS

This thesis is based on the following papers, which will be referred to in the text by their Roman Numerals:

I. Ian Krop, **Paula Maguire**, Jaana Lahti-Domenici, Gabriela Lodeiro, Andrea Richardson, Hrefna Kristin Johannsdottir, Heli Nevanlinna, Ake Borg, Rebecca Gelman, Rosa Bjork Barkardottir, Annika Lindblom, Kornelia Polyak. Lack of *HIN-1* methylation in BRCA1-linked and "BRCA1-like" breast tumours.
Cancer Research. 2003, 63, 2024-2027

II. **Paula Maguire**, Kerstin Holmberg, Maria Kost-Alimova, Stefan Imreh, Lambert Skoog and Annika Lindblom. CGH analysis of familial non-BRCA1/2 breast tumours and mutation screening of a candidate locus on chromosome 17q11.2-12.
International Journal of Molecular Medicine. 2005, 16(1): 135-141

III. **Paula Maguire**, Barbara Kremeyer, Annika Lindblom. A common haplotype on chromosome 6q14 shared by two non-BRCA1/2 breast cancer families; analysis of a 2.8Mb candidate locus.
Manuscript

IV. **Paula Maguire***, Sarah Margolin*, Johanna Skoglund, Xiao-Feng Sun, Jan-Åke Gustafsson, Anne-Lise Børresen-Dale, Annika Lindblom. *Estrogen receptor beta (ESR2)* polymorphisms in familial and sporadic breast cancer.
Breast Cancer Research and Treatment. 2005, 94(2): 145-52.

OTHER RELATED PAPERS

Kristiansen M, Knudsen GP, **Maguire P**, Margolin S, Pedersen J, Lindblom A and Orstavik KH. High incidence of skewed X chromosome inactivation in young patients with familial non-BRCA1/BRCA2 breast cancer.
Journal of Medical Genetics. 2005, 42(11): 877-80

Kremeyer B, Soller M, Lagerstedt K, **Maguire P**, Mazoyer S, Nordling M, Wahlstrom J and Lindblom A. The *BRCA1* exon 13 duplication in the Swedish population.
Familial Cancer. 2005; 4(2): 191-4

Skoglund J, Margolin S, Zhou XL, **Maguire P**, Werelius B, and Lindblom A. The *estrogen receptor alpha* C975G variant in familial and sporadic breast cancer: a case-control study.
Submitted

CONTENTS

LIST OF ABBREVIATIONS

3'UTR	3' untranslated region
5'UTR	5' untranslated region
ACN	Acetonitrile
AF1	Activating function 1
AF2	Activating function 2
AT	Ataxia telangiectasia
ATM	Ataxia Telangiectasia mutated gene
Bp	Base pair
BRCA1	Breast cancer susceptibility gene 1
BRCA2	Breast cancer susceptibility gene 2
CGH	Comparative genomic hybridisation
DBD	DNA binding domain
ddNTP	Dideoxynucleoside triphosophate
DHPLC	Denaturing high performance liquid chromatography
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DSB	Double stranded breaks
ERE	Estrogen response element
ESR1	Estrogen receptor alpha
ESR2	Estrogen receptor beta
HR	Homologous recombination
HWE	Hardy Weinberg equilibrium
IBD	Identical by descent
Kb	Kilo base
LBD	Ligand binding domain
LD	Linkage disequilibrium
LOD	Logarithm of the odds
LOH	Loss of heterozygosity
Mb	Mega base
NHEJ	Non-homologous end joining
NPL	Non-parametric linkage
OCCR	Ovarian cancer cluster region
OR	Odds ratio
P53	Tumour protein 53 gene
PCR	Polymerase chain reaction
PPi	Inorganic pyrophosphate
PTEN	Phosphatase and tensin homologue deleted on chromosome 10
Rb1	Retinoblastoma gene
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RT-PCR	Reverse-transcriptase PCR
SNP	Single nucleotide polymorphism
STK11	Serine threonine kinase 11 gene
TEAA	Triethylamine acetate

1 INTRODUCTION

Despite major advances in our understanding of human diseases, the mechanisms underlying many common diseases, such as heart disease and cancer, remain elusive. Common diseases are often multifactorial in nature involving the complex interaction of genetic and environmental factors. Cancer is often described as a genetic disease as the transition from a normal cell to a cancerous cell involves the acquisition of a number of genetic alterations. The alterations that confer a growth advantage to the cell are selected for and a form of somatic evolution occurs ultimately resulting in malignant transformation. The genetic alterations in the cell may be acquired somatically or be present in the germline. Most cancer types arise through the acquisition of numerous somatic mutations in a particular tissue, however in some hereditary forms of cancer, a predisposing mutation is present in the germline. As such hereditary cancer syndromes often result in multiple cancer types and generally have an earlier age of onset than their sporadic counterparts. The progression from normal cell to cancer cell is thought to occur through a multi-step model whereby sequential mutations in critical cancer genes give rise to the cancer phenotype. To date, colon cancer represents the most well elucidated cancer model whereby stepwise mutations in *APC*, *KRAS*, *SMAD4* and *p53* are associated with a defined series of stages from normal colonic mucosa to colorectal carcinoma (Fearon and Vogestein, 1990; Vogelstein and Kinzler, 1993). However, this cancer progression model has not been as well clarified for other cancer types such as breast cancer. In an effort to define cancer progression in a logical manner Hanahan and Weinberg have identified six features necessary for a cell to develop a cancer phenotype: self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis (Hanahan and Weinberg, 2000). The aim of cancer biologists is to identify the genetic changes that accompany each of these hallmarks of cancer and to determine which are shared between different cancer types and which are unique to specific tissues. The focus of this thesis is to understand the genetic basis of non-BRCA1/2 familial breast cancer predisposition and so we first need to examine the types of genes that are altered in carcinogenesis.

1.1 ONCOGENES

Oncogenes were initially discovered as retrovirally transmitted tumourigenic agents (Huebner and Todaro, 1969). The cellular counterparts of viral-oncogenes are proto-oncogenes and they are generally considered to be positive regulators of cell growth and differentiation. The first proto-oncogene discovered was *c-Src* (Takeya and Hanafusa, 1983) and to date several hundred proto-oncogenes have been identified. These genes can be classified into five broad functional categories: (1) growth factors, (2) growth factor receptors, (3) signal transducers, (4) transcription factors and (5) apoptosis regulators. Mutations in oncogenes are gain-of-function mutations, whereby the protein product of the gene is highly expressed or constitutively active generally leading to un-regulated cell proliferation. There are three general mechanisms by which oncogenes are activated and lead to cancer; firstly by point mutation as is the case in the *RAS* gene family (Capon, *et al.*, 1983; Sukumar, *et al.*, 1983; Yuasa, *et al.*, 1983), second by gene amplification, an example of which is *ERBB2* amplification in breast cancer (Berger, *et al.*, 1988) and finally chromosomal rearrangements, which occur frequently in leukaemias and lymphomas e.g. the BCR-ABL translocation in chronic myelogenous leukaemia (de Klein, *et al.*, 1982) and the cMyc/IgG translocation in Burkitts lymphoma (Taub, *et al.*, 1982). All three mechanisms of oncogene activation lead to abnormal activity of the normal protein. The combination of activating mutations in oncogenes and loss-of-function mutations in their counterparts, tumour suppressor genes, leads to uncontrolled cell proliferation and subsequently cancer development.

1.2 TUMOUR SUPPRESSOR GENES

The normal physiological role of tumour suppressor genes is to suppress cell proliferation through regulation of the cell cycle, transcriptional regulation and apoptosis. The first tumour suppressor gene identified was the retinoblastoma gene, *Rb1* (Friend, *et al.*, 1987), which is mutated in the childhood cancer syndrome, Retinoblastoma. Mutations in tumour suppressor genes are loss-of-function mutations and at a genetic level are referred to as recessive when compared to dominant oncogene mutations. The recessive model for tumour suppressor gene mutation was first proposed after Knudsons observations in retinoblastoma (Knudson, 1971). As is the case with many cancers, retinoblastoma occurs in both an inherited and sporadic form. Knudson observed that two-rate limiting steps were needed for the development

of cancer and that in the inherited form of retinoblastoma the first step was already present in the germline. The general mechanism of tumour suppressor gene inactivation is the combination of a discrete mutation (point mutation, small insertion/deletion) on one allele and large chromosomal changes, such as whole chromosome deletion or deletion of a chromosome arm, on the other allele. The inactivation of both of the classical tumour suppressor genes, *Rb1* and *p53*, is accompanied by loss of chromosomal material as evidenced by loss of heterozygosity (LOH) on chromosomes 13 and 17 respectively (Friend*, et al.*, 1987; Baker*, et al.*, 1989). More recently the second allele has been found to be silenced by methylation. Feinberg and Vogelstein were the first to demonstrate that global hypomethylation was associated with the cancer phenotype (Feinberg and Vogelstein, 1983) whereas gene silencing by promoter methylation was first noted in the *Rb1* gene (Sakai*, et al.*, 1991) and subsequently in the *von Hippel Lindau* and *p16* genes (Herman*, et al.*, 1994; Merlo*, et al.*, 1995). Although the breast cancer susceptibility genes, *BRCA1* and *BRCA2*, are generally referred to as tumour suppressor genes, they may be more appropriately classified as DNA repair genes, due to their function in the maintenance of genome integrity.

1.3 DNA REPAIR GENES

Cancer arises through the step-wise accumulation of mutations in key genes. However, the progression from a normal cellular phenotype to malignant transformation is due not only to DNA damage but also from the lack of efficient DNA repair mechanisms. Genes responsible for maintaining genome integrity through the repair of DNA damage are also called ‘caretaker’ genes. The genes involved in repairing subtle mistakes made during replication or induced by exposure to mutagens are members of the mismatch repair pathway (MMR), nucleotide excision repair pathway (NER) and the base excision repair pathway (BER). Another set of genes is involved in maintaining genome integrity during mitotic recombination and chromosomal segregation. These genes are responsible for repairing larger errors in DNA, such as double stranded breaks. Mutations in DNA repair genes drive tumourigenesis as the lack of an efficient repair system leads to increased mutation rates, resulting in mutations in oncogenes and tumour suppressor genes and having the overall effect of promoting cell proliferation.

The key role of these different repair systems is highlighted when genes involved in these pathways are mutated in the germline and give rise to hereditary cancer syndromes. Germline mutations in the MMR genes *hMSH2, hMLH1* and *hMSH6*, predispose to an inherited form of colorectal cancer, Hereditary Non Polyposis Colorectal Cancer (HNPCC) (Peltomaki*, et al.*, 1993; Lindblom*, et al.*, 1993b; Miyaki*, et al.*, 1997). Defects in the NER genes cause Xeroderma Pigmentosa (XPA), which is characterised by a high incidence of skin cancer and extreme sensitivity to ultraviolet radiation (Friedberg, 2001) Recently bi-allelic germline mutations in the *MYH* gene, a member of the BER pathway, have been identified and predispose to colorectal cancer (Al-Tassan*, et al.*, 2002). Defective repair of larger chromosomal aberrations has also been implicated in cancer progression. The recognition and repair of double stranded breaks (DSB) in DNA is mediated by a number of genes, *ATM*, *MRE11*, *NBS1*, *BRCA1* and *BRCA2*, all of which when mutated predispose to cancer and many of which are involved in breast cancer predisposition.

2 BREAST CANCER

Of all cancer types, breast cancer is the most common form of cancer in women in the Western world and overall the second most common cancer type, after lung cancer, in both sexes (Parkin*, et al.*, 2005). With an estimated 1.15 million new cases in 2002 and being the most prevalent cancer worldwide, breast cancer is clearly a serious public health issue, and efforts to understand the etiology of the disease are essential. The highest incidence of breast cancer is seen in developed countries, such as North America, Western Europe and Australia, this high incidence rate may reflect the available screening programmes but also likely reflects environmental exposures. The lowest rates are seen in Asian countries, but these countries are currently showing an increase in breast cancer risk, with China reporting a 3-4% annual incidence in breast cancer risks (Parkin*, et al.*, 2005). Although breast cancer has a strong genetic component and mutations in certain genes (*BRCA1* and *BRCA2*) are known to increase womens risk for developing the disease, genetic factors alone cannot explain the excess of breast cancer risk. The role of environmental factors can be clearly seen by the increase in breast cancer risk associated with migrants from low risk countries, such as Asia, moving to Western countries. Breast cancer is a common and complex disease, which involves a strong interplay between genetic and environmental factors, some aspect of which are discussed below.

2.1 NON-GENETIC RISK FACTORS

2.1.1 Hormonal risk factors

The role of hormones in breast cancer etiology was first suggested in 1896 by the discovery that oopherectomy could cause cancer regression and further supported by epidemiological studies, which indicated that bilateral oopherectomy significantly reduces breast cancer risk, and the earlier the ovaries are removed the greater the risk reduction (Hanstein*, et al.*, 2004; Trichopoulos*, et al.*, 1972). Since then numerous studies have demonstrated that exposure to estrogen is directly associated with the risk for developing breast cancer, with prolonged or increased exposure being associated with increased risk for developing breast cancer whereas reducing exposure is thought to have a protective effect (Hulka and Moorman, 2001; Martin and Weber, 2000). Therefore factors that increase the number of menstrual cycles,

such as early age at menarche, nulliparity and late onset of menopause are associated with an increased likelihood of developing breast cancer. While decreasing the total number of ovulatory cycles can have a protective effect, oopherectomy, moderate levels of exercise and longer periods of lactation can achieve this (Martin and Weber, 2000).

The risk associated with exposure to exogenous hormones in the form of oral contraceptives or hormone replacement therapy is less clear and remains controversial. Results from data pooled from 54 studies indicated that the risk associated with ever having used oral contraceptives was very small (relative risk of 1.07) (Collaborative Group on Hormonal Factors in Breast Cancer, 1996). The use of hormonal replacement therapy during menopause is associated with a modest increase in breast cancer risk, whereas long-term use (> 5 years) is associated with a 30-50% increase in risk (Colditz*, et al.*, 1995; Collaborative Group on Hormonal Factors in Breast Cancer, 1997).

2.1.2 Non-hormonal risk factors

There are a number of well established non-hormonal risk factors associated with breast cancer risk. Age is an established and important risk factor for developing breast cancer as breast cancer risk is known to increase steadily with age and in the United States over 66% of breast cancer cases are diagnosed in women aged 55 or over (Hulka and Moorman, 2001). Ionising radiation is associated with increased breast cancer risk, and young women exposed to ionising radiation as a treatment for childhood Hodgkins disease have a substantially increased risk of developing breast cancer (Bhatia*, et al.*, 1996). In addition, survivors of the atomic bomb blasts in Hiroshima, Japan have a very high incidence of breast cancer, particularly those exposed during adolescence, a period of active breast development (Tokunaga*, et al.*, 1994). Women with a history of benign breast lesions, particularly atypia have a three- to four-fold increased breast cancer risk (Hulka and Moorman, 2001). Postmenopausal women attending mammographic screening who present with a high percentage of density in their breast are at increased risk of developing breast cancer (Hulka and Moorman, 2001).

While obesity has been associated with breast cancer risk, this increase in risk may be hormone related as obesity and central fat distribution are believed to act through endocrine intermediates such as the steroid hormones (Pujol*, et al.*, 1997). Alcohol intake has been consistently related to an increased risk of breast cancer with women taking an occasional drink having a modestly increased risk while those having up to

four or more drinks a day have a substantially increased risk (Garfinkel, *et al.*, 1988; Bowlin, *et al.*, 1997). Recent findings suggest that among women with high plasma folate levels, there is no adverse effect of alcohol (Zhang, *et al.*, 2003). Therefore by increasing plasma folate levels through the use of multivitamins or fortification of the food supply breast cancer risk may be reduced, particularly in women who are consuming alcohol.

There is some evidence to suggest that some dietary components may play a role in breast cancer risk, however this remains a controversial topic. Certain studies have identified a relationship between the consumption of animal fat and breast cancer while an inverse relationship has been seen between vegetable fat and breast cancer (Colditz, 2005). It has also been proposed that soy protein is associated with a decreased breast cancer risk and that this may account for the reduced risk seen in Asian countries, where soy protein is a main constituent of the diet. Certain micronutrients such as Vitamin E have also been associated with reduced breast cancer risk.

2.2 GENETIC RISK FACTORS

Family history of breast cancer remains the single most important risk factor for developing the disease. Breast cancer is approximately twice as common in women with an affected first-degree relative and this risk increases with the number of affected relatives and is greater for women with relatives affected at a young age (Collaborative Group on Hormonal Factors in Breast Cancer, 2001). In addition, a Nordic twin study has estimated that 27% of the variation in breast cancer among monozygotic and dizygotic twins can be explained by heritable causes (Lichtenstein, *et al.*, 2000). The familial aggregation of breast cancer accounts for between 5-10% of all breast cancer cases. To date, there have been two highly penetrant predisposing genes identified *BRCA1* (Miki, *et al.*, 1994) and *BRCA2* (Wooster, *et al.*, 1995; Tavtigian, *et al.*, 1996). Mutations in these two genes are responsible for the majority of breast cancer in large, early-onset breast and breast/ovarian cancer families. However, they account for only a small proportion of the total familial cases and so the genetic factors increasing breast cancer risk in the remaining families are still to be identified.

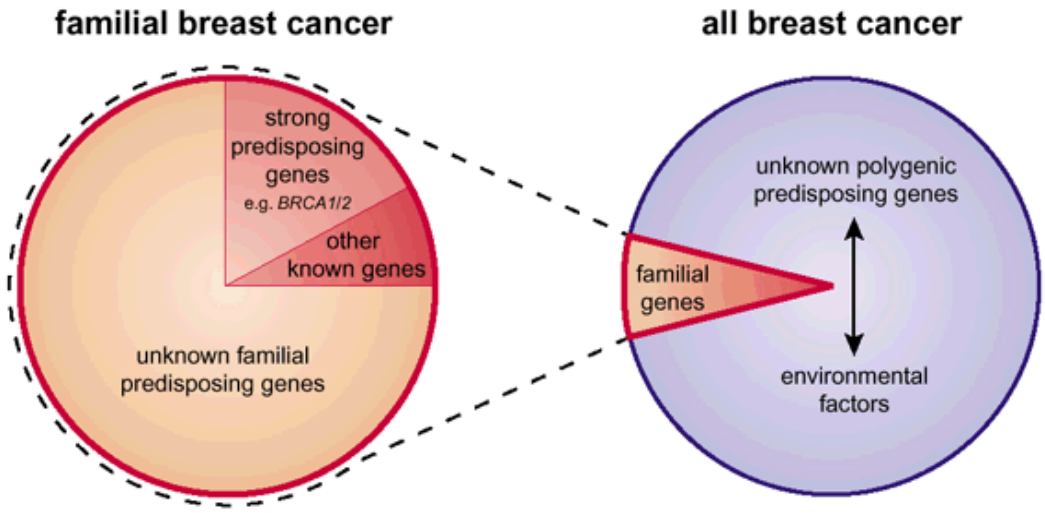

Figure 1. Ninety percent of breast cancers are sporadic, while familial cancer constitutes 5-10% of all breast cancer (right). Of these familial cases mutations in known high-penetrant genes (*BRCA1*, *BRCA2*, *p53*) account for approximately 25% of the families, the remaining familial risk may be due to unknown moderate- to high-penetrant genes (BRCA3) or a number of low penetrant alleles (*CHEK2*).
Modified from Balmain et al., Nat Genet. 2003 (33) Suppl: 238-44

2.3 HIGH PENETRANT BREAST CANCER SUSCEPTIBILITY GENES

2.3.1 *BRCA1* and *BRCA2*

It has been over a decade since the identification of the breast cancer predisposing genes, *BRCA1* and *BRCA2*. *BRCA1* was first localised, by linkage analysis (Hall, *et al.*, 1990), to chromosome 17q21 and three years later the *BRCA1* gene was successfully cloned (Miki, *et al.*, 1994). Linkage analysis of large early-onset breast and breast-ovarian cancer families yielded a logarithm of odds (LOD) score close to 6 for a locus on 17q21. Subsequently cloning of the *BRCA1* gene and identification of pathogenic mutations in breast cancer families established *BRCA1* as the first breast cancer predisposing gene. Less than a year later a second locus was identified on chromosome 13q12-13 in families with breast cancer not linked to 17q21 (Wooster, *et al.*, 1994) and soon thereafter the *BRCA2* gene was cloned (Wooster, *et al.*, 1995; Tavtigian, *et al.*, 1996). The *BRCA1* gene encodes an 1863 amino acid protein, and consists of 22 coding exons distributed over 80kb of genomic DNA. The *BRCA2* gene encodes a very large protein of 3418 amino acids, which is encoded by 27 exons. Although both breast

cancer predisposing genes are large and seem to be widely expressed in most proliferating tissues, they are not homologous to any known protein. Neither are the two proteins homologous to each other, although both BRCA1 and BRCA2 have some features in common.

