REVIEW ARTICLE

RECENT ADVANCES IN THE GENETICS OF HERITABLE BREAST CANCER

D. E. PORTER AND C. M. STEEL
Breast Unit, Western General Hospital, Edinburgh

KEY WORDS Breast cancer Genetics p53 BRCA1 Linkage

INTRODUCTION

There have been a number of recent advances in the genetics of heritable breast cancer. These are already beginning to influence risk assessment in individuals who are considered potentially susceptible to malignancy on the basis of their family history. It is anticipated that continued progress in this field will eventually lead to the identification of a small, very high risk group of breast cancer gene carriers. Diagnostic organ screening, surgical, and pharmacological prophylaxis will be more cost-effective in this well-defined and not insignificant subpopulation.

The early findings of Anderson (1974, 1977) and Lynch et al. (1976, 1979) that first degree relatives of breast cancer patients were at particularly high risk if the index case had either malignancy at young age or bilateral disease have recently been reaffirmed (Houlston et al., 1992; Tulinius et al., 1992a). Houlston et al. provide graphs which estimate, for first degree relatives of breast cancer patients, probability of disease over 10 years and also lifetime probability, according to age of diagnosis of the index case and age of the sister, mother or daughter requiring assessment. Such relative risks are accurate for the population as a whole, yet absolute risk will obviously vary from family to family if some breast cancer pedigrees represent chance clusters of a common malignancy while others are due to a highly penetrant gene.

Confirmation that there is indeed heterogeneity (both a genetic and a non-genetic component) in the aetiology of breast cancer pedigrees has come from several statistical analyses of cancer distribution within the families of breast cancer probands. These studies are known as segregation analyses. The relative likelihood of obtaining the observed distribution of affected and unaffected pedigree members is computed under different genetic models such as single gene autosomal dominant, recessive or polygenic inheritance. In most studies the maximum likelihood score achieved by application of the model of a single autosomal dominant gene with 80–90 per cent lifetime penetrance and carried by 1–6 in every 1000 women has been sufficient to displace other genetic possibilities (Newman et al., 1988; Claus et al., 1991; Iselius et al., 1991). In breast cancer pedigrees, some individuals affected by malignancy may be nongene-carriers whereas some unaffected relatives may be gene-carriers who have yet to develop the disease. The relative likelihood that any received 1 November 1992
© 1993 ASFRA
Revised 11 December 1992
individual within a pedigree carries the susceptibility gene is described by a combination of age-incidence curves for gene-carriers and nongene-carriers. Segregation analysis provides a model for the putative susceptibility gene which also defines these age-incidence curves. Depending on the actual population gene frequency, carriers will account for between 1 per cent and 6 per cent of all female breast cancer. The different age-incidence curves for gene-carriers and nongene-carriers, however, means that this proportion rises to between 4 per cent and 20 per cent of women diagnosed under age 45.

LOCATING THE BREAST CANCER SUSCEPTIBILITY GENE

Linkage analysis

This is a statistical analysis which aims to utilize molecular genetic typing of pedigree members to locate genes which cause disease. Its successful application relies both on an accurate model for the putative susceptibility gene, as defined through segregation analysis, and also on the existence of multiple polymorphic markers of known location throughout the genome which can be typed both in affected individuals and in close relatives unaffected by malignancy. Linkage analysis computes the LOD (logarithm of odds) score for each pedigree, which is the logarithmic odds in favour of the susceptibility locus existing in close proximity to the marker typed. In general, a LOD score more positive than +3 is required to assert linkage at any given distance between gene and marker locus. This allows a false positive rate of less than 5 per cent (Risch, 1992). A score of +4 would assert linkage with a false positive rate of 0.5 per cent. LOD scores more negative than −2 are generally required to reject the hypothesis that a genetic trait exhibits linkage to the marker locus.

For many years, markers used in linkage analysis have been DNA fragments inherited as single copies and which demonstrate allelic variation or polymorphism. Most sequence variations are single base-pair changes which create or destroy a cleavage site for a specific restriction enzyme, causing a change in length of DNA fragment. Thus the term restriction fragment length polymorphism (RFLP) has been applied to this type of polymorphism. The marker produced by a single such variant exhibits only one or two alleles resulting in very limited data for linkage analysis purposes (Nakamura et al., 1987). In recent years, DNA base sequences containing a variable number of tandemly repeating elements (known as VNTRs or minisatellites) have been discovered. These polymorphisms are highly informative for linkage analysis purposes since they exhibit multiple alleles (Nakamura et al., 1987, 1988). Both RFLPs and VNTRs require a relatively large amount of DNA (10 µg) for DNA digestion and Southern Blot analysis. The use of microsatellites, however, which are very small repeating elements, often only a few base-pairs in length, allows typing of pedigree members through the utilization of minute quantities of blood lymphocyte DNA or through extraction of genomic material from archival paraffin-embedded tissue obtained from pathology specimens. This conservation of DNA resources is achieved through fragment amplification by means of polymerase chain reaction (Mullis and Faloona, 1987) and provides highly informative marker data. Microsatellites are situated every 30–60 kilobases
throughout the genome (Weber and May, 1989); intervals which are short enough for accurate linkage analysis.