Essentially all families with an apparent autosomal dominant mode of inheritance to both breast and ovarian cancer are accounted for by mutations in *BRCA1* and *BRCA2*. Mutations in both of these genes are considered highly penetrant and are associated with a lifetime risk of breast cancer of between 60-85% and a lifetime risk of ovarian cancer of between 15-40% (Brose*, et al.*, 2002; Thompson and Easton, 2002b). A recent meta analysis of 22 studies unselected for family history has estimated the cumulative breast cancer risk by age 70 to be 65% in *BRCA1* mutation carriers and an ovarian cancer risk of 39%, while for *BRCA2* mutation carriers the cumulative risk by age 70 was estimated to be 45% for breast cancer and 11% for ovarian cancer (Antoniou*, et al.*, 2003). For *BRCA1* mutation carriers their risk increases with age until 50 years and then decreases, however for *BRCA2* mutation carriers their risk continues to increase with age. There is also evidence that different mutations confer different cancer risks, a study by the Breast Cancer Linkage Consortium (BCLC) has shown that *BRCA2* mutations occurring in the ovarian cancer cluster region (OCCR) are associated with a low risk of breast cancer and a higher risk of ovarian cancer (Thompson and Easton, 2001). There is also evidence for a genotype-phenotype correlation with *BRCA1* mutations, with mutations in the 3' end of the gene being associated with a lower ovarian cancer risk (Gayther*, et al.*, 1995), while mutations in the central region of *BRCA1* are associated with a lower breast cancer risk (Thompson and Easton, 2002a). The reason for this phenotypic variation is currently unclear but may be explained following further elucidation of the functional roles of both BRCA1 and BRCA2.

2.3.2 BRCA3 and novel breast cancer genes

Mutations in *BRCA1* and *BRCA2* do account for the majority of large, early-onset breast and breast-ovarian cancer families, however, these *BRCA1/2* families account for only a small proportion (15-20%) of all breast cancer families. In a study carried out by the Breast Cancer Linkage Consortium, analysis of 237 breast cancer families revealed that 67% of these families were linked to neither *BRCA1* nor *BRCA2* (Ford*, et al.*, 1998). These results and results from other studies demonstrating the low frequency of

BRCA1 and *BRCA2* mutations in certain populations indicate that novel breast cancer genes remain to be identified. Several candidate loci have been proposed to harbour novel breast cancer predisposing genes, including a region on 8p12-22 (Kerangueven*, et al.*, 1995; Seitz*, et al.*, 1997), chromosome 13q21 (Kainu*, et al.*, 2000), and chromosome 2q32.2 (Huusko*, et al.*, 2004). The region on 8p12-22 was analysed following the observation of allele loss in sporadic breast cancer, and subsequent linkage analysis in both French and German breast cancer families yielded modest positive LOD scores. This finding has not been independently replicated in any further studies and the analysis of 31 non-BRCA1/2 breast cancer families failed to find evidence of linkage to this locus (Rahman*, et al.*, 2000). Kainu *et el*. suggested a locus on 13q21, distinct form *BRCA2* and *Rb*, as a putative novel breast cancer locus (Kainu*, et al.*, 2000). The identification of this locus was based on CGH analysis of familial non-BRCA1/2 breast cancers, which identified a specific deletion of 13q21-q22 shared by patients from one family. Subsequent linkage analysis of 77 families indicated a LOD score of 3.46. No predisposing gene has thus far been identified and there has been no further evidence that this locus harbours a common breast cancer predisposing gene (Du*, et al.*, 2002; Thompson*, et al.*, 2002). Finally, although a number of genome wide linkage studies of non-BRCA1/2 breast cancer families have been undertaken, in our own lab and in the UK and Netherlands, to date the only published report is a Finnish study of 14 high-risk breast cancer families. The results of this study suggested a region on 2q32, which gave a LOD score of 1.61 (Huusko*, et al.*, 2004). No predisposing gene has been identified and this study has yet to be replicated.

The search for BRCA3 has lead to some conclusions; firstly, the remaining breast cancer families not attributable to mutations in *BRCA1* or *BRCA2* are genetically heterogeneous, secondly, it is apparent that as no distinct phenotype is available to classify the remaining non-BRCA1/2 families novel approaches, most probably based on molecular profiling, are needed to group families into more genetically homogenous sub-groups and thirdly, the difficulty in identifying novel high-to-moderate penetrant genes may be because they do not exist and the residual familial risk is instead due to low-penetrance alleles, which are of course difficult if not impossible to identify by conventional genetic analysis. Several low-penetrance alleles may act in an additive or multiplicative fashion to increase a womans risk for breast cancer. Candidate low penetrance genes are proto-oncogenes and genes involved in carcinogen metabolism, estrogen metabolism and immunomodulatory pathways.

2.3.3 Rare Cancer Syndromes

2.3.3.1 P53 and Li Fraumeni Syndrome

The *p53* cell cycle checkpoint gene is one of the most commonly mutated genes in human cancers, with mutations estimated to occur in up to 50% of all cancers including breast cancer. Li-Fraumeni syndrome is an autosomal dominant disorder caused by germline mutations in *p53* gene and characterised by an increased risk of soft tissue and osteosarcomas, leukaemias, brain tumours, adenocortical carcinomas, and breast cancers (de Jong, *et al.*, 2002). Germline mutations in *p53* are a rare cause of breast cancer, accounting for less than 1% of breast cancer cases (Patel, *et al.*, 1995; Zelada-Hedman, *et al.*, 1997). However, in approximately 40% of human breast cancers *p53* is somatically mutated and approximately 30-42% of breast cancers exhibit loss of heterozygosity at the *p53* locus on chromosome 17p (Greenblatt, *et al.*, 1994). Although 70% of Li-Fraumeni families have mutations in *p53*, mutations in *CHEK2* are responsible for a proportion of families (Bell, *et al.*, 1999) and recently a third locus on chromosome 1q23 has been suggested in two non-*p53/CHEK2* families (Bachinski, *et al.*, 2005). Interestingly, the family in which the initial *CHEK2* mutation was identified contained multiple cases of early-onset and bilateral breast cancer, indicating at the time that *CHEK2* may be a novel breast cancer predisposing gene.

2.3.3.2 PTEN and Cowdens Syndrome

Mutations in the *PTEN* gene, which encodes a lipid phosphatase on chromosome 10q23.3, predispose to Cowdens syndrome and are present in 80% of Cowden syndrome families (de Jong, *et al.*, 2002). Cowden syndrome is an autosomal dominant disorder characterised by hamartomas and a risk of breast, thyroid and endometrial cancer (Eng, 2003). In women with Cowden syndrome the lifetime risk of developing breast cancer ranges from 25-50% (Eng, 2003). Loss of heterozygosity at the *PTEN* locus has been observed in up to 40% of sporadic breast cancers (de Jong, *et al.*, 2002) and 11% of familial breast cancers (Lindblom, *et al.*, 1993a). However, mutations in the *PTEN* gene are rare both at the somatic level in sporadic breast cancers and in breast cancer families not associated with Cowden syndrome (Chen, *et al.*, 1998; Feilotter, *et al.*, 1999; Freihoff, *et al.*, 1999).

2.3.3.3 STK11 and Peutz-Jegher syndrome

Peutz-Jegher syndrome is a rare autosomally dominant inherited condition characterized by predisposition to benign hamartomatous polyps, in addition to an increased risk of cancer of the breast, gastrointestinal tract, testis and ovaries (Hemminki, *et al.*, 1997). In addition to increased cancer susceptibility, Peutz-Jegher syndrome is characterized by mucocutaneous pigmentation, which usually affects the lips, buccal mucosa and digits. Using a combined approach of comparative genomic hybridization of benign hamartomas and linkage analysis, the predisposing locus was mapped to chromosome 19p. In Peutz-Jegher families, pathogenic mutations were identified in the *STK11* gene, which codes for a serine-threonine kinase (Hemminki, *et al.*, 1998). Mutations in the *STK11* gene appear to be associated with breast cancer risk only in the context of the Peutz-Jegher syndrome. No somatic *STK11* mutations have been described in breast cancer (Bignell, *et al.*, 1998) and no germline mutations have been identified in familial breast cancers (Chen and Lindblom, 2000).

2.3.3.4 ATM and Ataxia Telangiectasia

Ataxia-telangiectasia (AT) is a rare autosomal recessive early childhood disorder, characterized by progressive cerebellar ataxia, skin and ocular telangiectasia, immunodeficiency, chromosomal instability, extreme radiosensitivity and an increased risk of cancer. The *ATM* gene was localised to chromosome 11q22-23 and encodes a large 3056 amino acid protein (Savitsky, *et al.*, 1995). *ATM* is involved in the sensing and repair of DNA damage and subsequently phosphorylates a number of key cell cycle regulators including *BRCA1* and *CHEK2*. Several studies have indicated that women heterozygous for *ATM* mutations have an increased risk of developing breast cancer. However, the role of *ATM* in breast cancer predisposition has been the source of some controversy as the high prevalence of heterozygotes in the population led to the idea that screening mammography, a source of ionising radiation, could increase the penetrance of these mutations (Nathanson, *et al.*, 2001). There has also been speculation that cancer risk may be associated with the mutation type, that is, that truncating and missense mutations would confer different breast cancer risks (de Jong, *et al.*, 2002). A recent study of 1160 relatives of A-T patients from 132 families did confirm *ATM* heterozygotes are at increased risk for developing breast cancer particularly at a young age and that this risk decreases with age and the same study showed no correlation between mutation type and cancer risk (Thompson, *et al.*, 2005).

2.4 LOW PENETRANT GENES

As mentioned earlier the search for novel high penetrant breast cancer predisposing genes has thus far been unsuccessful. One explanation for the lack of a BRCA3 gene is that the remaining familial breast cancer risk is in fact due to polygenic inheritance, whereby susceptibility to breast cancer is conferred by several alleles each of which confer only a moderate risk, but which act together to increase the breast cancer risk (Pharoah*, et al.*, 2002). The difficulty with the polygenic model of inheritance is that the number and type of susceptibility genes is highly variable and therefore a candidate gene approach may be necessary, which in contrast to positional cloning methods relies on some prior knowledge of the candidate gene (i.e. location and function). To date most studies assessing the affect of low-penetrance alleles have focused on genes known to be involved in processes such as DNA repair, estrogen and carcinogen metabolism. The difficulty arising from the current onslaught of association studies is that relatively few investigations report significant findings and if they do these findings are often not confirmed in subsequent studies. Below are some examples of the types of low-penetrance genes thought to play a role in breast cancer predisposition.

2.4.1 CHEK2

The *CHEK2* gene, a recent success in the search for novel breast cancer genes, was identified through traditional linkage analysis, although the identified variant in this gene represents a low-penetrant allele in breast cancer. The search for novel breast cancer predisposing genes lead the Breast Cancer Consortium to carry out linkage analysis on their largest non-BRCA1/2 breast cancer family, identifying a locus on chromosome 22q with a LOD score of 1.2 (Meijers-Heijboer*, et al.*, 2002). The *CHEK2* gene is located on chromosome 22q and encodes the mammalian homolog of Rad52, a checkpoint kinase that is a key regulator of the cellular response to DNA damage. A truncating germline mutation, 1100delC, was identified in 7 members of this family affected with breast cancer and this was the same mutation identified in the original Li-Fraumeni family. This mutation occurs in the kinase domain of the CHEK2 protein and has been found to abolish its function. The *CHEK2**1100delC mutation is associated with an increased breast cancer risk particularly in patients with a family history of the disease (Vahteristo*, et al.*, 2002; CHEK2 Breast Cancer Case Control Consortium, 2004)**.** The frequency of the *CHEK2**1100delC mutation is approximately 1% in controls and 4-5% in breast cancer cases with a family history (Meijers-Heijboer*, et al.*,

2002; Vahteristo, *et al.*, 2002). The identification of this variant in families with relatively few affected members, its incomplete segregation with disease and its presence in the general population point to a role for *CHEK2* as a low penetrance breast cancer gene.

2.4.2 Other low-penetrance alleles

This section outlines some examples from the literature of other low-penetrance alleles associated with breast cancer risk, this is by no means a complete review as the list of low-penetrance alleles studied to date in relation to breast cancer is exhaustive. Mutations in proto-oncogenes lead to unregulated cell cycle and abnormal growth and proliferation and as such variants in proto-oncogenes are likely to play a role in cancer susceptibility. The *HRAS1* proto-oncogene is flanked by a polymorphic minisatellite at the 3' end, and this minisatellite marker is composed of four common alleles and numerous intermediate and rare alleles (de Jong, *et al.*, 2002). Several studies have examined this *HRAS1* minisatellite polymorphism and breast cancer risk and found an association with an increased breast cancer risk with odds ratios (OR) of approximately 2 (with one study reporting an OR of 7) (Garrett, *et al.*, 1993; Krontiris, *et al.*, 1993; Gosse-Brun, *et al.*, 1999). However, a recent study found no increase in breast cancer risk associated with rare *HRAS1* alleles although an increased risk was observed for one rare large allele (Tamimi, *et al.*, 2003).

A number of genes involved in metabolic pathways have been studied in relation to cancer risks, including the *cytochrome P450* family, the *GST* family and the *NAT1* and *NAT2* genes. The cytochrome P450 family are phase I enzymes and in general these enzymes activate carcinogens and therefore an increased enzyme activity may be associated with increased cancer risk. The GST family are phase II enzymes and act to metabolically inactivate carcinogens, a genotype associated with decreased enzyme activity might therefore increase breast cancer risk. In a meta-analysis performed by Dunning *et al*., polymorphisms in the phase II enzymes GSTP1 and GSTM1 were significantly associated with breast cancer risk. The GSTP1 Ile105Val polymorphism was associated with a moderately elevated breast cancer risk (OR=1.6 p=0.02) and the GSTM1 polymorphic gene deletion was associated with post-menopausal breast cancer (OR=1.33 p=0.04) (Dunning, *et al.*, 1999). A polymorphism in the cell cycle checkpoint gene *p53*, Arg72Pro, was also associated with a slightly increased risk of breast cancer (OR=1.27 p=0.03).

As mentioned earlier, both endogenous and exogenous hormone exposure is a risk factor for breast cancer and as such, genes involved in estrogen metabolism and signalling are potential candidates. The *CYP19* gene is a member of the cytochrome P450 family involved in estrogen metabolism, and the *CYP19* $(TTTA)_n$ polymorphism has been associated with an increased breast cancer risk (OR=2.33 p=0.002) (Dunning, *et al.*, 1999). Many additional studies have been carried out on the *CYP17* gene, the COMT family of enzymes, the steroid hormone receptors, *ESR1*, *PR* and *AR*, however the results from many of these studies remain contradictory and their association with breast cancer risk remains to be validated. Recently, a number of studies have indicated that the *ESR2* gene may be associated with breast cancer risk (Paper IV) (Zheng, *et al.*, 2003; Gold, *et al.*, 2004).

3 MOLECULAR ASPECTS OF BREAST CANCER

3.1 FUNCTIONS OF BRCA1 AND BRCA2

Since their discovery over a decade ago, the breast cancer predisposing genes *BRCA1* and *BRCA2* have been the focus of intensive research, however, to date a complete picture of their functional roles has not been attained. Investigations of the physiological role of BRCA1 have lead to the identification of several functions, which may underlie its role in carcinogenesis. These roles include DNA repair, cell-cycle checkpoint control, protein ubiquitylation and chromatin remodelling. In contrast, there is still a limited knowledge of BRCA2 function, besides its involvement in homologous recombination and more recently in cell cytokinesis. Some of the functions of BRCA1 and BRCA2 are outlined in this chapter and it is clear how these functions can play a role in the tumour suppressor/caretaker function associated with these genes, what remains unclear however is how disruption of these fundamental roles in essential cellular process can lead to the tissue specific cancer phenotype associated with mutations in these genes. It is evident that further investigations are needed to determine the tissue specific role of these genes and to identify specific interacting partners, which may hold the link to their role in breast and ovarian carcinogenesis.

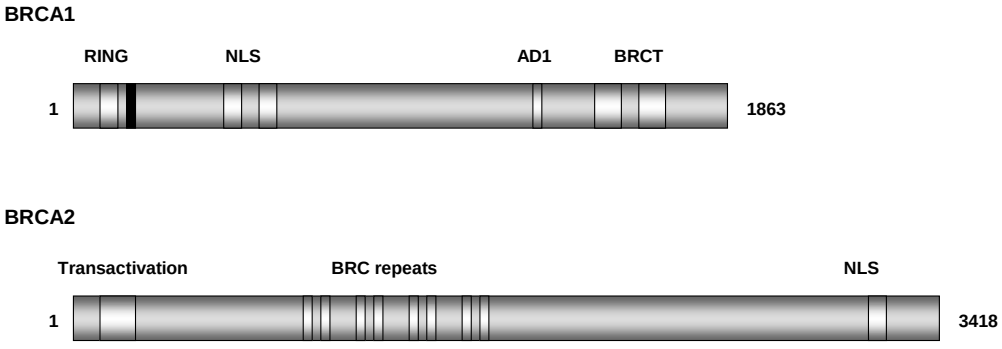

Figure 2. Structural and Functional Aspects of BRCA1 and BRCA2

The BRCA1 N-terminal RING domain, nuclear localisation signal (NLS) and C-terminal BRCT domains are shown. The BRCA2 transactivation domain, eight BRC repeat motifs and the NLS are also indicated.

3.1.1 DNA repair

Both breast cancer predisposing genes are involved in the repair of double stranded breaks (DSB) by homologous recombination. Although, the involvement of BRCA2 appears to be more direct than that of BRCA1, their functional abrogation leads to gross chromosomal abnormalities presumably due to the incorrect repair of DSBs. DSBs can be induced by a number of different mechanisms including ionizing radiation, certain metabolites produced during normal cellular reactions and also during DNA replication. DSBs are also natural intermediates during essential cellular processes such as meiotic recombination or immunoglobulin/TCR receptor maturation. Several mechanisms have evolved to repair DSBs, the two main pathways being homologous recombination (HR) and non-homologous end joining (NHEJ). HR is an accurate repair pathway which uses an undamaged complementary sister chromatid as a template for repair, whereas NHEJ is an error-prone process whereby inaccurate nucleotide substitutions are tolerated at the site of DNA damage. Cells deficient in *Brca1* and *Brca2* exhibit spontaneous chromosome breakage and severe aneuploidy and centrosome amplification, (Patel*, et al.*, 1998; Shen*, et al.*, 1998; Tutt*, et al.*, 1999; Xu*, et al.*, 1999b) and this chromosomal instability appears to be due to a deficiency in homologous recombination (Moynahan*, et al.*, 1999; Tutt*, et al.*, 2001). Examination of *BRCA1* and *BRCA2* deficient tumours also demonstrated evidence of numerous gross chromosomal aberrations (Tirkkonen*, et al.*, 1997).

One of the first indications that BRCA1 was involved in DNA repair came from its association with Rad51, the eukaryotic homolog of RecA. The BRCA1 protein colocalises with Rad51 in nuclear dots during S phase and both are re-localised to sites of repair in response to DNA damage (Scully*, et al.*, 1997). Rad51 is a member of the RAD52 group of proteins, which includes RAD50, RAD51, RAD52, RAD54 and MRE11, which are key players involved in the repair of DSBs by homologous recombination (HR). Rad51 plays a central role in HR, it coats single stranded DNA to form a nucleoprotein filament that invades and pairs with a homologous DNA duplex, initiating strand exchange between the paired DNA molecules (Venkitaraman, 2002). Although it is clear that BRCA1 plays a role along with Rad51 in DSB repair, the exact nature of the BRCA1 function remains to be elucidated and it does not appear to directly regulate Rad51. In contrast, it has been shown BRCA2 can directly bind Rad51 and this interaction is mediated primarily through the ~40 amino acid BRC motifs in BRCA2. BRCA2 appears to regulate the activity of Rad51 *in vitro*, with

Rad51 being sequestered by BRCA2, which suppresses its ability to form nucleoprotein filaments. DNA damage may be the trigger that releases Rad51 from this inactive BRCA2-bound state and re-localizes it to sites of DNA damage (Venkitaraman, 2002).

It appears that BRCA1 acts in a more general and fundamental role at the level of sensing and signalling DNA damage to the cell. One of the earliest responses to DNA damage is the phosphorylation of the histone, H2A-X, which is subsequently localised over a large region spanning the site of DNA damage, suggesting that chromatin remodelling may occur to facilitate access of the repair machinery. BRCA1 co-localises to the sites of H2A-X phosphorylation, consistent with a role in the early events of DNA repair (Paull, *et al.*, 2000).

3.1.2 Checkpoint Control

Although *BRCA1* has been implicated in the early response to DNA damage, it can also be seen to work further downstream in the repair process through its role in the cell cycle checkpoints. *Brca1* deficient cells show a defect in the S-phase and G2/M-phase cell cycle checkpoint controls (Xu, *et al.*, 1999b; Xu, *et al.*, 2001) and this phenotype is associated with sensitivity to radiation. BRCA1 is rapidly phosphorylated after DNA damage in dividing cells by a group of protein kinases ATM, ATR and CHEK2 and phosphorylation of BRCA1 contributes to its function in checkpoint control. ATM and CHEK2 phosphorylate BRCA1 after ionising radiation, whereas ATR is more specifically activated after UV treatment or replication arrest. BRCA1 has also recently been implicated in the decatanation checkpoint control, which monitors the degree of chromatid decatanation before cells pass into mitosis (Lou, *et al.*, 2005). Although a direct role for BRCA2 in control of cell cycle has not been established, recent evidence points to role for BRCA2 in cell cytokinesis (Daniels, *et al.*, 2004). Cells lacking functional BRCA2 take a longer time to progress from mitotic anaphase to the completion of daughter-cell separation.

3.1.3 Protein ubiquitylation

Classically, ubiquitylation is the process by which proteins are tagged for degradation by the proteasome. The ubiquitin pathway covalently modifies target proteins by the addition of a 76-amino acid ubiquitin molecule to lysine residue(s) of the target molecule. Target proteins can either be mono-ubiquinated or polyubiquinated. The BRCA1 protein contains a ring-finger motif at the N-terminal domain, a feature

conserved in many proteins with E3 ubiquitin ligase activity. BRCA1 has been demonstrated to form a heterodimer with BARD1 associating via the RING-finger domains and adjacent α helices and this BRCA1:BARD1 complex has the ability to ligate ubiquitin *in vitro* (Mallery, *et al.*, 2002). However, although this function of BRCA1 has been established the key question remains, what are the targets for this ubiquitylation activity? Another point is that perhaps the ubiquitination of target proteins does not always result in protein degradation and may instead be a signalling mechanism, which is transient and more difficult to detect. Ubiquitin has seven lysine residues that can be potential donors on each ubiquitin monomer. Lysine-48 linked ubiquitin chains are targeted to the proteasome for degradation, lysine-63 linked chains confer non-proteolytic signals that control various pathways, including DNA repair in yeast and regulation of certain protein kinases (Spence, *et al.*, 1995; Wang, *et al.*, 2001). This link between protein ubiquitylation and DNA repair has been strengthened by the recent study by Morris *et al*., which demonstrated that *in vivo* BRCA1 co-localises with conjugated ubiquitin at replication forks and at sites of irradiation induced DNA damage (Morris and Solomon, 2004). The formation of the conjugated ubiquitin by BRCA1 requires lysine-6 of ubiquitin, whether lysine-6 linked ubiquitin chains are the only form of BRCA1 conjugated chains *in vivo* and whether they are targeted for degradation or towards other pathways remains to be determined. Recently, a role for BRCA1 in DNA decatanation has been suggested (Lou, *et al.*, 2005) and the mechanism of this BRCA1 function may occur through BRCA1-dependant TopIIα ubiquitination, and regulate the activity of TopIIα. Defects in BRCA1 function may therefore reduce TopIIα activity and result in a DNA decatanation defect (Ashworth, 2005).

3.1.4 Chromatin remodelling

It is known that chromatin is remodelled at DSB sites, presumably to facilitate the DNA repair process. Interestingly, a role for BRCA1 in chromatin remodelling has been suggested due to its interaction with a number of proteins with known remodelling functions. BRCA1 is a member of BASC, BRCA1-associated genome surveillance complex, this multiprotein complex contains several tumour suppressor genes and genes involved in DNA repair in addition to the DNA remodelling helicase and Blooms syndrome gene, BLM (Wang, *et al.*, 2000). BRCA1 has also been found to interact directly with BRG1 a member of the SWI/SNF chromatin remodelling complex (Bochar, *et al.*, 2000). BRCA1 interacts, through its BRCT repeats, with another DNA

helicase BACH1 (Cantor, *et al.*, 2001). Taken together these data suggest that the role of BRCA1 in DNA repair may be related to its chromatin remodelling ability.

Interestingly, BRCA1 has been linked to a role in establishing heterochromatin as it has been localised to the inactive X (Xi) chromosome (Ganesan, *et al.*, 2002). Both BRCA1 and its heterodimeric partner, BARD1 interact directly or indirectly with the Xi-specific transcript (XIST) RNA. In primary tumours lacking BRCA1, XIST failed to form Xi-associated foci. Whether the loss of BRCA1 and its subsequent effect of Xi plays a role in tumour development or whether the loss of Xi is a general consequence of the loss of the fundamental property of genome integrity remains to be established.

Although, outlined above are a sub-set of BRCA1 and BRCA2 functions and their interacting partners it is clear that the majority of known functions to date involve maintenance of genome integrity, whether directly through DNA repair or indirectly through regulation of different proteins, cell cycle control or BRCA1's chromatin remodelling activity. However, two fundamental questions remain; first, why do mutations in ubiquitously expressed genes involved in essential cellular processes predispose specifically to breast and ovarian cancer and second, does loss of BRCA1 and BRCA2 expression play a role in breast and ovarian cancer progression in general (i.e. sporadic forms). Some potential elements, which may be involved in the answer to these questions, are discussed below.

3.2 ESTROGEN RECEPTORS

As discussed earlier, estrogen exposure is a risk factor for developing breast cancer and this association was first seen as early as 1896 when the British surgeon, Sir George Beatson, noted that oopherectomy could lead to tumour regression (Hanstein, *et al.*, 2004). Furthermore, the estrogen receptor (ER) antagonist tamoxifen, which blocks the actions of estrogens in the breast, prevents primary and recurring breast tumour development (Hilakivi-Clarke, 2000). It is generally assumed that the association of estrogen with increased cancer risk may be due to its role in increasing cell proliferation. This increase in cell proliferation may lead to the accumulation of mutations, which can cause cancer. In addition, the cellular metabolites of estrogen are known carcinogenic compounds and may themselves result in direct DNA damage (Liehr, 2000). These factors, taken together with the role of estrogen in the activation of

certain tumour suppressor genes, such as *BRCA1*, and the direct interaction of the estrogen receptor alpha with BRCA1 may provide a link to explain why mutations in *BRCA1* and *BRCA2* predispose to the estrogen-responsive cancers of the breast and ovary. Estrogen exerts its biological effects through binding of the steroid hormone receptors, estrogen receptor alpha (ESR1) (Greene*, et al.*, 1986) and the more recently discovered estrogen receptor beta (ESR2) (Mosselman*, et al.*, 1996). ESR1 and ESR2 belong to the nuclear receptor superfamily of ligand inducible transcription factors, which regulate transcription in association with coregulators via binding to DNA enhancer elements (EREs) located within the promoters of target genes. To date the majority of studies have focused on the role of ESR1 in breast cancer and its application as a prognostic factor as well a therapeutic target in breast cancer management.

The estrogen receptors contain six functional domains (A-F). Estrogen binds to the ligand binding domain (LBD in domain E) and induces a conformational change which unmasks the activating function 1 domain (AF1 in domain A/B). This conformational change allows dimerisation of the receptor to occur and activation of the activating function 2 (AF2 in domain E). The estrogen receptor can thus bind to the ERE of target DNA via domain C. Figure 3 illustrates a schematic of the structure and function of ESR1 and ESR2. *ESR1* and *ESR2* share about 95% homology in the DNA binding domain (domain C), which is capable of binding to EREs within target genes and display 55% homology in the ligand binding domain, exhibiting similar but not identical ligand binding properties (Cowley and Parker, 1999).

Despite the similarity in structure, the ER's appear to have distinct and non-overlapping functions. The divergence at the N-terminal region results in the AF1 domain of ESR2 having negligible transcriptional activation activity in comparison to that of ESR1 (Cowley and Parker, 1999). Interestingly, BRCA1 has been demonstrated to inhibit ligand-dependant and ligand-independent estrogen activity of ESR1 and this inhibition is due both to the direct interaction of BRCA1 and ESR1 and to the downregulation of ESR1 transcriptional coactivator p300 (Ma*, et al.*, 2005). This BRCA1 dependant regulation of ESR1 activity may play a role in the tissue specific cancer progression associated with BRCA1 mutations. The lack of a functional BRCA1 protein may lead to increased ESR1 activity, thereby causing increased cell proliferation in addition to

the accumulation of genotoxic metabolites that result in DNA damage that is inappropriately repaired in the absence of BRCA1.

ESR2 is also expressed in human breast tissues and in tumour samples suggesting a role for ESR2 in breast cancer (Leygue*, et al.*, 1998; Speirs*, et al.*, 1999). ESR2 is usually co-expressed with ESR1 in breast tumours and studies indicate that ESR2 can heterodimerise with ESR1 (Cowley*, et al.*, 1997). The presence of numerous ESR2 splice variants and their ability to heterodimerise with ESR1 indicates that ESR2 may have an ESR1 regulatory function. The Cx isoform of ESR1 has been shown to act as a potential inhibitor of ESR1 transactivation, possibly through heterodimer formation (Ogawa*, et al.*, 1998). A number of studies have suggested that decreased ESR2 expression is associated with breast cancer and that perhaps the balance between ESR1:ESR2 has a role in breast carcinogenesis (Bardin*, et al.*, 2004). Therefore, the complex interplay between the estrogen receptor signalling pathways and their inherent role in breast cancer remains to be elucidated. The determination of the role of ESR2 in breast cancer development may lead to identification of novel therapeutic targets and enhanced clinical management of breast cancer. In an effort to determine the role of *ESR2* in both sporadic and hereditary breast cancer development, we have examined three single nucleotide polymorphisms (SNPs) in the *ESR2* for their association to breast cancer (Paper IV).

Figure 3. Schematic illustration of the modular design of the estrogen receptors. A-F represent the six functional domains of the estrogen receptors. The A/B domain contains the activating function-1 domain (AF1). Domain C contains the DNA binding domain (DBD) and the C-terminal domain E contains the ligand binding domain (LBD) and activating function-2 domain.

3.3 EPIGENETICS IN BREAST CANCER

Epigenetics is defined as stable alterations in the genome, heritable through cell division, that do not involve changes in the DNA sequence itself. Epigenetic alterations include both alterations at the DNA and protein level, through methylation of CpG dinucleotides (either hypo- or hyper-methylation) or by acetylation or deacetylation of core histone proteins. The involvement of epigenetic changes in human cancers was first noted in 1983 by Feinberg *et al*., where global hypomethylation at CpG dinucleotides was observed in colorectal tumours (Feinberg and Vogelstein, 1983). Later hypermethylation of the retinoblastoma tumour suppressor gene in sporadic cases of the disease was described in 1991. (Sakai, *et al.*, 1991). The consequence of hypomethylation is the activation of genes that are normally inactivated by methylation of CpG islands near or within their promoters. In contrast, hypermethylation of CpG islands has been associated with gene silencing and a number of classical tumour suppressor genes have been identified in which hypermethylation of their respective promoters is associated with disease progression.

The historical classification of *BRCA1* and *BRCA2* as tumour suppressor genes has often been challenged due to the rare occurrence of somatic mutations in the sporadic forms of breast and ovarian cancer (Futreal, *et al.*, 1994; Lancaster, *et al.*, 1996). However, as high levels of LOH are observed in sporadic breast and ovarian carcinomas (Futreal, *et al.*, 1992; Russell, *et al.*, 2000) and BRCA1 transcript and protein levels are decreased in sporadic breast and ovarian cancer (Thompson, *et al.*, 1995; Wilson, *et al.*, 1999) an alternative to direct mutation may be epigenetic alterations of the second allele. Dobrovic *et al*., were the first to observe that the *BRCA1* promoter was hypermethylated in a subset of sporadic breast tumours (Dobrovic and Simpfendorfer, 1997). This finding has since been confirmed by a number of studies and hypermethylation of the *BRCA1* promoter has been associated with decreased BRCA1 expression (Catteau, *et al.*, 1999; Esteller, *et al.*, 2000; Rice, *et al.*, 2000). To date, there are no reports detailing *BRCA2* promoter hypermethylation (Collins, *et al.*, 1997). The *HIN-1* gene is a recently identified putative tumour suppressor gene that is silenced by methylation in the majority of breast carcinomas. In Paper I we investigated the mutational status and the promoter methylation status of the *HIN-1* gene in breast cancer.

4 STRATEGIES FOR CANCER GENE DISCOVERY

Breast cancer genes may be defined as genes in which germline mutations predispose to breast cancer or in the broader sense genes that play a role in the breast cancer pathogenesis. We have already discussed the known breast cancer predisposition genes and their possible functional roles in breast cancer development, however as mentioned before, these known genes account for only a small proportion of familial breast cancer cases. The difficulty to date in identifying novel breast cancer genes most likely results from the absence of remaining common high-penetrance genes. While highly penetrant genes may exist in individual families these genes are difficult to detect by conventional analysis. Recent evidence points to the existence of multiple moderate- to low-penetrance genes involved in breast cancer pathogenesis. The possible interaction of these low-penetrance alleles with environmental factors contributes to overall breast cancer risk and as such labels familial non-BRCA1/2 breast cancer as a complex disease. While family based linkage studies can be useful for the identification of high to moderate-penetrance genes, the identification of low-penetrance genes currently relies on a candidate gene approach. Some strategies for identifying novel cancer genes are discussed in particular relation to familial breast cancer.

4.1 LINKAGE ANALYSIS

Parametric linkage analysis is the analysis of the cosegregation of genetic loci in pedigrees. Two genetic loci are said to be linked if they are transmitted together from parent to offspring more often than would be expected by the rules of independent assortment. The closer two loci are on a chromosome the more likely it is they will be transmitted together, the further apart they are on the chromosome, the more likely it is that a recombination event during meiosis will separate them. The recombination fraction θ refers to the probability of recombination between two loci at meiosis. Two loci are said to be in complete linkage if no recombination between them is observed ($\theta = 0$), there is some degree of linkage if the recombination fraction in less than 50% ($\theta < 0.5$), while two loci are not linked if the recombination fraction is 50% ($= 0.5$). By genotyping polymorphic genetic markers and studying their segregation through pedigrees, it is possible to infer their position relative to each other on the genome. The number and type of genetic markers has varied throughout the years and currently

linkage mapping sets are available which are composed of several hundred (>400) evenly spaced microsatellite markers.

A statistical measure of linkage needs to be applied to determine if there is significant evidence of linkage between two loci and LOD score analysis is a likelihood-based parametric linkage approach for the determination of linkage. The LOD score represents the ratio of two likelihoods, the likelihood of observing linkage ($\theta<0.5$) versus the likelihood of observing no linkage ($\theta=0.5$). Large positive LOD scores are evidence of linkage and negative LOD scores are evidence against linkage. In order to calculate a LOD score a number of parameters must first be defined, including a mode of inheritance for the disease, disease allele frequency, marker allele frequency and a full marker map for each chromosome. Originally a LOD score of 3 was proposed as evidence of significant linkage and LOD scores below –2 were evidence against linkage (Morton, 1955). However a LOD score of 3 has been determined to be equivalent to a genome-wide significance of only 0.09 and therefore higher thresholds of linkage have been suggested with a LOD score of 3.3 being equivalent to a genome wide significance of 0.05 (Lander and Kruglyak, 1995).

There have been a number of success stories in the search for the genes responsible for inherited diseases including familial cancer syndromes. Many of these successes have focused on diseases segregating in large families and typically displaying a single, simple mode of inheritance either dominant, recessive or X-linked. Positional cloning, as this method is generally known as, has lead to the identification of a number of high penetrant genes responsible for several inherited cancers, including breast and ovarian cancer (*BRCA1* and *BRCA2*) (Hall, *et al.*, 1990; Wooster, *et al.*, 1994), adenomatous polyposis colon cancer (*APC*) (Bodmer, *et al.*, 1987), HNPCC (*MSH2*, *MLH1*) (Peltomaki, *et al.*, 1993; Lindblom, *et al.*, 1993b) melanoma (*CDNK2A*) (Cannon-Albright, *et al.*, 1992) and testicular cancer (*TCG1*) (Rapley, *et al.*, 2000).

The past successes of linkage analysis have mainly relied on disease phenotypes exhibiting classical mendelian-like inheritance patterns. The discovery of *BRCA1* and subsequently *BRCA2* relied on the distinct phenotypes exhibited by the families, without the selection for cases with early-onset breast cancer or male breast cancer the demonstration of true linkage would not have been possible. Herein lies the dilemma with the remaining non-BRCA1/2 breast cancer families, to date no identifiable

phenotype has been suggested that accurately classifies these families into genetically homogenous sub-groups. As a result, no novel genes have so far been identified by linkage analysis and any suggestive loci have failed to be replicated. The difficulty in using linkage analysis for the detection of genes causing complex disease, which breast cancer is, arises from a number of sources; firstly the remaining non-BRCA1/2 families are inherently small in size, which reduces the power of linkage. Secondly, inaccurate phenotyping, with the inclusion of phenocopies in the analysis, can obscure the linkage data. Thirdly, genetic heterogeneity, which certainly exists and complex inheritance patterns also detract from the ability of linkage analysis to detect novel disease loci.

Some of the difficulties associated with traditional parametric linkage analysis may be overcome by using the alternative non-parametric linkage (NPL) linkage analysis method. NPL is currently the method of choice for mapping complex diseases where several genes and environmental factors might contribute to disease risk and where there is no clear mode of inheritance. The sib-pair approach studies affected sibling pairs whereby according to the null hypothesis, at any locus the probability that the sibs share no alleles identical by descent (IBD) is 0.25, that they share one allele is 0.5 or that they share two is 0.25. Linkage is suggested when the affected sibs share significantly more alleles than would be expected by chance. In NPL analysis a LOD score of 3.6 corresponds to a genome wide significance level (Lander and Kruglyak, 1995).