_Cytogenetic and 'loss of heterozygosity' studies_

Linkage analysis can only reasonably expect to locate genes in close proximity (within about 30cM) of the marker under analysis. The chance that the breast cancer gene lies within this distance of any randomly chosen marker is only about 2 per cent (Risch, 1992). It is obvious that a great deal of work, which might involve the typing of pedigrees at more than 50 marker loci, would be necessary to scan the whole genome in the absence of intuitive knowledge as to possible sites for the gene.

One approach which has previously identified specific chromosome bands as sites of susceptibility loci in heritable cancers involves karyotypic analysis. The genes responsible for familial retinoblastoma and familial adenomatous polyposis (FAP) were localized to 13q14 and 5q21-22, respectively, on the basis of constitutional cytogenetic abnormalities in affected individuals (Lele _et al._, 1963; Grace _et al._, 1971; Herrera _et al._, 1986). To date, however, no case of familial breast cancer associated with a constitutional chromosome aberration has come to light (Steel _et al._, 1991).

Cytogenetic analysis of breast tumours (in contrast to normal lymphocyte DNA) yields many abnormalities. However, chromosomal rearrangements do seem to occur in some regions more frequently than others and a review of the literature by Mitchell and Santibanez-Koref (1990) reveals that 21 chromosome bands are involved in structural alterations in more than 5 per cent of breast tumours.

Analysis of blood and tumour DNA pairs from the same individual occasionally reveals chromosomal regions in which the two alleles of a polymorphic marker in blood DNA have been reduced to one allele in the tumour sample. This is known as loss of constitutional heterozygosity. There are at least 12 reported loci in which loss of heterozygosity occurs in a large proportion of human breast tumours (Steel _et al._, 1992).

Although there is only limited concordance between the candidate sites identified in cytogenetic and loss of heterozygosity studies, one or more of these loci may be associated with loss of tumour suppressor gene function important in the aetiology of breast neoplasia. Unfortunately, the ability of these studies to reveal candidate susceptibility loci for subsequent linkage analysis assessment has been compromised by the presence of a great deal of background noise from other possible tumour suppressor gene sites. The large number of such loci means that only through a combined collaborative approach by many laboratories can repetition of work be avoided. In recent times this has been made possible through the efforts of the International Breast Cancer Linkage Consortium. This consortium of laboratories has agreed to co-ordinate its activities in the search for a heritable breast cancer gene. Linkage analysis data for candidate loci have been made generally available at an early stage to facilitate this process. Loci which have appeared to give positive evidence for linkage have then been further examined by means of joint analysis of pooled family data.

_IS p53 THE ANSWER?_

Li-Fraumeni syndrome has as its major features early age of diagnosis of malignancy with excess soft tissue and bony sarcoma, adrenocortical carcinoma,
acute leukaemia, CNS malignancy and breast cancer (Birch, 1990). The discovery that perhaps over 50 per cent of Li-Fraumeni syndrome families, as defined by Li et al., (1988), contain affected members with a germline p53 gene mutation has led to the search for mutations in this gene in young breast cancer patients and breast cancer families. Linkage analysis is largely superfluous in this process as the p53 gene, located on chromosome 17p13, has already been fully sequenced.

One report of 126 consecutive breast cancer patients diagnosed under age 40 and screened for constitutional mutations revealed only a single case with p53 gene mutation. This patient gave a family history which included a further two cases of breast cancer and other malignant types in first and second degree relatives (Sidransky et al., 1992). An analysis of 136 breast cancer patients of all ages revealed a single constitutional mutant case who again had a positive family history with breast and other malignancies in close relatives (Prosser et al., 1992). A Norwegian group examined 167 unselected breast cancer probands and additionally 40 patients diagnosed with breast cancer under age 35. In both series an individual with germline p53 mutation was identified and, again, significant family histories of three or four cases of breast cancer in close relatives were obtained from both probands. This group also examined 30 patients with specifically heritable breast cancer but found no p53 mutations (Andersen, 1992). Similarly Prosser et al. (1991) examined five Edinburgh pedigrees (four of which have subsequently exhibited linkage to a 17q gene with a posterior probability greater than 85 per cent) and found no germline p53 mutation in any of the affected individuals screened. In the analyses described above then, a total of four pedigrees with germline p53 mutations have been discovered. Interestingly, although all contained several cases of breast cancer, none fulfilled the strict criteria for Li-Fraumeni syndrome which requires a sarcoma or other embryonal tumour at young age in the family.