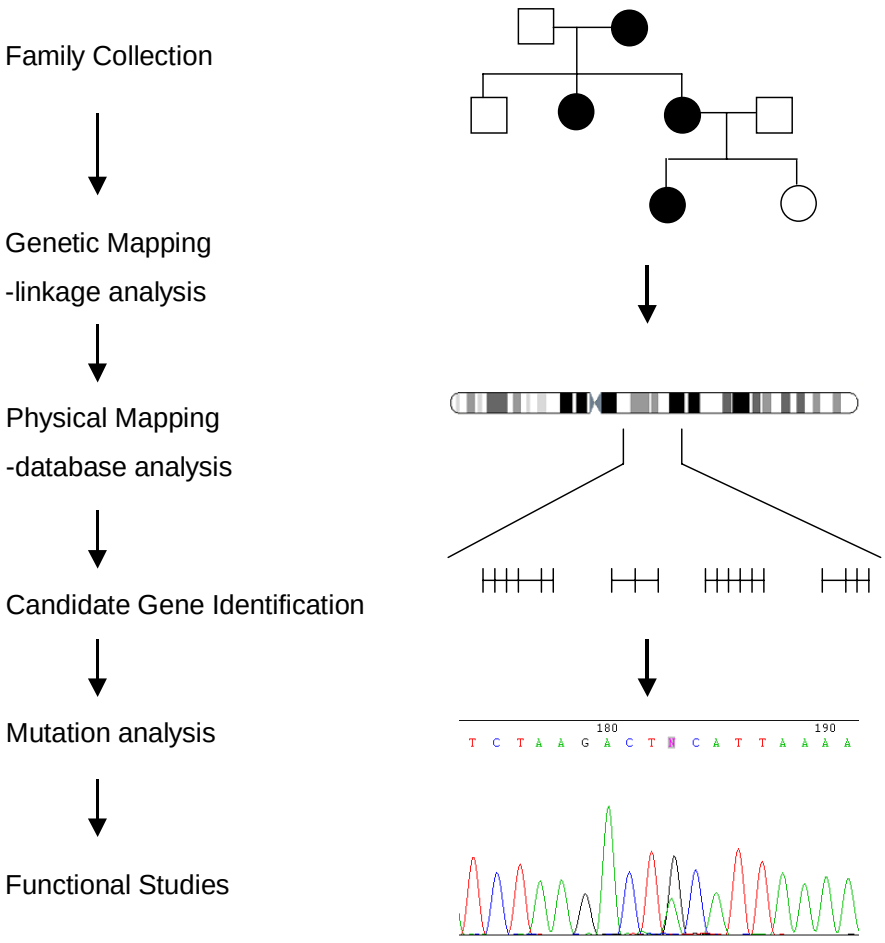

Figure 4. Positional cloning as a method for the localization and identification of novel genes causing genetic disease.

4.2 ASSOCIATION ANALYSIS

In 2002 Pharaoh *et al*., suggested a polygenic susceptibility model to explain breast cancer risk (Pharoah, *et al.*, 2002). This model was proposed based on the study of a population-based series of 1,484 affected individuals. The results from this analysis pointed to two genetic models which best fit the data, first the polygenic model favoured by the authors and second a recessive model, which has been suggested previously (Cui, *et al.*, 2001). Although both models fit the population based series well, the recessive model did not fit the multiple-case families as well as the polygenic model and so the polygenic model was favoured. The polygenic model suggests that susceptibility to breast cancer is conferred by a large number of alleles, each conferring a small genotypic risk, which act additively or multiplicatively to increase a womans risk. It was also suggested that half of all breast cancers occur in the most susceptible 12% of the population (Pharoah, *et al.*, 2002). This model has both an impact on the clinical management and prediction/prevention strategies but also on experimental designs aimed at detecting breast cancer genes.

Due to the low relative risks associated with each low-penetrant allele, identifying these genes is beyond the practical scope of traditional linkage analysis. Therefore the focus has turned to association or linkage disequilibrium (LD) studies. In contrast to family based linkage studies, association studies are designed to detect associations at the population level, which represent shared common ancestral chromosomes. One of the main differences in the strategies of linkage analysis and association studies is that association analyses currently rely on *a priori* knowledge of biologically plausible candidates or loci of interest. Single nucleotide polymorphisms (SNPs) in candidate genes may act as low-penetrant alleles, these SNPs may be coding or non-coding and may have a direct functional effect or may be associated through linkage disequilibrium (LD) with a functional variant. Large-scale SNP discovery projects aim to identify a significant proportion of the SNP variation in the human genome, which will facilitate genome wide association analysis for complex diseases. Although the total number of SNPs estimated in the human genome is close to 10 million, the International HapMap project aims at identifying the extent of LD between SNPs and designate tagging SNPs capable of capturing all the variation within a specific chromosomal loci (Sachidanandam, *et al.*, 2001; Altshuler, *et al.*, 2005). This data should simplify future genome-wide association studies reducing the current bias towards assessment of

candidate genes. A recent low-density whole genome analysis was carried out for breast cancer and identified several potential novel breast cancer loci (Ellis*, et al.*, 2005).

The most common method for identification of genetic association is the case-control study, whereby the particular genotype frequencies are compared in a number of unrelated affected cases and healthy controls. Other study designs may be family-based, such as the case-parent triad design and these designs have the advantage of counteracting the confounding effects due to population stratification that can occur in case-control studies (Cordell and Clayton, 2005). Allelic association is present when the distribution of genotypes differs in cases and controls and such an association provides evidence that the locus under study, or a neighbouring locus, is related to disease susceptibility. The statistical tests used to determine significant evidence of association may be logistic regression, chi-square analysis and odds ratio estimations. As discussed earlier, there are currently a large number of articles being published based on this "common disease common variant hypothesis", however many of these studies are underpowered and fail to be replicated. The main difficulties in carrying out accurate association studies are having adequate power to detect moderate effects, correctly matched controls and avoiding population stratification. Larger sample sizes along with careful selection of candidate SNPs may improve the chances of successfully identifying true associations.

SNP selection has become a matter of debate and there appears to be two schools of thought: first, a map based approach, which focuses on LD between SNPs and second a sequence based approach, focusing on SNPs in coding regions of genes (Botstein and Risch, 2003). The map-based approach is theoretically similar to traditional linkage analysis in that no assumption is made about the type or position of sequence change leading to disease susceptibility. Based on the assumption of LD blocks throughout the genome, it is assumed that a number of tagging SNPs may be selected which could identify the genetic variation of the entire genome. Although this method sounds appealing in its lack of positional and functional bias, estimates of between 500,000 to 1,000,000 SNPs may need to be genotyped in European populations (Botstein and Risch, 2003), although with the completion of the second phase of the HapMap this number may be considerably smaller (Altshuler*, et al.*, 2005). In contrast, the sequence-based approach focuses on identifying SNPs within the coding region of genes either

missense or nonsense mutations, splice site mutations or mutations in regulatory/promoter sequences. The number of coding SNPs in the genome has been estimated to be between 50,000-100,000 and as such represents a vast difference in genotyping costs. The sequence-based approach also offers the advantage of directly identifying causal genetic variants. However, both approaches suffer from individual drawbacks and a decision on the best way forward for the mapping of common diseases remains to be established. For common cancer syndromes a combined strategy of linkage mapping and tumour studies may provide some solutions in the search for novel predisposing loci.

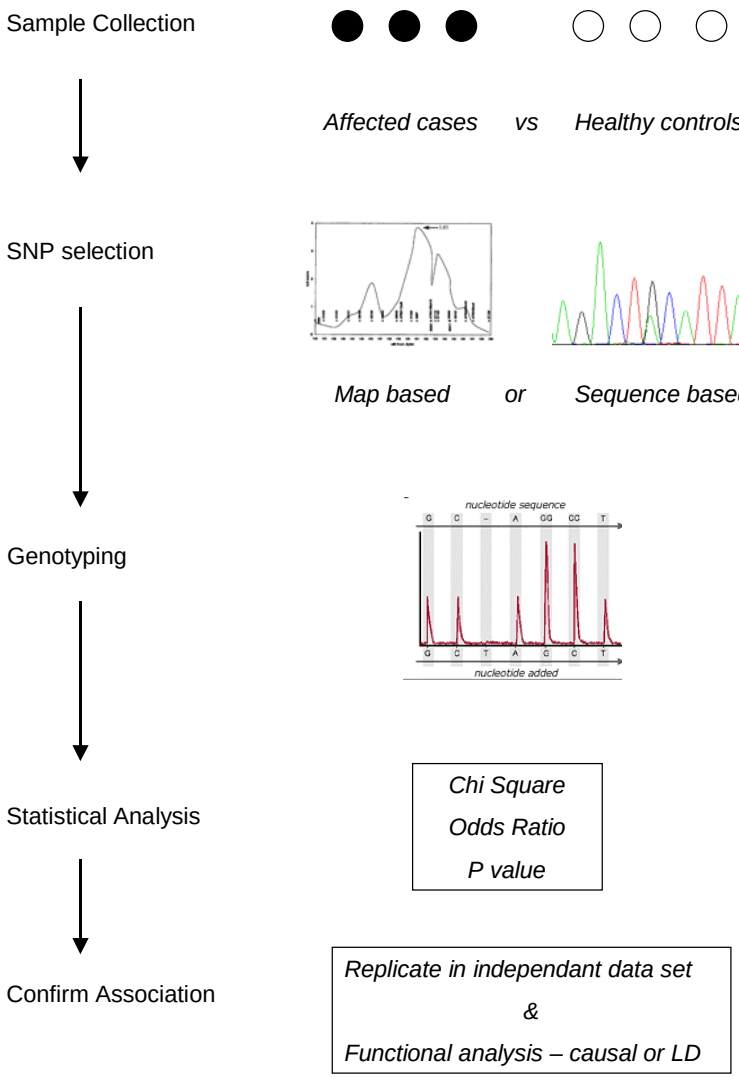

Figure 5. Schematic outline of an association analysis using the case-control study design.

4.3 CYTOGENETICS

There are a number of techniques that may be used to identify cytogenetic abnormalities. Although constitutional chromosomal defects are uncommon in patients with inherited cancers, rare cases when observed have helped to localise a number of cancer genes. Constitutional chromosomal aberrations involving interstitial deletions as well as balanced translocations helped to identify the retinoblastoma (Motegi*, et al.*, 1982) and neurofibromatosis 1 gene respectively. Techniques such as G-banding can be used to detect gross chromosomal aberrations, while fluorescent in situ hybridisation (FISH) has higher resolution for the detection of smaller aberrations, and multi-colour FISH or spectral karyotyping (SKY) can be used to identify complex karyotypes. Cytogenetic techniques have been very successful in detecting cancer related chromosomal translocations, typically associated with leukaemias and lymphomas. The identification of the Philadelphia chromosome BCR-ABL translocation is a prime example. Identification of the fusion genes involved in this leukaemia specific translocation has lead to the successful development of a therapeutic agent. Signal transduction inhibitors (STI) accurately target the kinase domain of specific proteins (e.g. BCR-ABL, KIT) blocking the binding of ATP and hence disrupt the aberrant signalling properties of these proteins, and inhibit tumour specific cell proliferation (Tibes*, et al.*, 2005). Although constitutional aberrations are uncommon in hereditary cancers, somatic aberrations are a characteristic of breast cancer and cytogenetic techniques are useful in identifying gross chromosomal aberrations in addition to copy number changes, which may play a role in carcinogenesis.

4.4 SOMATIC ALTERATIONS IN TUMOURS

In addition to the localisation and identification of germline alterations in cancer, somatic alterations in tumours may provide insight into novel predisposing genes. Traditionally loss of heterozygosity (LOH) has been used as a tool to map tumour deletions. LOH involves the genotyping of microsatellite markers at particular loci in both constitutional and tumour DNA. Apparent homozygosity of particular marker(s) in the tumour that are constitutively heterozygous points to a loss of genetic material in the tumour. The detection of regions of the genome that are frequently and consistently deleted in tumour samples can point to a tumour suppressor locus (Lasko*, et al.*, 1991). Germline mutations in the breast cancer tumour suppressor genes *p53*, *BRCA1* and *BRCA2* are associated with LOH on chromosomes 17p, 17q21

and 13q12-13 respectively (Baker, *et al.*, 1990; Smith, *et al.*, 1992; Collins, *et al.*, 1995). In an analogous method the *PTEN* tumour suppressor gene was discovered based on representational difference analysis of a breast tumour followed by mapping of the deleted fragment on 10q23 and candidate gene analysis (Li, *et al.*, 1997). Numerous LOH studies have been carried out examining chromosomal aberrations in breast cancer and a number of loci have been suggested to be involved in breast carcinogenesis. A genome-wide analysis of sporadic breast cancers observed the highest frequency of LOH at 1q, 4p, 8p, 8q, 11q, 13q, 16q, 17p, 17q and 22q (Shen, *et al.*, 2000). While LOH studies of familial breast cancers detected LOH most frequently at 8p, 16q, 17p, 17q and 19p (Lindblom, *et al.*, 1993a). The fact that a number of these loci are common between sporadic and familial breast cancer suggests they may play an important role in breast cancer pathogenesis. We have examined a locus on 17q, distinct from *BRCA1*, for its role in familial breast cancer predisposition (Paper II).

Comparative genomic hybridisation (CGH) is a technique for the detection of chromosomal aberrations, which allows the simultaneous detection of genomic deletions and amplifications (Kallioniemi, *et al.*, 1992). Tumour DNA and reference DNA are labelled in different colours and mixed prior to hybridising to metaphase chromosomes on a slide. The labelled DNA competes for binding to its complementary sequence within the genome and any dosage change (deletion/amplification) may be visualised by a colour change at the locus in question. CGH has been successful in identifying numerous disease-associated chromosomal loci, including those involved in microdeletion/duplication syndromes, genetic diseases and cancers. The locus responsible for Peutz-Jegher syndrome (*STK11* gene) was initially localised by CGH studies in tumours from affected patients (Hemminki, *et al.*, 1997). The advent of high-resolution CGH and array CGH allows the identification of subtle changes in DNA copy number previously undetectable by metaphase CGH (Pollack, *et al.*, 1999). In addition to identifying specific predisposing loci, CGH is a powerful tool for identification of tumour specific alteration patterns. *BRCA1* and *BRCA2* tumour profiling has revealed specific sets of genomic aberrations associated with these tumour types (Wessels, *et al.*, 2002; van Beers, *et al.*, 2005). *BRCA1* and *BRCA2* mutations may therefore cause breast cancer through different pathways and the identified CGH profiles may point to loci involved in disease progression. In addition, the classification of tumours based on their specific CGH profiles may be a tool with which to classify

non-BRCA1/2 associated familial breast cancer and in combination with other techniques may allow the identification of novel predisposing loci. CGH analysis of 18 familial non-BRCA1/2 breast tumours identified two main regions of loss one on chromosome 6q and a second on chromosome 17q (Paper II). The region on 17q was investigated further and small region on 17q11.2-12 was excluded as harbouring a predisposing mutation in the investigated families. Interestingly, chromosome 6q is the focus of a more recent study as a candidate locus in another set of familial non-BRCA1/2 families (Paper III).

Microarray analysis, in contrast to the strategies mentioned above, is a method of global gene expression profiling. Microarray analysis is also a hybridisation based technique, the difference being that total RNA from the tissue of interest in labelled and hybridised to cDNA or oligonucleotide probes in an effort to detect differences in expression. In the context of tumour profiling, typically reduction in gene expression is associated with tumour suppressor genes, while amplifications are classically associated with oncogene activations. Global expression profiling has a general role in the classification of tumours; this classification may be clinically useful, classifying tumours based on their prognosis or response to certain therapies (Sorlie*, et al.*, 2001; van 't Veer*, et al.*, 2002). Although this technique is not based on the direct detection of disease associated alterations, the expression patterns of tumours may be useful in identifying clinical and pathological breast cancer sub-types and in the classification of non-BRCA1/BRCA2 familial breast tumours. Expression profiling of breast cancer tumours has successfully distinguished breast tumours of the BRCA1, BRCA2 or non-BRCA1/2 (BRCA3) sub-type (Hedenfalk*, et al.*, 2001; Hedenfalk*, et al.*, 2003), suggesting that predisposing germline mutations may influence the genetic progression of the tumours resulting in distinct expression profiles. We have investigated two non-BRCA1/BRCA2 families, which exhibited similar expression profiles by microarray analysis. These two families were determined to share a common ancestral haplotype on chromosome 6q**,** suggesting they may be genetically homogenous with a common breast cancer predisposition (Paper III).

The current inability to detect novel breast cancer genes may result in part from the lack of distinct phenotypes in the remaining families. Thus, the combined approach of tumour profiling and traditional genetic linkage analysis may lead to the classification

of these families into sub-groups, which are more genetically homogenous and hence facilitate the identification of novel moderate penetrance breast cancer susceptibility genes.

4.5 MOUSE MODELS

Animal models afford us the ability to recapitulate human disease, thereby allowing us to study and understand the disease process in the animal under study. This understanding of biological disease mechanisms may yield insight into important developmental and pathogenic disease pathways enabling the development of therapeutic strategies, which can be extended to human disease prevention. Although distinct and appreciable species-specific differences exist, animal models allow the ability to detect novel cancer susceptibility genes in addition to the possibility of studying the functional aspects of known disease genes. Rodents are commonly used models for human disease and the rat represents a more comparable model for human cancer than the mouse, (Anisimov, *et al.*, 2005), however, mice are studied more frequently, as they are more amenable to genetic manipulation.

Currently mouse models exist to study the function of known breast cancer genes and the consequences of their mutation. Initial efforts to produce knock-out *Brca1* and *Brca2* mice for functional studies met with difficulties as homozygous mutant mice die in early embryogenesis (Moynahan, 2002). The embryos are characterised by severe developmental delay and defects in cell proliferation. In addition, the embryos of both *Brca1* and *Brca2* knockout mice exhibit chromosomal abnormalities suggestive of a defect in DSB repair (Deng and Brodie, 2001). The similarity in the phenotypes of *Brca1* and *Brca1* knockout embryos suggested a functional link between the two genes. The ability to study the *Brca1* and *Brca2* function *in vivo* came with the advent of new techniques; the development of conditional knockout mice, where the gene of interest in knocked out in a tissue specific manner. The use of conditional knockouts for *Brca1* and *Brca2* lead to the first evidence that *Brca1* and *Brca2* were tumour suppressor genes in the mouse. Mice with a deletion of *Brca1* exon 11, which was conditionally expressed in mammary epithelial cells, developed mammary tumours (Xu, *et al.*, 1999a). The first conditional knockout of *Brca2* in mammary epithelium created a mutation in exons 3 and 4, and resulted in 77% of mice developing mammary tumours compared to no tumours in control animals (Ludwig, *et al.*, 2001).

Mouse models for human disease may allow us not only to study the function of known genes but also to identify novel genes that may modify the penetrance of these known genes. This has been clearly demonstrated in an animal model of colon cancer, the *MIN* mouse. A modifier of the *Apc* mutant Min-1 gene, Mom-1 strongly inhibited polyp formation in mice (Dietrich, *et al.*, 1993). This has also been observed in breast cancer mouse models where the defects observed in *Brca1*-deficient embryos are partially rescued by *p53*-deficient background (Deng and Brodie, 2001). This finding highlights the role of genetic background in the development of appropriate animal models of human disease and may also lead to the identification of novel modifiers of human disease phenotypes. The ability to develop conditional knockout mice to study the function of novel genes will be of great importance in developing our understanding of normal development and disease processes. However, the difficulty remains in the choice of candidates genes and in the economic ability to generate large numbers of mutant mice and determine their exact phenotypes.

A number of different techniques have been outlined thus far that can aid in the identification of genes involved in human disease, including breast cancer. However the difficulties to date in identifying a novel BRCA3 gene underscore the need to develop new strategies for cancer gene identification. Currently, a combined approach utilising a number of these techniques may offer an increased opportunity to detect novel genes involved in breast cancer. This thesis aims to identify novel genes involved in familial non-BRCA1/2 breast cancer and has used the combined strategies of linkage analysis and tumour profiling to this aim. In addition, the role of low-penetrance alleles has also been examined in relation to non-BRCA1/2 familial breast cancer.

5 AIMS

The main aim of this thesis work was to investigate the genetic basis of breast cancer in families where the disease is not attributable to mutations in either of the known breast cancer genes *BRCA1* or *BRCA2*.

The specific aims were to:

Determine if germline mutations in the *HIN-1* gene were responsible for the increased familial risk attributed to non-BRCA1/2 families that exhibited suggestive linkage to the candidate locus on chromosome 5q.

Identify chromosomal aberrations associated with non-BRCA1/2 breast tumours and, in combination with linkage data, determine novel candidate loci that may harbour a breast cancer predisposing gene in these specific families.

Re-evaluate genetic linkage data based on identification of family groups hypothesized to be genetically homogenous and further investigate putative chromosomal loci of interest in these families.

Investigate the contribution of polymorphisms in the *ESR2* gene as low penetrant breast cancer susceptibility alleles in sporadic and familial breast cancer cases.

6 MATERIALS & METHODS

6.1 BREAST CANCER FAMILIES

This thesis is based on the study of familial non-BRCA1/2 breast cancer families and the majority of breast cancer patients included were from families included in a previous genome wide linkage analysis of 102 non-BRCA1/2 families. Breast cancer families were counselled at the Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. Two hundred and eighty five pedigree members of whom 245 were breast cancer patients were included in the linkage analysis. Twenty-five families had two affected cases, forty-three families had three cases, twenty-four families had four cases, eight families had five cases and two families had six cases. The average age at diagnosis was 54.7 years, with an age range of 30-72 years. All familial cases proceeded through genetic counselling and those who met the criteria for *BRCA1/2* testing were screened negative for mutations while the remaining samples did not fulfil the criteria for *BRCA1* or *BRCA2* testing. DNA was extracted from peripheral blood lymphocytes by standard procedures. All breast tumour samples were collected from patients after surgery, snap frozen and stored at –80° C until use.

In Paper I, breast cancer patients from ten families displaying a positive LOD score and suggestively linked haplotypes for the two markers flanking the *HIN-1* gene were screened for germline mutations. A total of 226 sporadic breast tumours and 24 familial breast tumours were analysed in this study. Of the 226 sporadic breast tumours 14 were classified as BRCA1-like based on them being ER- and HER2-negative, high-grade tumours. Of the 24 familial tumours, 18 were from *BRCA1* mutation carriers and 6 were from non-BRCA1/2 families.

In Paper II, CGH was carried out on a total of eighteen familial non-BRCA1/2 breast tumours. Twelve from high-risk families, defined as having ≥ 3 first or second-degree relatives with breast cancer. The remaining six tumours were from low-risk families, defined as having only 2 first-degree relatives with breast cancer. Ten non-BRCA1/BRCA2 breast cancer families were subsequently analysed for germline mutations in ten candidate genes on chromosome 17q11.2-12. Five of these families had tumours that had been included in the CGH analysis and displayed loss of

chromosome 17. The other five families exhibited loss of chromosome 17 in LOH studies.

In Paper III, two non-BRCA1/2 breast cancer families were analysed initially by linkage and fine-mapping and a candidate locus on chromosome 6 was identified (Families 6006 and 6043). Following genotyping of all 102 breast cancer families for the four microsatellite markers on chromosome 6q14, an additional 29 breast cancer families were added to the analysis. Controls were 95 blood donors from the Stockholm region of Sweden. Tumour RNA was available for families 6006 and 6043 and was used for reverse-transcriptase PCR analysis.

In Paper IV, 400 familial and 323 sporadic breast cancer patients were examined. The sporadic patients and 141 of the familial cases were collected as a population based breast cancer cohort at the Clinic of Oncology at Södersjukhuset and Karolinska University Hospital, Stockholm, Sweden while the remaining 259 familial cases were collected at the Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. The 400 familial cases were divided into two groups, namely cases from high-risk families (n=212) and cases from low-risk families (n=188), based on pedigree analysis. Families with multiple affected family members and an apparent dominant mode of inheritance are termed familial high-risk, while those families with two affected women and an unclear mode of inheritance are termed familial low-risk. This sub-classification of families into two groups is consistent with empirical risk estimates, where high-risk families have a 3-5 fold increased risk of developing breast cancer while low-risk families have a 2-3 fold increased risk (Claus, et al., 1996; Gail, *et al.*, 1989). For all samples from Södersjukhuset and Karolinska University Hospital, a family history of breast cancer was obtained and all samples were screened for mutations in *BRCA1* exon 11, which accounts for the majority of Swedish mutations (Margolin, *et al.*, 2004). Controls were 480 blood donors from the Stockholm region of Sweden.

6.2 METHODS

6.2.1 Methylation-specific PCR (MSP)

Methylation-specific PCR distinguishes unmethylated from methylated alleles in a given gene. The ability to detect methylated and unmethylated alleles is based on sequence changes produced following bisulfite treatment of DNA, which converts unmethylated, but not methylated, cytosines to uracil. Subsequently PCR analysis can distinguish the alleles, using specific primers designed for either methylated or unmethylated DNA. In order to determine the location of methylated cytosines, DNA was extracted from cells, bisulfite treated, and purified. PCR amplification was initially performed by using primers designed to amplify the coding strand of bisulfite treated DNA. PCR products were subcloned into pZERO (Invitrogen) vector and four to six individual clones were sequenced for each PCR product. Based on sequence analysis, primers were designed for the amplification of unmethylated and methylated genomic DNA. Placental DNA treated *in vitro* with *Sss*I bacterial methylase was used as a positive control for methylated alleles. DNA from normal lymphocytes was used as a negative control for methylated genes.

6.2.2 Linkage Analysis & Fine Mapping

Linkage analysis is a powerful tool for the localisation of genes involved in inherited syndromes. This family based technique relies solely on the knowledge that the disease phenotype is inherited in the family. The principle of linkage analysis relies on the co-segregation of a genetic marker with the disease. The closer the genetic marker and the disease loci are the less likely they are to be separated by a recombination event. The marker and disease trait are linked if they segregate together in a family pedigree. The tools to map the disease loci are based on the common variation within the human genome. Genetic markers are polymorphic sites in the genome, which can be easily assayed and are specifically mapped. Previously bi-allelic restriction fragment length polymorphisms (RFLPs) were used for linkage analysis. Nowadays, however, most genome-wide linkage scans rely on dense sets of highly polymorphic microsatellite markers. Microsatellite markers are either di-, tri-, tetra- or penta-nucleotide tandem repeat polymorphisms, which are located throughout the genome at relatively even spacing. They are easily assayed by PCR amplification of the repeat and flanking regions and individual genotypes are clearly distinguished as differences in the amplicon sizes.

A genome-wide linkage scan was carried out on the 102 familial non-BRCA1/2 breast cancer families. A total of 380 fluorescently labelled microsatellite markers covering the entire genome with an average spacing of 10cM were used. The average heterozygosity of the markers was 0.76 in our sample set. Results from the genome-wide scan revealed no overall positive LOD scores. However, haplotype and genotype data was available for all 102 families. This data could then be used to investigate families in which disease was potentially linked to candidate loci, as was the case with the *HIN-1* gene in Paper I. These data were also used to determine in which families disease was linked to chromosome 17 (Paper II). Analysis of genotype and shared haplotype data for families 6006 and 6043 revealed that breast cancer in both families was suggestively linked to three chromosomes (Paper III).

The 380 original microsatellite markers used in the genome-wide scan were spaced at roughly 10cM intervals, therefore chromosomal loci of interest were further investigated by fine-mapping the regions for the families in question. Additional microsatellite markers were selected based on their position with respect to the original linked markers and on their reported heterozygosity. Each fluorescently labelled microsatellite marker was amplified independently by PCR. Amplified fragments were pooled based on the size of the amplified fragment and separated on an ABI Prism 377 automated DNA sequencer (PE Applied Biosystems, Foster City, CA), together with internal size standard. Electrophoretic data was analysed using Genescan3.1® and Genotyper2.0® software programs (PE Applied Biosystems, Foster City, CA). Fine-mapping of a region on chromosome 17q was carried out to determine the minimal possible region of overlapping linkage in ten families (Paper II). Fine-mapping of three chromosomal regions was carried for families 6006 and 6043 in Paper III.

A statistical test is needed to measure the significance of linkage and the determination of linkage involves the calculation of LOD scores. These calculations are carried out by sophisticated linkage analysis software packages such as the Genehunter or Simwalk2 programs (Kruglyak*, et al.*, 1996; Sobel and Lange, 1996). Parametric linkage analysis relies on the specification of certain parameters, pertaining to a known mode of inheritance. Non-parametric linkage analysis does not rely on a known genetic model. In the original genome-wide linkage analysis both parametric and non-parametric linkage analysis were carried out with the Genehunter

program. When carrying out parametric linkage analysis an autosomal dominant mode of inheritance was assumed with a disease allele frequency of 0.0001 and an equal female to male recombination rate. The penetrance for homozygous normal, heterozygous and homozygous affected was 0.05, 0.80 and 0.80 respectively. Both the Genehunter and SimWalk2 programs were used for generating haplotypes in addition to manual haplotype estimation.

6.2.3 Comparative genomic hybridisation (CGH)

CGH of metaphase chromosomes enables genome-wide analysis of gross DNA copy number changes. This technique was first reported in the 1990s (Kallioniemi, *et al.*, 1992) and is widely used for the analysis of tumour genomes and constitutional aberrations. Tumour DNA and normal DNA are differentially labelled with fluorochromes and co-hybridised to normal metaphase spreads. Due to the competitive nature of the binding, regions of the genome that are amplified are detected by an increase in intensity of the tumour fluorochrome, while regions of genetic loss are identified by increased intensity of the normal fluorochrome. The presence of gains and losses of genetic material may signify the position of oncogenes or tumour suppressor genes respectively. Traditional metaphase CGH has a low-resolution of between 10-20Mb and aims to map chromosomal aberrations to their physical position on the chromosome. Recently a microarray based approach has been implemented, array-CGH, which enables the mapping of copy number alterations relative to the genome sequence and the resolution is determined only by the spacing of the clones (Pollack, *et al.*, 1999).

In Paper II we utilised traditional metaphase CGH to analyse copy number changes in 18 familial non-BRCA1/2 breast tumours. CGH was performed according to Kallioniemi *et al.*, (Kallioniemi, *et al.*, 1994) with minor modifications. Test and reference DNA were labelled, by nick translation, with digoxygenin-11-dUTP and biotin-14-dATP respectively. Unlabelled Cot-1 DNA was used to suppress hybridisation to highly repetitive sequences. Fluorescence intensity ratios were calculated and a test:reference ratio of <0.8 was considered evidence for loss of a chromosomal region, while a test:reference ratio >1.2 indicated the gain of a region.

6.2.4 DHPLC

DHPLC (Denaturing high performance liquid chromatography) allows the automated detection of single base substitutions as well as small insertions and deletions. Under partially denaturing conditions DNA heteroduplexes are formed between mutated and normal DNA molecules by mixing, denaturing and reannealing and can be distinguished from homoduplexes consisting of normal DNA only, by separation on a liquid chromatography column under appropriate conditions. Positively charged TEAA (triethylamine acetate) ions are adsorbed to the non-polar solid phase DNASep column, renatured PCR products can then be injected into the column and the negatively charged dsDNA molecules are bound to the positively charged surface of the column. The strength of the binding is based on the number of ion pairs formed between the negatively charged DNA and the positively charged TEAA. The column is then subjected to increasing concentrations of the organic solvent ACN (Acetonitrile), which results in dissociation of the ampiphilic ions and the dsDNA from the column. Heteroduplexes which have incorrect base-pairing at the mutation site will have fewer ion-pairing bonds than homoduplexes with no mutation. Therefore, heteroduplexes will be eluted from the column earlier than homoduplexes and are typically displayed as more than one peak on the chromatogram as compared to one peak, which represents normal homoduplexes.

DHPLC was carried out using a Transgenomic Wave DNA Fragment analysis system, an automated instrument equipped with a DNASep column (Transgenomic, Crewe, United Kingdom). Renatured PCR products were loaded directly onto the system and the optimal column running temperature, and concentrations of TEAA and ACN for each specific PCR amplicon were determined directly by the WAVEMAKER 3.4 software. Abnormal elution profiles were identified by visual inspection of the chromatograms. DHPLC can only detect the presence of a mismatched base and not the location or chemical nature of the mismatch. Therefore PCR products exhibiting aberrant profiles were noted and re-amplified from genomic DNA for direct sequencing. DHPLC and direct sequencing was used for mutation analysis of the ten genes on chromosome 17 (Paper II).

6.2.5 Direct Sequencing

Today the most common form of direct sequencing is cycle sequencing, which is a modification of the traditional Sanger sequencing method (Sanger, *et al.*, 1977).

Cycle sequencing relies on the use of chemically modified nucleotides in the sequencing reaction; these dideoxynucleotides are usually fluorescently labelled in four different colours representing the four different bases (A, T, G, C). The sequencing reaction requires purified DNA template, usually in the form of a PCR product, an appropriate sequencing primer, a thermostable sequencing enzyme and a mix of dNTPs and ddNTPs. During the sequencing reaction, the double stranded PCR product is denatured and at an appropriate temperature the sequencing enzyme begins to polymerise the addition of new bases complementary to the template strand. This process continues until a fluorescent ddNTP is incorporated. The chemical modification of the ddNTPs prohibits further polymerisation reactions. Subsequent cycling reactions create new templates for the polymerase and additional dNTPs and ddNTPs are added to each new chain. Finally, the sequencing product is composed of a collection of DNA strands of different lengths each with a ddNTPs at the 3'end. These DNA fragments are size fractionated on a polyacrylamide gel and the fluorescently labelled ddNTPs are analysed by the use of the laser within an automated DNA sequencer. Sequences can be directly read from generated chromatograms and base pair substitutions are easily recognisable as overlapping peaks in the chromatogram.

Direct sequence analysis is an accurate and specific method for the detection of base substitutions and insertions/deletion mutations. It does not allow the identification of gross gene rearrangements. The genomic structure of all genes studied in this thesis was determined by analysis of public sequence databases (NCBI, UCSC and Ensembl). Primers were designed to amplify all exons including exon/intron boundaries, the 5'UTR and 3'UTR regions and in some cases the putative promoter regions using the online Primer3 software package.

Amplified PCR products were cleaned prior to sequencing by incubation with the ExoSap enzyme (GE Healthcare). Sequencing reactions were carried out using the ABI Big Dye Terminator v3.1 kit (Applied Biosystems, Foster City, California). Cleaned sequencing products were electrophoresed in an ABI377 automated sequencer or an ABI 3730 XL capillary sequencer.

6.2.6 Reverse-Transcriptase PCR (RT-PCR)

RT-PCR is an RNA based method for studying the expression of specific mRNAs within tissues. Total RNA from the tissue of interest is first converted into single-stranded complementary DNA (cDNA), using reverse transcriptase enzyme and either random hexamer priming method (used in Paper III) or oligo dT primer method. Primers are designed for the gene(s) of interest in a manner such that at least one primer is placed covering two exons, in an effort to reduce genomic DNA contamination.

In Paper III cDNA was prepared from total RNA extracted from breast tumours and EBV transformed cell lines, available from families 6006 and 6043. Amplified RT-PCR products were size fractionated on an agarose gel along with normal control samples in order to detect additional or aberrant bands. RT-PCR was carried out for all six genes in order to detect mutations affecting splicing or the presence of large insertion/deletion mutations.

6.2.7 mRNA *in Situ* Hybridization

The cellular expression of specific genes can be determined by hybridization with labeled probes complementary to the mRNA of interest. The use of tissue sections enables accurate localization and visualization *in situ* of the relevant molecules. To generate templates for *in vitro* transcription reactions, full-length human *HIN-1* cDNA was PCR amplified and subcloned into pZERO 1.0 (Invitrogen, Carlsbad, CA) and used for the generation of sense and antisense digitonin-labeled riboprobes followed by mRNA *in situ* hybridizations. The hybridized sections were observed with a Nikon microscope, and images were obtained using a SPOT charge-coupled device camera and processed with Adobe Photoshop. Hybridizations using the sense probe were carried out as a control for non-specific hybridizations. mRNA *in situ* hybridization was used to determine HIN-1 expression in breast tissues (Paper I).

6.2.8 Pyrosequencing

Pyrosequencing is a real-time sequencing technique for the determination of nucleic acid sequences (Ronaghi, *et al.*, 1996). Pyrosequencing involves the monitoring of an enzymatic cascade that occurs during the sequencing reaction, this cascade begins with the release of inorganic pyrophosphate (PPi) during nucleotide incorporation. The released PPi is converted to ATP by ATP sulfyrase and this reaction provides energy for luciferase to oxidize luciferin resulting in light generation that is registered as a peak

in the pyrogram. Unincorporated nucleotides are degraded by apyrase prior to the addition of the next nucleotide, allowing the repeated addition of nucleotides. Since the added nucleotide is known, the sequence of the template can be determined. This technique is widely used for the genotyping of SNPs, the pyrosequencing primer used in the reaction is designed so that the 3' end of the primer hybridises just a few bases away from the polymorphic site. Genotypes can be accurately distinguished by visual inspection of the pyrogram readouts.

Following PCR amplification, the rs1256049 and rs4986938 SNPs were genotyped by Pyrosequencing (Paper IV). Biotinylated PCR templates were immobilized on streptavidin-coated paramagnetic Sepharose beads in Binding Buffer. The bead-template complexes were denatured, and subsequently incubated with a mixture of Annealing Buffer and sequencing primer. Annealing took place at 80°C for 2 min followed by cooling to room temperature. Real-time pyrosequencing was carried out in an automated 96-well pyrosequencer using PSQ SNP96MA enzymes and substrates (Biotage, Uppsala, Sweden). Pyrogram readouts were converted to numerical values for peak heights using a software module designed for this purpose (Biotage, Uppsala, Sweden). Genotypes were analyzed manually by visual inspection of each pyrogram by two independent researchers.

6.2.9 Restriction fragment length polymorphism (RFLP)

RFLP analysis allows the distinction of bi-allelic sequence variants that either create or abolish restriction enzyme recognition sites. Amplified PCR fragments are digested with a suitable restriction enzyme that recognizes specific sequence surrounding the polymorphism of interest. The restriction patterns created following digestion allow easy genotype identification. The rs928554 SNP was evaluated using restriction enzyme digestion instead of pyrosequencing due to the sub-optimal sequence surrounding this variant. The G > A change generates a restriction site for the enzyme Tsp509I. A 265bp amplified PCR products was digested at 65° for 1 hour and separated on a 4.5% agarose gel. Samples homozygous for the G allele demonstrated 2 bands upon digestion, one band of 187bp and a second of 78bp. Heterozygote samples produced 4 bands of size, 178, 149, 78 and 38bp. The homozygous variant samples produced 3 bands of size 149 bp, 78 bp and 38bp.

6.2.10 Association analysis

Genotypic and allelic data for each of the three polymorphisms were compared between the cases and controls using chi-square analysis. Odds ratios with 95% confidence intervals were calculated using wt/wt genotype as reference genotype and comparing wt/var and var/var to this reference. All three polymorphisms were tested for adherence to Hardy-Weinberg equilibrium (HWE) in both cases and controls.

Estimation of linkage disequilibrium (LD) and haplotype analysis was carried out using both the Haploview v3.1.1 1 (Barrett, *et al.*, 2005) program and the UNPHASED program (Dudbridge, 2003). The UNPHASED program was run through the GLUE interface at http://www.rfcgr.mrc.ac.uk/. Genotype and marker data were loaded in linkage format files into the Haploview v 3.1.1 program for estimation of LD in the region and generation of inferred haplotypes. The default algorithm used by the program is based on the work of Gabriel *et al*. (Gabriel, *et al.*, 2002), where 95% confidence bounds on D' are generated and each comparison is called strong LD, inconclusive or strong recombination. A block is generated if 95% of informative comparisons are in strong LD.

Association analysis of inferred haplotypes was also carried out using the COCAPHASE program within the UNPHASED package. This program uses standard unconditional logistic regression identical to the modelfree method of T5 of EHPLUS and the log-linear modeling. The EM algorithm is used to obtain haplotype frequency estimates.

7 RESULTS & DISCUSSION

7.1 PAPER I

Lack of *HIN-1* methylation in BRCA1-linked and "BRCA1-like" breast tumours

HIN-1 is a putative tumour suppressor gene located on chromosome 5q, a region frequently lost in *BRCA1* tumours (Nathanson, *et al.*, 2002). The *HIN-1* gene was identified by SAGE experiments comparing normal mammary epithelial cells and ductal carcinoma *in situ* cells (Krop, *et al.*, 2001). *HIN-1* appears to be a secreted growth inhibitory cytokine with a role in epithelial cell differentiation. Based on its location, putative function and the discovery that *HIN-1* is silenced by methylation in sporadic breast tumours, we hypothesized that *HIN-1* may act as a low to moderate penetrant breast cancer gene in families displaying suggestive linkage to chromosome 5q. In addition, *HIN-1* may be inactivated by a genetic mechanism in *BRCA1* tumours.

In order to determine if germline alterations in the *HIN-1* influence breast cancer risk, we analysed the coding region of the gene for mutations. Ten non-BRCA1/2 breast cancer families displaying suggestive linkage to the *HIN-1* locus on chromosome 5 were analysed for germline mutations. No sequence alterations were found in the three *HIN-1* coding exons. In addition, fifteen *BRCA1* associated tumours and thirty-five sporadic breast tumours were also investigated for alterations and none were found.

As no mutations were detected in the *BRCA1* tumours, we investigated whether, similar to sporadic breast carcinomas, HIN-1 expression is silenced by methylation in these *BRCA1* tumors. In striking contrast to sporadic breast tumors, *HIN-1* was completely unmethylated in most *BRCA1* tumors. Although the number of *BRCA1* tumors analyzed was relatively small, the difference in the frequency of *HIN-1* methylation between *BRCA1* and sporadic tumors was statistically significant ($P < 0.0001$). mRNA *in situ* hybridization confirmed that HIN-1 was expressed in *BRCA1* tumours, whereas sporadic tumours were largely negative for HIN-1 expression.

A sub-set of sporadic tumours displaying *BRCA1* like characteristics (steroid receptor and HER2 negative, high histologic grade tumors), were also analysed by MSP to determine if, similar to *BRCA1* tumours, the *HIN-1* gene was unmethylated. *HIN-1* was

found to be unmethylated in the majority of these tumours and the frequency of *HIN-1* methylation was statistically significantly different between BRCA1-like and other types of sporadic tumors ($P<0.01$).

Our results suggest that genetic variations in the coding region of *HIN-1* are not likely to influence breast cancer risk and do not appear to play a major role in breast tumourigenesis. In contrast, silencing of *HIN-1* due to methylation is a frequent and early event in the majority of sporadic breast carcinomas, but not in *BRCA1* and BRCA1-like tumors. Interestingly, BRCA1 associated and BRCA-1 like sporadic tumours were unmethylated at the HIN-1 promoter suggesting that similar genetic pathways may be disrupted in these tumour types.

Lack of HIN-1 expression has also been noted in a number of other malignancies, including lung cancer and nasopharyngeal cancer (Marchetti, *et al*., 2004; Wong, *et al*., 2003). In the case of lung cancer, lack of HIN-1 expression has been correlated with poor clinical outcome and suggests that *HIN-1* status may be a useful biomarker in certain malignancies (Marchetti, *et al*., 2004). The lack of HIN-1 expression has not been correlated with any sequence alterations and methylation of the promoter region appears to be the mechanism of gene silencing (Krop, *et al*., 2001; Shigematsu, *et al*., 2005) Previously, a tumour suppressor function has been ascribed to *HIN-1* and recent data further support this idea and suggests that *HIN-1* is a potent inhibitor of anchorage-dependant and independent cell growth (Krop*, et al.*, 2005). This *HIN-1* function may be due to the combined functions of apoptosis induction and regulation of cell cycle entry. Although *HIN-1* may not have a predisposing role in breast cancer, its altered expression may have consequences in the initiation and progression of breast cancers.

7.2 PAPER II

CGH analysis of familial non-BRCA1/2 breast tumours and mutation screening of a candidate locus on chromosome 17q11.2-2.

In an effort to define novel predisposing breast cancer genes in non-BRCA1/2 breast cancer families we used the combined approach of comparative genomic hybridisation (CGH), and genetic linkage analysis. CGH has been successful in the mapping of several disease genes, including the gene responsible for Peutz-Jegher syndrome, *STK11* on chromosome 19p (Hemminki*, et al.*, 1997). CGH allows DNA copy number changes in the entire genome to be assessed in a single hybridisation, regions of genetic loss may harbour potential tumour suppressor genes.

Eighteen non-BRCA1/2 familial tumours were analysed by CGH for copy number changes. The eighteen tumours were derived from high-risk families (n=12) and low-risk families (n=6). In high-risk families the most frequent chromosomal alterations were loss of one copy of chromosome 17 (n=5) or loss of chromosome 6q (n=6) and these alterations appeared to be mutually exclusive. In low-risk families the most frequent chromosomal aberrations was gain of chromosome 8q (n=4). Although the frequencies of chromosome 17 and chromosome 6q losses were comparable, tumours that exhibited loss of chromosome 17 tended to have fewer aberrations. This data suggested that loss of chromosome 17 may represent an early event in tumourigenesis and thus may be related to a predisposing event. In addition, loss of chromosome 17 had been observed in previous LOH studies on similar material (Lindblom, et al., 1993a; Zelada-Hedman*, et al.*, 1994). Therefore chromosome 17 was the focus of our investigation in this study.

Data from our previous genome-wide linkage analysis (unpublished data) was available for 102 non-BRCA1/2 breast cancer families, including the 18 families included in the CGH analysis and a proportion of the families used in the previous LOH studies. Analysis of genetic linkage data revealed that for the ten families whose tumours exhibited loss of chromosome 17, disease was suggestively linked to chromosome 17. Breast cancer in these ten families was linked to different regions of chromosome 17, however the minimal overlapping region of linkage was between D17S1293 and D17S1294. An additional 6 microsatellite markers were added between these original markers and genotyping revealed a small region of possible

linkage shared between the families. This region of suggestive linkage spanned the microsatellite markers D17S1880 and D17S1293.

Database analysis revealed that ten genes were located within this candidate region on chromosome 17q11.2-12, *TADA1*, *ACCN1*, *TLK2* and a cluster of 7 C-C chemokine genes. The ten families were analysed for germline mutations in these candidate genes. DHPLC analysis of all coding exons, including exon/intron boundaries followed by direct sequencing of aberrant profiles was carried out for each of the ten genes within the region. A total of ten sequence variants were identified in these genes, four of which represented silent mutations in coding exons, five were intronic variants and the last was a mutation in the regulatory region of *CCL1*. All available family members were sequenced for the identified variants and two variants in the *ACCN1* gene and one variant in the *CCL2* gene were found in all available affected family members. However, the variants in *ACCN1* are currently reported in the Ensembl database as SNP rs2228990 and SNP rs2228989, with heterozygote frequencies of about 30% in the normal population. The frequency of the *CCL2* variant, which affects the DNA sequence at the C-C amino acid motif, has not yet been determined in the normal population but is reported in the Ensembl database.

In this study we have examined familial non-BRCA1/2 breast tumours and determined a pattern of chromosomal aberrations; namely losses of chromosome 6q, 11q22-qter, 17 and 18 and gains of 1q, 8q and 16p. The loss of 6q and 11q and the gains of 1q, 8q and 16p were also seen in a previous CGH study by Kainu et el., which examined 37 familial non-BRCA1/2 tumours (Kainu*, et al.*, 2000). It may become apparent with additional studies on similar materials, that like *BRCA1* and *BRCA2* tumours, non-BRCA1/2 tumours may harbour a specific pattern of chromosomal alterations (Wessels, *et al*., 2002; van Beers, *et al*., 2005; Johnsson, *et al*., 2005). Although the results from this study showed no evidence for a predisposing mutation in any of the ten genes in the region on chromosome 17q11.2-12, the combined approach of CGH, LOH and genetic linkage analysis offers the ability to define chromosomal loci of interest in certain sets of families.

7.3 PAPER III

A common haplotype on chromosome 6q14 shared by two non-BRCA1/2 breast cancer families; analysis of a 2.8 Mb region.

The inability to detect a novel "BRCA3" gene necessitates new study designs, which combine multiple approaches with the hope of identifying novel predisposing loci. In this study we combined data from our previous genome wide linkage scan with data from a cDNA microarray study of fourty eight familial non-BRCA1/2 tumours. The microarray study revealed five family groups that consistently clustered together through repeated analysis. We hypothesized that these family groups may represent genetically homogenous sub-groups and that their shared tumour profiles may reflect a shared genetic predisposition. As such, linkage data for these family groups was re-examined to determine regions of shared linkage between the families.

This study focuses on one family group consisting of two families; family 6006 and 6043. Initial re-evaluation of the linkage data for these families identified suggestive linkage to three chromosomal loci. Fine-mapping of all three loci, by genotyping additional microsatellite markers in the three regions, ruled out chromosome 7 and chromosome 10. Breast cancer in both families was linked to chromosome 6, with suggestive LOD scores of 1.48 and 0.78 for family 6006 and 6043 respectively. The total region of shared linkage encompasses 43.8Mb and contains over 400 known genes. Both families were determined to share a common four marker linked haplotype (2-7-5-2) within this region, this haplotype spanned 2.8Mb and genotyping an additional ten SNP markers further supported this shared haplotype on chromosome 6q14.1. The entire set of 100 non-BRCA1/2 breast cancer families were also typed for these markers to determine if this was a common haplotype. Although no other family shared this haplotype, an additional 29 families did exhibit suggestive linkage to the region and were added to the study.

Database analysis revealed the presence of six known genes within the 2.8Mb region on chromosome 6q14.1; *HTR1B*, *IRAK1BP1*, *PHIP*, *HMGN3*, *C6ORF152* and *SH3GBRL2*. All of these genes are expressed in mammary tissue and have interesting functional roles that could possibly act as putative breast cancer predisposing genes. Initially one affected member from each of the 31 breast cancer families was selected

for germline mutation screening of the six genes. The entire coding region including exon/intron boundaries and regulatory regions such as the 3' and 5' UTR and putative promoter regions were sequenced. RT-PCR was carried out for families 6006 and 6043 in order to detect aberrant splicing mutations or insertion/deletion mutations.

A total of 53 sequence alterations were identified in the six genes, no frameshift or nonsense mutations were identified, 19 coding mutations were identified, ten of which were non-synonymous missense mutations and 9 were synonymous missense mutations. Twenty-five mutations were identified in the regulatory regions of genes, either the 3' or 5' UTR of putative promoter regions and nine intronic alterations were detected. The focus of this current study was families 6006 and 6043 and as they shared a common haplotype it was hypothesized that they may share a common predisposing founder mutation. Eleven sequence variants were shared by families 6006 and 6043, one of which was a missense mutation in the PHIP gene**,** Leu1093Pro, while the remainder of the shared alterations were mainly intronic or silent mutations. All eleven variants were frequent in the other 29 breast cancer families and when tested in 95 normal healthy controls had heterozygote frequencies of approximately 50%.

These results suggest, that although families 6006 and 6043 were found to share a common haplotype on chromosome 6q14.1, this region is unlikely to harbour a highly penetrant predisposing mutation in these families. Breast cancer in the families is still linked to the larger 43.8Mb region on chromosome 6 and further studies are needed to determine possible candidate genes within this large region. In conclusion, we have employed global gene expression profiling of familial tumours in an effort to determine a molecular phenotype that may allow classification of familial non-BRCA1/2 breast tumours. The classification of families into sub-groups allowed us to determine common regions of suggestive linkage in these families and may be a tool for future studies of families believed to carry high- to moderate-penetrant predisposing mutations.

7.4 PAPER IV

Estrogen receptor beta (*ESR2*) polymorphisms in familial and sporadic breast cancer

Estrogen plays a role in normal mammary development and in breast carcinogenesis and its biological effects are mediated mainly through the estrogen receptors; estrogen receptor alpha (ESR1) and estrogen receptor beta (ESR2). These two receptors are members of the nuclear receptor superfamily of ligand inducible transcription factors, which regulate transcription in association with coregulators via binding to DNA enhancer elements (EREs) located within the promoters of target genes. As the search for high to moderate penetrant genes continues, the focus in genetic research has shifted to the localization and identification of low-penetrance genes. These genes may be common in the general population and modestly increase womens risk of breast cancer, however the combined effect of several low-penetrant genes may have a substantial impact on breast cancer risk. Due to their functional roles in breast biology and estrogen metabolism, the estrogen receptors are candidate low-penetrance breast cancer susceptibility genes.

In this study we examined three common polymorphisms in the *estrogen receptor beta* (*ESR1*) gene; rs1256049: a G>A polymorphism at position 1082 in exon 5, rs4986938: a G>A polymorphism at position 1730 in the 3'UTR of exon 8 and rs928554: a G>A polymorphism located 56 bases 3' of the *ESR2* alternative transcript, Cx exon 9. These SNPs were investigated for association with breast cancer risk in 723 breast cancer cases and 480 healthy controls. The breast cancer cases were composed of 323 sporadic cases and 400 familial cases and the 400 familial cases were further divided into 212 familial high-risk cases and 188 familial low-risk cases.

All three SNPs were successfully genotyped by pyrosequencing or RFLP analysis and were determined to be in Hardy-Weinberg equilibrium in cases and controls. The genotype and allele frequencies for each of the polymorphisms were tested for association in each breast cancer group compared to the healthy controls. There was no overall statistically significant difference in the genotype frequencies between cases and controls. However, in the familial low-risk breast cancer cases the

heterozygote genotype of the rs4986938 SNP showed a suggestively protective effect on breast cancer risk (OR = 0.72, 95% CI; 0.50-1.05) whereas the G allele of the rs928554 SNP appeared to associated with a moderately elevated breast cancer risk (OR = 1.26, 95% CI; 0.99-1.61). In the sporadic breast cancer cases the A allele of rs1256049 appeared to be associated with a decrease in breast cancer risk (OR = 0.63, 95% CI; 0.37-1.07). No suggestive associations were seen in the familial high-risk cases.

As no significant associations were found when investigating the three SNPs individually we investigated the extent of linkage disequilibrium (LD) in the region and determined that all 3 SNPs were in strong LD, with D' values ranging from 0.86 to 1.00. Investigation of the possible haplotypes formed by these three SNPs for association with breast cancer revealed one common haplotype G-A-G which was associated with increased risk in the sporadic breast cancer cases (OR = 3.0 p=0.03). The contrary haplotype A-G-A was associated with decreased risk (OR = 0.4 p=0.03), which supports the finding of the individual A allele of rs1256049 having a protective effect. No significant difference in the haplotype frequencies was seen for the familial breast cancer cases.

Examination of the individual variants did not reveal any significant differences in their genotype distribution and as these variants represent a silent SNP and two 3'UTR variants it is difficult to assign them any functional role. A number of studies have also examined different SNPs individually for association to breast cancer and found no overall statistically significant associations (Forsti, *et al*., 2003; Zheng , *et al*., 2003; Gold, *et al*., 2004). However, Zheng *et al*., found that breast cancer risk was increased in women with an *ESR2* risk genotype and high hormone levels (Zheng, *et al.*, 2003). While a study by Gold *et al*., found an association between *ESR2* haplotypes with breast cancer in the Ashkenazi Jewish population (Gold, *et al.*, 2004). These results along with our current finding support a role for the *ESR2* locus in breast carcinogenesis. Due to the strong amount of LD in the region it is possible that the SNPs examined in this study are in LD with an as yet unidentified variant within or close to the *ESR2* gene. Therefore it will be worthwhile to investigate the locus in more detail in an attempt to identify novel breast cancer susceptibility alleles and further understand the complex etiology of breast carcinogenesis.

8 CONCLUSIONS

This thesis work aimed to identify the genetic basis of familial non-BRCA1/2 breast cancer. The main findings of the studies in this thesis conclude that:

The *HIN-1* gene is not responsible for the increased risk associated with the families exhibiting suggestive linkage to the candidate locus on chromosome 5q. In addition, the methylation status of *HIN-1* is significantly different in sporadic and *BRCA1* tumours. However, a sub-set of BRCA1-like sporadic tumours are unmethylated at the *HIN-1* promoter. Although *HIN-1* is unlikely to play a major role in breast cancer predisposition its lack of expression may have a consequence for breast cancer pathogenesis.

Comparative genomic hybridisation of familial non-BRCA1/2 breast tumours identified that the loss of chromosome 6q, 11q22-qter, 17 and gains of chromosome 1q, 8q and 16p were frequent events in these non-BRCA1/2 tumours. In high-risk non-BRCA1/2 breast cancers families loss of chromosome 17 or chromosome 6q were mutually exclusive events. Analysis of a region on chromosome 17q11.2-2 in ten families exhibiting suggestive linkage to the region revealed no pathogenic germline mutations.

Re-evaluation of linkage data two non-BRCA1/2 families determined that breast cancer in both families was linked to chromosome 6. These families were sub-grouped based on tumour profiling. The classification of families into distinct groups can be a useful aid in genetic linkage analysis. These two families shared a common founder haplotype on chromosome 6q14. Mutation analysis of the region identified a number of common variants shared by the two families, none of them obviously pathogenic.

Analysis of polymorphisms in the *ESR2* gene revealed no overall significant differences in genotype of allele distributions. However, as there is strong LD across the *ESR2* locus, haplotype analysis was carried out and it was determined that one common haplotype G-A-G was associated with an increased risk of sporadic breast cancer.

9 ACKNOWLEDGEMENTS

This work was carried out at the Department of Molecular Medicine, Karolinska Institute and I would like to express my sincere gratitude to all those who made this possible especially:

Professor Annika Lindblom, my supervisor, for accepting me as a PhD student in your group. Thank you for sharing your knowledge of cancer genetics, for always being available to answer questions big or small and for allowing me to grow both personally and as a scientist into an independent person. Thank you also for accepting me into your home for my first months in Sweden, it was very generous of you and it's a time I will always remember. Finally thank you for having such a nice lab, filled with good people and great friendships, this has been a truly memorable time for me.

Professor Magnus Nordenskjöld for creating a stimulating scientific atmosphere and for always showing an interest and being so good humoured.

All the members of Annikas lab: **Barbro** for all your help, technical or otherwise, for taking care of us all and always being so cheerful. **Xiao Lei** for being a really great friend, for always answering my questions no matter what you were doing, for many laughs, interesting stories about China and enjoyable trips to the pub! **Tanja** for your invaluable friendship, thanks for always being there for a chat, a laugh, or a cry. For many fun times and great memories! I know I would never have made it without you. **Johanna** thanks for a nice friendship and lots of good times over the years. **Jana**, for being a good friend and my thesis-writing buddy! **Barbara** for being so open and friendly and for nice picnics and barbeques, **Martina** for good work and fun times on the project over the summers and **Quan, Tao, Liping, Helene, Moraima, Sara** and **Berith** for good company.

All the present and former members of CMM especially; **Frederik, Hahn, Ana B, Anna S**, **Tiina, Michela**, **Virpi**, **Micke, Keng-Ling**, **Biying**, **Emma T, Huma**, **Clara**, **Sofia**, **Kazuko**, **Lina**, **Husameldin, Louise**, **Jacqueline, Elisabeth, Eva, Eric, Helena, Peter, Jill, Britt-Marie, Maria, Zheng, Feng Qing, Katja** and **Anya.** Friends from floor 01; **Theo, Jan, Trish Andrea, Lui, Emma F, Lasse, Brita** thank you all for bringing friendship and a positive atmosphere to CMM.

Dr Fabio Sanchez, for always taking time to listen to me, answer my questions and to explain complex things to me.

Dr Günther Weber, thank you for taking the time to be helpful and for organizing journal club.

Professors Anna Weddel, **Agneta Nordenskjöld** and **Catharina Larsson** for always being interested and helpful.

Anki, Maggan, Anna-Lena, Christina, Sigrid & Catarina thanks for all your help and support both technical and personal throughout my stay in Sweden**.**

Thanks to all the administration staff especially **Yvonne Cowan** for always taking care of everything, and **Chritina Bremer, Kerstin Florell, Delphi Post, Gunilla Risberg** and **Majalena Granqvist** for help with questions big or small.

Lennart Helleday, thanks for being so helpful and cheerful all the time. For coming to my rescue on numerous computer-related occasions!

Ingrid Kössler and the staff at **Bröstcancerföreningarnas Riksorganisation-BRO** for your interest and for supplying pink ribbons.

All my friends at **National Centre for Medical Genetics**, Dublin, Ireland, where I first started down the road of medical genetics. Thanks for all your help, friendship and good times.

Dr Sean Ennis, thanks for all your advice both scientifically and about life in general and for always having the time for a chat and a coffee when I made it home.

Ulrika and **Ulf**, thanks for being great friends, hope to see you and the family in Ireland some day!

Declan my first Irish contact here in Sweden, thanks for taking me to the Loft for the first time and for being such a good friend.

Cathy & Dawn – the cup of tea that saved my life! Thanks for being so carefree and fun-loving and for being good friends when I moved to Jägargatan.

Hanna Karlsson, thanks for being a good friend, for many nice chats and nights out and for memorable ice-skating trips to Kungstagården, hopefully we will see you in Dublin soon.

Emma and **Per-Axel** my roommates from Jägargatan and Bagarmossen! Thanks for being good roommates and friends.

Gail, well what can I say, we made it through Jägargatan, Yoga, early morning swims, Swedish winters and finally our thesis and still managed to have a laugh! So thanks for always being there for girly chats and moans and I look forward to plenty more, back in Dublin!

Dragan for being a great friend, you and Tanja are like my family here and I am going to miss you both, I'm sure we'll meet in Ireland or Belgrade some day!

The "Irish" gang: **Paula C, Maura, Pontus, Oonagh, Daniel, Jim, Jay, Sara-Jayne, Padraig, Mary, Máire, Helen, Steve, Charlotte, Joachim** for being good friends and for many enjoyable nights in various Irish pubs all over Stockholm.

The "International" gang: **Anthony, Kristjana, Mark W, Brad, Jenny & Mark I, Þórgunnur & Frikki.** Thanks for being great friends, for memorable parties and a great trip to Iceland.

My old roommates, from No.32, **Keith No.1** and **Keith No.1**, **France** and **Ad** thanks for great friendship, great fun, lots of laughs and making it so hard to leave Dublin when I did.

My friends at home, **Sinéad & Ian, Allison & John, Jacqui, Megan and Valerie**, for being my friends all these years and for allowing our friendship to stay strong even when I've been far away.

My family, **Mum & Dad** for love and support and encouragement throughout the years and for always thinking of me, without you both, this would not have been possible. My brother and his family **Kevin, Karyn & Jack** for coming to visit during my first winter here and for lots of chats, advice and encouragement. My fabulous sisters, **Teresa, Geraldine & Anne** for always having the time to listen to me, for being yourselves and for being so much fun.

Elsku **Gummi**, þú ert minn einasti eini! Thank you for being my best friend, for listening to me and encouraging me, for always keeping me laughing and most of all for loving me.

10 REFERENCES

Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, Hodges AK, Davies DR, David SS, Sampson JR and others. **2002**. Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. Nat Genet 30(2):227-232.

Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. **2005**. A haplotype map of the human genome. Nature 437(7063):1299-1320.

Anisimov VN, Ukraintseva SV, Yashin AI. **2005**. Cancer in rodents: does it tell us about cancer in humans? Nat Rev Cancer 5(10):807-819.

Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A and others. **2003**. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet 72(5):1117-1130.

Ashworth A. **2005**. Oh what a tangled web it weaves: BRCA1 and DNA decatenation. Cancer Cell 8(2):95-97.

Bachinski LL, Olufemi SE, Zhou X, Wu CC, Yip L, Shete S, Lozano G, Amos CI, Strong LC, Krahe R. **2005**. Genetic mapping of a third Li-Fraumeni syndrome predisposition locus to human chromosome 1q23. Cancer Res 65(2):427-431.

Baker SJ, Fearon ER, Nigro JM, Hamilton SR, Preisinger AC, Jessup JM, vanTuinen P, Ledbetter DH, Barker DF, Nakamura Y and others. **1989**. Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. Science 244(4901):217-221.

Baker SJ, Preisinger AC, Jessup JM, Paraskeva C, Markowitz S, Willson JK, Hamilton S, Vogelstein B. **1990**. p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. Cancer Res 50(23):7717-7722.

Bardin A, Boulle N, Lazennec G, Vignon F, Pujol P. **2004**. Loss of ERbeta expression as a common step in estrogen-dependent tumor progression. Endocr Relat Cancer 11(3):537-551.

Barrett JC, Fry B, Maller J, Daly MJ. **2005**. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21(2):263-265.

Balmain A, Gray J, Ponder B. **2003**. The genetics and genomics of cancer. Nat Genet 33 Suppl:238-244.

Bell DW, Varley JM, Szydlo TE, Kang DH, Wahrer DC, Shannon KE, Lubratovich M, Verselis SJ, Isselbacher KJ, Fraumeni JF and others. **1999**. Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. Science 286(5449):2528-2531.

Berger MS, Locher GW, Saurer S, Gullick WJ, Waterfield MD, Groner B, Hynes NE. **1988**. Correlation of c-erbB-2 gene amplification and protein expression in human breast carcinoma with nodal status and nuclear grading. Cancer Res 48(5):1238-1243.

Bhatia S, Robison LL, Oberlin O, Greenberg M, Bunin G, Fossati-Bellani F, Meadows AT. **1996**. Breast cancer and other second neoplasms after childhood Hodgkin's disease. N Engl J Med 334(12):745-751.

Bignell GR, Barfoot R, Seal S, Collins N, Warren W, Stratton MR. **1998**. Low frequency of somatic mutations in the LKB1/Peutz-Jeghers syndrome gene in sporadic breast cancer. Cancer Res 58(7):1384-1386.

Bochar DA, Wang L, Beniya H, Kinev A, Xue Y, Lane WS, Wang W, Kashanchi F, Shiekhattar R. **2000**. BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. Cell 102(2):257-265.

Bodmer WF, Bailey CJ, Bodmer J, Bussey HJ, Ellis A, Gorman P, Lucibello FC, Murday VA, Rider SH, Scambler P and others. **1987**. Localization of the gene for familial adenomatous polyposis on chromosome 5. Nature 328(6131):614-616.

Botstein D, Risch N. **2003**. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat Genet 33 Suppl:228-237.

Bowlin SJ, Leske MC, Varma A, Nasca P, Weinstein A, Caplan L. **1997**. Breast cancer risk and alcohol consumption: results from a large case-control study. Int J Epidemiol 26(5):915-923.

Brose MS, Rebbeck TR, Calzone KA, Stopfer JE, Nathanson KL, Weber BL. **2002**. Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program. J Natl Cancer Inst 94(18):1365-1372.

Cannon-Albright LA, Goldgar DE, Meyer LJ, Lewis CM, Anderson DE, Fountain JW, Hegi ME, Wiseman RW, Petty EM, Bale AE and others. **1992**. Assignment of a locus for familial melanoma, MLM, to chromosome 9p13-p22. Science 258(5085):1148-1152.

Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, Wahrer DC, Sgroi DC, Lane WS, Haber DA and others. **2001**. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. Cell 105(1):149-160.

Capon DJ, Seeburg PH, McGrath JP, Hayflick JS, Edman U, Levinson AD, Goeddel DV. **1983**. Activation of Ki-ras2 gene in human colon and lung carcinomas by two different point mutations. Nature 304(5926):507-513.

Catteau A, Harris WH, Xu CF, Solomon E. **1999**. Methylation of the BRCA1 promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics. Oncogene 18(11):1957-1965.

CHEK2 Breast Cancer Case Control Consortium. **2004**. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. Am J Hum Genet 74(6):1175-1182.

Chen J, Lindblom A. **2000**. Germline mutation screening of the STK11/LKB1 gene in familial breast cancer with LOH on 19p. Clin Genet 57(5):394-397.

Chen J, Lindblom P, Lindblom A. **1998**. A study of the PTEN/MMAC1 gene in 136 breast cancer families. Hum Genet 102(1):124-125.

Claus EB, Schildkraut JM, Thompson WD, Risch NJ. **1996**. The genetic attributable risk of breast and ovarian cancer. Cancer 77(11):2318-2324.

Colditz GA. **2005**. Epidemiology and prevention of breast cancer. Cancer Epidemiol Biomarkers Prev 14(4):768-772.

Colditz GA, Hankinson SE, Hunter DJ, Willett WC, Manson JE, Stampfer MJ, Hennekens C, Rosner B, Speizer FE. **1995**. The use of estrogens and progestins and the risk of breast cancer in postmenopausal women. N Engl J Med 332(24):1589-1593.

Collaborative Group on Hormonal Factors in Breast Cancer. **1996**. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Collaborative Group on Hormonal Factors in Breast Cancer. Lancet 347(9017):1713-1727.

Collaborative Group on Hormonal Factors in Breast Cancer. **1997**. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer. Lancet 350(9084):1047-1059.

Collaborative Group on Hormonal Factors in Breast Cancer. **2001**. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet 358(9291):1389-1399.

Collins N, McManus R, Wooster R, Mangion J, Seal S, Lakhani SR, Ormiston W, Daly PA, Ford D, Easton DF and others. **1995**. Consistent loss of the wild type allele in breast cancers from a family linked to the BRCA2 gene on chromosome 13q12-13. Oncogene 10(8):1673-1675.

Collins N, Wooster R, Stratton MR. **1997**. Absence of methylation of CpG dinucleotides within the promoter of the breast cancer susceptibility gene BRCA2 in normal tissues and in breast and ovarian cancers. Br J Cancer 76(9):1150-1156.

Cordell HJ, Clayton DG. **2005**. Genetic association studies. Lancet 366(9491):1121-1131.

Cowley SM, Hoare S, Mosselman S, Parker MG. **1997**. Estrogen receptors alpha and beta form heterodimers on DNA. J Biol Chem 272(32):19858-19862.

Cowley SM, Parker MG. **1999**. A comparison of transcriptional activation by ER alpha and ER beta. J Steroid Biochem Mol Biol 69(1-6):165-175.

Cui J, Antoniou AC, Dite GS, Southey MC, Venter DJ, Easton DF, Giles GG, McCredie MR, Hopper JL. **2001**. After BRCA1 and BRCA2-what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer. Am J Hum Genet 68(2):420-431.

Daniels MJ, Wang Y, Lee M, Venkitaraman AR. **2004**. Abnormal cytokinesis in cells deficient in the breast cancer susceptibility protein BRCA2. Science 306(5697):876-879.

de Jong MM, Nolte IM, te Meerman GJ, van der Graaf WT, Oosterwijk JC, Kleibeuker JH, Schaapveld M, de Vries EG. **2002**. Genes other than BRCA1 and BRCA2 involved in breast cancer susceptibility. J Med Genet 39(4):225-242.

de Klein A, van Kessel AG, Grosveld G, Bartram CR, Hagemeijer A, Bootsma D, Spurr NK, Heisterkamp N, Groffen J, Stephenson JR. **1982**. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. Nature 300(5894):765-767.

Deng CX, Brodie SG. **2001**. Knockout mouse models and mammary tumorigenesis. Semin Cancer Biol 11(5):387-394.

Dietrich WF, Lander ES, Smith JS, Moser AR, Gould KA, Luongo C, Borenstein N, Dove W. **1993**. Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. Cell 75(4):631-639.

Dobrovic A, Simpfendorfer D. **1997**. Methylation of the BRCA1 gene in sporadic breast cancer. Cancer Res 57(16):3347-3350.

Du Q, Luo L, von Wachenfeldt A, Kockum I, Luthman H, Lindblom A. **2002**. No evidence for a familial breast cancer susceptibility gene at chromosome 13q21 in Swedish breast cancer families. Int J Cancer 98(5):799-800.

Dudbridge F. **2003**. Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 25(2):115-121.

Dunning AM, Healey CS, Pharoah PD, Teare MD, Ponder BA, Easton DF. **1999**. A systematic review of genetic polymorphisms and breast cancer risk. Cancer Epidemiol Biomarkers Prev 8(10):843-854.

Ellis NA, Kirchhoff T, Mitra N, Ye TZ, Chuai S, Huang H, Nafa K, Norton L, Neuhausen S, Gordon D and others. **2005**. Localization of breast cancer susceptibility loci by genome-wide SNP linkage disequilibrium mapping. Genet Epidemiol.

Eng C. **2003**. PTEN: one gene, many syndromes. Hum Mutat 22(3):183-198.

Esteller M, Silva JM, Dominguez G, Bonilla F, Matias-Guiu X, Lerma E, Bussaglia E, Prat J, Harkes IC, Repasky EA and others. **2000**. Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. J Natl Cancer Inst 92(7):564-569.

Fearon ER, Vogelstein B. **1990**. A genetic model for colorectal tumorigenesis. Cell 61(5):759-767.

Feilotter HE, Coulon V, McVeigh JL, Boag AH, Dorion-Bonnet F, Duboue B, Latham WC, Eng C, Mulligan LM, Longy M. **1999**. Analysis of the 10q23 chromosomal region and the PTEN gene in human sporadic breast carcinoma. Br J Cancer 79(5-6):718-723.

Feinberg AP, Vogelstein B. **1983**. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 301(5895):89-92.

Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J and others. **1998**. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet 62(3):676-689.

Forsti A, Zhao C, Israelsson E, Dahlman-Wright K, Gustafsson JA, Hemminki K. **2003**. Polymorphisms in the estrogen receptor beta gene and risk of breast cancer: no association. Breast Cancer Res Treat 79(3):409-413.

Freihoff D, Kempe A, Beste B, Wappenschmidt B, Kreyer E, Hayashi Y, Meindl A, Krebs D, Wiestler OD, von Deimling A and others. **1999**. Exclusion of a major

role for the PTEN tumour-suppressor gene in breast carcinomas. Br J Cancer 79(5-6):754-758.

Friedberg EC. **2001**. How nucleotide excision repair protects against cancer. Nat Rev Cancer 1(1):22-33.

Friend SH, Horowitz JM, Gerber MR, Wang XF, Bogenmann E, Li FP, Weinberg RA. **1987**. Deletions of a DNA sequence in retinoblastomas and mesenchymal tumors: organization of the sequence and its encoded protein. Proc Natl Acad Sci U S A 84(24):9059-9063.

Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki Y and others. **1994**. BRCA1 mutations in primary breast and ovarian carcinomas. Science 266(5182):120-122.

Futreal PA, Soderkvist P, Marks JR, Iglehart JD, Cochran C, Barrett JC, Wiseman RW. **1992**. Detection of frequent allelic loss on proximal chromosome 17q in sporadic breast carcinoma using microsatellite length polymorphisms. Cancer Res 52(9):2624-2627.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M and others. **2002**. The structure of haplotype blocks in the human genome. Science 296(5576):2225-2229.

Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ. **1989**. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 81(24):1879-1886.

Ganesan S, Silver DP, Greenberg RA, Avni D, Drapkin R, Miron A, Mok SC, Randrianarison V, Brodie S, Salstrom J and others. **2002**. BRCA1 supports XIST RNA concentration on the inactive X chromosome. Cell 111(3):393-405.

Garfinkel L, Boffetta P, Stellman SD. **1988**. Alcohol and breast cancer: a cohort study. Prev Med 17(6):686-693.

Garrett PA, Hulka BS, Kim YL, Farber RA. **1993**. HRAS protooncogene polymorphism and breast cancer. Cancer Epidemiol Biomarkers Prev 2(2):131-138.

Gayther SA, Warren W, Mazoyer S, Russell PA, Harrington PA, Chiano M, Seal S, Hamoudi R, van Rensburg EJ, Dunning AM and others. **1995**. Germline mutations of the BRCA1 gene in breast and ovarian cancer families provide evidence for a genotype-phenotype correlation. Nat Genet 11(4):428-433.

Gold B, Kalush F, Bergeron J, Scott K, Mitra N, Wilson K, Ellis N, Huang H, Chen M, Lippert R and others. **2004**. Estrogen receptor genotypes and haplotypes associated with breast cancer risk. Cancer Res 64(24):8891-8900.

Gosse-Brun S, Sauvaigo S, Daver A, Page M, Lortholary A, Larra F, Bignon YJ, Bernard-Gallon D. **1999**. Specific H-Ras minisatellite alleles in breast cancer susceptibility. Anticancer Res 19(6B):5191-5196.

Greenblatt MS, Bennett WP, Hollstein M, Harris CC. **1994**. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res 54(18):4855-4878.

Greene GL, Gilna P, Waterfield M, Baker A, Hort Y, Shine J. **1986**. Sequence and expression of human estrogen receptor complementary DNA. Science 231(4742):1150-1154.

Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King MC. **1990**. Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250(4988):1684-1689.

Hanahan D, Weinberg RA. **2000**. The hallmarks of cancer. Cell 100(1):57-70.

Hanstein B, Djahansouzi S, Dall P, Beckmann MW, Bender HG. **2004**. Insights into the molecular biology of the estrogen receptor define novel therapeutic targets for breast cancer. Eur J Endocrinol 150(3):243-255.

Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP and others. **2001**. Gene-expression profiles in hereditary breast cancer. N Engl J Med 344(8):539-548.

Hedenfalk I, Ringner M, Ben-Dor A, Yakhini Z, Chen Y, Chebil G, Ach R, Loman N, Olsson H, Meltzer P and others. **2003**. Molecular classification of familial non-BRCA1/BRCA2 breast cancer. Proc Natl Acad Sci U S A 100(5):2532-2537.

Hemminki A, Markie D, Tomlinson I, Avizienyte E, Roth S, Loukola A, Bignell G, Warren W, Aminoff M, Hoglund P and others. **1998**. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. Nature 391(6663):184-187.

Hemminki A, Tomlinson I, Markie D, Jarvinen H, Sistonen P, Bjorkqvist AM, Knuutila S, Salovaara R, Bodmer W, Shibata D and others. **1997**. Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. Nat Genet 15(1):87-90.

Herman JG, Latif F, Weng Y, Lerman MI, Zbar B, Liu S, Samid D, Duan DS, Gnarra JR, Linehan WM and others. **1994**. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc Natl Acad Sci U S A 91(21):9700-9704.

Hilakivi-Clarke L. **2000**. Estrogens, BRCA1, and breast cancer. Cancer Res 60(18):4993-5001.

Huebner RJ, Todaro GJ. **1969**. Oncogenes of RNA tumor viruses as determinants of cancer. Proc Natl Acad Sci U S A 64(3):1087-1094.

Hulka BS, Moorman PG. **2001**. Breast cancer: hormones and other risk factors. Maturitas 38(1):103-113; discussion 113-106.

Huusko P, Juo SH, Gillanders E, Sarantaus L, Kainu T, Vahteristo P, Allinen M, Jones M, Rapakko K, Eerola H and others. **2004**. Genome-wide scanning for linkage in Finnish breast cancer families. Eur J Hum Genet 12(2):98-104.

Jonsson G, Naylor TL, Vallon-Christersson J, Staaf J, Huang J, Ward MR, Greshock JD, Luts L, Olsson H, Rahman N and others. **2005**. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res 65(17):7612-7621.

Kainu T, Juo SH, Desper R, Schaffer AA, Gillanders E, Rozenblum E, Freas-Lutz D, Weaver D, Stephan D, Bailey-Wilson J and others. **2000**. Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. Proc Natl Acad Sci U S A 97(17):9603-9608.

Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D. **1992**. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258(5083):818-821.

Kallioniemi OP, Kallioniemi A, Piper J, Isola J, Waldman FM, Gray JW, Pinkel D. **1994**. Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors. Genes Chromosomes Cancer 10(4):231-243.

Kerangueven F, Essioux L, Dib A, Noguchi T, Allione F, Geneix J, Longy M, Lidereau R, Eisinger F, Pebusque MJ and others. **1995**. Loss of heterozygosity and linkage analysis in breast carcinoma: indication for a putative third susceptibility gene on the short arm of chromosome 8. Oncogene 10(5):1023-1026.

Knudson AG, Jr. **1971**. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 68(4):820-823.

Krontiris TG, Devlin B, Karp DD, Robert NJ, Risch N. **1993**. An association between the risk of cancer and mutations in the HRAS1 minisatellite locus. N Engl J Med 329(8):517-523.

Krop I, Parker MT, Bloushtain-Qimron N, Porter D, Gelman R, Sasaki H, Maurer M, Terry MB, Parsons R, Polyak K. **2005**. HIN-1, an inhibitor of cell growth, invasion, and AKT activation. Cancer Res 65(21):9659-9669.

Krop IE, Sgroi D, Porter DA, Lunetta KL, LeVangie R, Seth P, Kaelin CM, Rhei E, Bosenberg M, Schnitt S and others. **2001**. HIN-1, a putative cytokine highly expressed in normal but not cancerous mammary epithelial cells. Proc Natl Acad Sci U S A 98(17):9796-9801.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. **1996**. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58(6):1347-1363.

Lancaster JM, Wooster R, Mangion J, Phelan CM, Cochran C, Gumbs C, Seal S, Barfoot R, Collins N, Bignell G and others. **1996**. BRCA2 mutations in primary breast and ovarian cancers. Nat Genet 13(2):238-240.

Lander E, Kruglyak L. **1995**. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11(3):241-247.

Lasko D, Cavenee W, Nordenskjold M. **1991**. Loss of constitutional heterozygosity in human cancer. Annu Rev Genet 25:281-314.

Leygue E, Dotzlaw H, Watson PH, Murphy LC. **1998**. Altered estrogen receptor alpha and beta messenger RNA expression during human breast tumorigenesis. Cancer Res 58(15):3197-3201.

Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, Puc J, Miliaresis C, Rodgers L, McCombie R and others. **1997**. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science 275(5308):1943-1947.

Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K. **2000**. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 343(2):78-85.

Liehr JG. **2000**. Is estradiol a genotoxic mutagenic carcinogen? Endocr Rev 21(1):40-54.

Lindblom A, Skoog L, Rotstein S, Werelius B, Larsson C, Nordenskjold M. **1993a**. Loss of heterozygosity in familial breast carcinomas. Cancer Res 53(18):4356-4361.

Lindblom A, Tannergard P, Werelius B, Nordenskjold M. **1993b**. Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. Nat Genet 5(3):279-282.

Lou Z, Minter-Dykhouse K, Chen J. **2005**. BRCA1 participates in DNA decatenation. Nat Struct Mol Biol 12(7):589-593.

Ludwig T, Fisher P, Murty V, Efstratiadis A. **2001**. Development of mammary adenocarcinomas by tissue-specific knockout of Brca2 in mice. Oncogene 20(30):3937-3948.

Ma YX, Tomita Y, Fan S, Wu K, Tong Y, Zhao Z, Song LN, Goldberg ID, Rosen EM. **2005**. Structural determinants of the BRCA1 : estrogen receptor interaction. Oncogene 24(11):1831-1846.

Mallery DL, Vandenberg CJ, Hiom K. **2002**. Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains. Embo J 21(24):6755-6762.

Marchetti A, Barassi F, Martella C, Chella A, Salvatore S, Castrataro A, Mucilli F, Sacco R, Buttitta F. **2004**. Down regulation of high in normal-1 (HIN-1) is a frequent event in stage I non-small cell lung cancer and correlates with poor clinical outcome. Clin Cancer Res 10(4):1338-1343.

Margolin S, Werelius B, Fornander T, Lindblom A. **2004**. BRCA1 mutations in a population-based study of breast cancer in Stockholm County. Genet Test 8(2):127-132.

Martin AM, Weber BL. **2000**. Genetic and hormonal risk factors in breast cancer. J Natl Cancer Inst 92(14):1126-1135.

Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, Hollestelle A, Houben M, Crepin E, van Veghel-Plandsoen M and others. **2002**. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet 31(1):55-59.

Merlo A, Herman JG, Mao L, Lee DJ, Gabrielson E, Burger PC, Baylin SB, Sidransky D. **1995**. 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. Nat Med 1(7):686-692.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W and others. **1994**. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266(5182):66-71.

Miyaki M, Konishi M, Tanaka K, Kikuchi-Yanoshita R, Muraoka M, Yasuno M, Igari T, Koike M, Chiba M, Mori T. **1997**. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. Nat Genet 17(3):271-272.

Morris JR, Solomon E. **2004**. BRCA1 : BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. Hum Mol Genet 13(8):807-817.

Morton NE. **1955**. Sequential tests for the detection of linkage. Am J Hum Genet 7(3):277-318.

Mosselman S, Polman J, Dijkema R. **1996**. ER beta: identification and characterization of a novel human estrogen receptor. FEBS Lett 392(1):49-53.

Motegi T, Komatsu M, Nakazato Y, Ohuchi M, Minoda K. **1982**. Retinoblastoma in a boy with a de novo mutation of a 13/18 translocation: the assumption that the retinoblastoma locus is at 13q141, particularly at the distal portion of it. Hum Genet 60(2):193-195.

Moynahan ME. **2002**. The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans. Oncogene 21(58):8994-9007.

Moynahan ME, Chiu JW, Koller BH, Jasin M. **1999**. Brca1 controls homology-directed DNA repair. Mol Cell 4(4):511-518.

Nathanson KL, Shugart YY, Omaruddin R, Szabo C, Goldgar D, Rebbeck TR, Weber BL. **2002**. CGH-targeted linkage analysis reveals a possible BRCA1 modifier locus on chromosome 5q. Hum Mol Genet 11(11):1327-1332.

Nathanson KL, Wooster R, Weber BL. **2001**. Breast cancer genetics: what we know and what we need. Nat Med 7(5):552-556.

Ogawa S, Inoue S, Watanabe T, Orimo A, Hosoi T, Ouchi Y, Muramatsu M. **1998**. Molecular cloning and characterization of human estrogen receptor betacx: a potential inhibitor ofestrogen action in human. Nucleic Acids Res 26(15):3505-3512.

Parkin DM, Bray F, Ferlay J, Pisani P. **2005**. Global cancer statistics, 2002. CA Cancer J Clin 55(2):74-108.

Patel KJ, Yu VP, Lee H, Corcoran A, Thistlethwaite FC, Evans MJ, Colledge WH, Friedman LS, Ponder BA, Venkitaraman AR. **1998**. Involvement of Brca2 in DNA repair. Mol Cell 1(3):347-357.

Patel UA, Perry M, Crane-Robinson C. **1995**. Screening for germline mutations of the p53 gene in familial breast cancer patients. Eur J Clin Invest 25(2):132-137.

Paull TT, Rogakou EP, Yamazaki V, Kirchgessner CU, Gellert M, Bonner WM. **2000**. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. Curr Biol 10(15):886-895.

Peltomaki P, Aaltonen LA, Sistonen P, Pylkkanen L, Mecklin JP, Jarvinen H, Green JS, Jass JR, Weber JL, Leach FS and others. **1993**. Genetic mapping of a locus predisposing to human colorectal cancer. Science 260(5109):810-812.

Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. **2002**. Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31(1):33-36.

Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. **1999**. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 23(1):41-46.

Pujol P, Galtier-Dereure F, Bringer J. **1997**. Obesity and breast cancer risk. Hum Reprod 12 Suppl 1:116-125.

Rahman N, Teare MD, Seal S, Renard H, Mangion J, Cour C, Thompson D, Shugart Y, Eccles D, Devilee P and others. **2000**. Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. Oncogene 19(36):4170-4173.

Rapley EA, Crockford GP, Teare D, Biggs P, Seal S, Barfoot R, Edwards S, Hamoudi R, Heimdal K, Fossa SD and others. **2000**. Localization to Xq27 of a susceptibility gene for testicular germ-cell tumours. Nat Genet 24(2):197-200.

Rice JC, Ozcelik H, Maxeiner P, Andrulis I, Futscher BW. **2000**. Methylation of the BRCA1 promoter is associated with decreased BRCA1 mRNA levels in clinical breast cancer specimens. Carcinogenesis 21(9):1761-1765.

Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P. **1996**. Real-time DNA sequencing using detection of pyrophosphate release. Anal Biochem 242(1):84-89.

Russell PA, Pharoah PD, De Foy K, Ramus SJ, Symmonds I, Wilson A, Scott I, Ponder BA, Gayther SA. **2000**. Frequent loss of BRCA1 mRNA and protein expression in sporadic ovarian cancers. Int J Cancer 87(3):317-321.

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL and others. **2001**. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409(6822):928-933.

Sakai T, Toguchida J, Ohtani N, Yandell DW, Rapaport JM, Dryja TP. **1991**. Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. Am J Hum Genet 48(5):880-888.

Sanger F, Nicklen S, Coulson AR. **1977**. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74(12):5463-5467.

Savitsky K, Bar-Shira A, Gilad S, Rotman G, Ziv Y, Vanagaite L, Tagle DA, Smith S, Uziel T, Sfez S and others. **1995**. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. Science 268(5218):1749-1753.

Scully R, Chen J, Plug A, Xiao Y, Weaver D, Feunteun J, Ashley T, Livingston DM. **1997**. Association of BRCA1 with Rad51 in mitotic and meiotic cells. Cell 88(2):265-275.

Seitz S, Rohde K, Bender E, Nothnagel A, Pidde H, Ullrich OM, El-Zehairy A, Haensch W, Jandrig B, Kolble K and others. **1997**. Deletion mapping and linkage analysis provide strong indication for the involvement of the human chromosome region 8p12-p22 in breast carcinogenesis. Br J Cancer 76(8):983-991.

Shen CY, Yu JC, Lo YL, Kuo CH, Yue CT, Jou YS, Huang CS, Lung JC, Wu CW. **2000**. Genome-wide search for loss of heterozygosity using laser capture microdissected tissue of breast carcinoma: an implication for mutator phenotype and breast cancer pathogenesis. Cancer Res 60(14):3884-3892.

Shen SX, Weaver Z, Xu X, Li C, Weinstein M, Chen L, Guan XY, Ried T, Deng CX. **1998**. A targeted disruption of the murine Brca1 gene causes gamma-irradiation hypersensitivity and genetic instability. Oncogene 17(24):3115-3124.

Shigematsu H, Suzuki M, Takahashi T, Miyajima K, Toyooka S, Shivapurkar N, Tomlinson GE, Mastrangelo D, Pass HI, Brambilla E and others. **2005**. Aberrant methylation of HIN-1 (high in normal-1) is a frequent event in many human malignancies. Int J Cancer 113(4):600-604.

Smith SA, Easton DF, Evans DG, Ponder BA. **1992**. Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. Nat Genet 2(2):128-131.

Sobel E, Lange K. **1996**. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 58(6):1323-1337.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS and others. **2001**. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98(19):10869-10874.

Speirs V, Malone C, Walton DS, Kerin MJ, Atkin SL. **1999**. Increased expression of estrogen receptor beta mRNA in tamoxifen-resistant breast cancer patients. Cancer Res 59(21):5421-5424.

Spence J, Sadis S, Haas AL, Finley D. **1995**. A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. Mol Cell Biol 15(3):1265-1273.

Sukumar S, Notario V, Martin-Zanca D, Barbacid M. **1983**. Induction of mammary carcinomas in rats by nitroso-methylurea involves malignant activation of H-ras-1 locus by single point mutations. Nature 306(5944):658-661.

Takeya T, Hanafusa H. **1983**. Structure and sequence of the cellular gene homologous to the RSV src gene and the mechanism for generating the transforming virus. Cell 32(3):881-890.

Tamimi RM, Hankinson SE, Ding S, Gagalang V, Larson GP, Spiegelman D, Colditz GA, Krontiris TG, Hunter DJ. **2003**. The HRAS1 variable number of tandem repeats and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 12(12):1528-1530.

Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, Aaronson S, Leder P. **1982**. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc Natl Acad Sci U S A 79(24):7837-7841.

Tavtigian SV, Simard J, Rommens J, Couch F, Shattuck-Eidens D, Neuhausen S, Merajver S, Thorlacius S, Offit K, Stoppa-Lyonnet D and others. **1996**. The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. Nat Genet 12(3):333-337.

Thompson D, Duedal S, Kirner J, McGuffog L, Last J, Reiman A, Byrd P, Taylor M, Easton DF. **2005**. Cancer risks and mortality in heterozygous ATM mutation carriers. J Natl Cancer Inst 97(11):813-822.

Thompson D, Easton D. **2001**. Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. Am J Hum Genet 68(2):410-419.

Thompson D, Easton D. **2002a**. Variation in BRCA1 cancer risks by mutation position. Cancer Epidemiol Biomarkers Prev 11(4):329-336.

Thompson D, Easton DF. **2002b**. Cancer Incidence in BRCA1 mutation carriers. J Natl Cancer Inst 94(18):1358-1365.

Thompson D, Szabo CI, Mangion J, Oldenburg RA, Odefrey F, Seal S, Barfoot R, Kroeze-Jansema K, Teare D, Rahman N and others. **2002**. Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. Proc Natl Acad Sci U S A 99(2):827-831.

Thompson ME, Jensen RA, Obermiller PS, Page DL, Holt JT. **1995**. Decreased expression of BRCA1 accelerates growth and is often present during sporadic breast cancer progression. Nat Genet 9(4):444-450.

Tibes R, Trent J, Kurzrock R. **2005**. Tyrosine kinase inhibitors and the dawn of molecular cancer therapeutics. Annu Rev Pharmacol Toxicol 45:357-384.

Tirkkonen M, Johannsson O, Agnarsson BA, Olsson H, Ingvarsson S, Karhu R, Tanner M, Isola J, Barkardottir RB, Borg A and others. **1997**. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res 57(7):1222-1227.

Tokunaga M, Land CE, Tokuoka S, Nishimori I, Soda M, Akiba S. **1994**. Incidence of female breast cancer among atomic bomb survivors, 1950-1985. Radiat Res 138(2):209-223.

Trichopoulos D, MacMahon B, Cole P. **1972**. Menopause and breast cancer risk. J Natl Cancer Inst 48(3):605-613.

Tutt A, Bertwistle D, Valentine J, Gabriel A, Swift S, Ross G, Griffin C, Thacker J, Ashworth A. **2001**. Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. Embo J 20(17):4704-4716.

Tutt A, Gabriel A, Bertwistle D, Connor F, Paterson H, Peacock J, Ross G, Ashworth A. **1999**. Absence of Brca2 causes genome instability by chromosome breakage and loss associated with centrosome amplification. Curr Biol 9(19):1107-1110.

Vahteristo P, Bartkova J, Eerola H, Syrjakoski K, Ojala S, Kilpivaara O, Tamminen A, Kononen J, Aittomaki K, Heikkila P and others. **2002**. A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. Am J Hum Genet 71(2):432-438.

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT and others. **2002**. Gene expression profiling predicts clinical outcome of breast cancer. Nature 415(6871):530-536.

van Beers EH, van Welsem T, Wessels LF, Li Y, Oldenburg RA, Devilee P, Cornelisse CJ, Verhoef S, Hogervorst FB, van't Veer LJ and others. **2005**. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res 65(3):822-827.

Venkitaraman AR. **2002**. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 108(2):171-182.

Vogelstein B, Kinzler KW. **1993**. The multistep nature of cancer. Trends Genet 9(4):138-141.

Wang C, Deng L, Hong M, Akkaraju GR, Inoue J, Chen ZJ. **2001**. TAK1 is a ubiquitin-dependent kinase of MKK and IKK. Nature 412(6844):346-351.

Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J. **2000**. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. Genes Dev 14(8):927-939.

Wessels LF, van Welsem T, Hart AA, van't Veer LJ, Reinders MJ, Nederlof PM. **2002**. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 62(23):7110-7117.

Wilson CA, Ramos L, Villasenor MR, Anders KH, Press MF, Clarke K, Karlan B, Chen JJ, Scully R, Livingston D and others. **1999**. Localization of human BRCA1 and its loss in high-grade, non-inherited breast carcinomas. Nat Genet 21(2):236-240.

Wong TS, Kwong DL, Sham JS, Tsao SW, Wei WI, Kwong YL, Yuen AP. **2003**. Promoter hypermethylation of high-in-normal 1 gene in primary nasopharyngeal carcinoma. Clin Cancer Res 9(8):3042-3046.

Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G. **1995**. Identification of the breast cancer susceptibility gene BRCA2. Nature 378(6559):789-792.

Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D and others. **1994**. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science 265(5181):2088-2090.

Xu B, Kim S, Kastan MB. **2001**. Involvement of Brca1 in S-phase and G(2)-phase checkpoints after ionizing irradiation. Mol Cell Biol 21(10):3445-3450.

Xu X, Wagner KU, Larson D, Weaver Z, Li C, Ried T, Hennighausen L, Wynshaw-Boris A, Deng CX. **1999a**. Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. Nat Genet 22(1):37-43.

Xu X, Weaver Z, Linke SP, Li C, Gotay J, Wang XW, Harris CC, Ried T, Deng CX. **1999b**. Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. Mol Cell 3(3):389-395.

Yuasa Y, Srivastava SK, Dunn CY, Rhim JS, Reddy EP, Aaronson SA. **1983**. Acquisition of transforming properties by alternative point mutations within c-bas/has human proto-oncogene. Nature 303(5920):775-779.

Zelada-Hedman M, Borresen-Dale AL, Claro A, Chen J, Skoog L, Lindblom A. **1997**. Screening for TP53 mutations in patients and tumours from 109 Swedish breast cancer families. Br J Cancer 75(8):1201-1204.

Zelada-Hedman M, Torroella M, Mesquita R, Nordenskjold M, Skoog L, Lindblom A. **1994**. Loss of heterozygosity studies in tumors from families with breast-ovarian cancer syndrome. Hum Genet 94(3):231-234.

Zhang SM, Willett WC, Selhub J, Hunter DJ, Giovannucci EL, Holmes MD, Colditz GA, Hankinson SE. **2003**. Plasma folate, vitamin B6, vitamin B12, homocysteine, and risk of breast cancer. J Natl Cancer Inst 95(5):373-380.

Zheng SL, Zheng W, Chang BL, Shu XO, Cai Q, Yu H, Dai Q, Xu J, Gao YT. **2003**. Joint effect of estrogen receptor beta sequence variants and endogenous estrogen exposure on breast cancer risk in Chinese women. Cancer Res 63(22):7624-7629.