It appears from these reports that p53 mutated non Li-Fraumeni breast cancer pedigrees do exist and contribute in small degree to the total burden of breast cancer. They do not, however, explain the majority of heritable breast cancer. For this, another genetic cause must be sought.

BRCA1

In 1990, a seminal publication by Hall et al. reported genetic linkage of some breast cancer families to the polymorphic marker CMM86 on chromosome 17q21-22. CMM86 is a VNTR with more than 30 alleles. Hall’s summed LOD score for 23 breast cancer pedigrees was $-5.48$ assuming the gene was sited at CMM86, but became positive ($+2.35$) at 20cM recombination distance. By subdivision of families according to average age of onset of disease it became apparent that pedigrees with early age onset breast cancer were much more likely to be linked to this marker, so that at zero recombination a summated LOD score of $+5.98$ was achieved in families with a mean age of diagnosis of breast cancer under 47 years. The authors commented on the large number of possible candidate genes located proximal to CMM86 on the long arm of chromosome 17. These include c-erb B2 (her-2) which has the properties of a growth factor receptor and is amplified in many primary breast tumours, 17β oestradiol dehydrogenase (17HSD) which converts oestrone to oestradiol homeobox 2 (hox 2) critical in early development, nm23
expression of which is lost in association with breast cancer lymph node metastasis and INT4 which is homologous with the mouse mammary virus integration site and encodes a growth factor-like product.

Close on the heels of this publication followed a collaborative study from French and Canadian laboratories in which linkage to CMM86 was strongly positive in two out of five large breast/ovarian cancer pedigrees, the largest of which gave a maximal LOD score of $+2.71$ at 10cM recombination distance (Narod et al., 1991).

At the November 1990 and May 1991 meetings of the International Breast Cancer Linkage Consortium in London and Edinburgh there was excitement at the prospect of having located a major susceptibility gene. It was apparent, however, that the combined consortium family data set could not accurately define the site of the gene, denoted BRCA1, in this region. The data supported a range of possibilities from, at one extreme, a large proportion of families linked to a 17q gene up to 20cM from CMM86 to, at the other extreme, a small proportion of mainly early age onset families tightly linked to the locus (Easton et al., 1993).

In order to resolve this problem of location, the consortium made available a large number of highly polymorphic microsatellite oligonucleotide probes mapping to chromosome 17q which were distributed to collaborative laboratories. By means of multipoint analysis, loci which produced the most highly positive LOD scores were identified and probable flanking markers defined.

Early in the analysis of these marker data it became apparent that probes telomeric to CMM86 produced most negative LOD scores. Effort was therefore concentrated on more proximal loci where most of the candidate genes for breast cancer susceptibility were known to exist. Four probes typed in 15 Edinburgh breast cancer families were CMM86, NM23, 4206, and MFD188. These extend across a '12-28cM region of 17q12-21 (Easton et al., 1993). Two point LOD scores at 42D6 and MFD188 were the most highly positive. A LOD of $+5.62$ was registered with assumed genetic location 2.5cM from 42D6. A LOD of $+3.69$ was registered with assumed genetic location 7.5cM from MFD188. These two markers are situated 5–10cM apart. Allele data from Edinburgh pedigrees were utilized in the International Breast Cancer Linkage Consortium analysis which included a total of 214 breast cancer and breast/ovarian cancer families from 13 groups (Easton et al., 1993). Again, the most highly positive two-point LOD scores were achieved using markers 42D6 and MFD188 in the combined consortium data set.

One means of defining markers which actually flank the susceptibility gene is to identify families which are probably linked to the gene and determine if there are any individuals with breast cancer who appear not to have inherited a marker allele which segregates elsewhere in the pedigree, with the disease. Having identified such individuals there are three reasons why this phenomenon might occur. First, the patient concerned could exhibit a sporadic (non-genetic) form of the disease. This possibility is low when the individual has had breast cancer diagnosed at early age or has had bilateral disease or ovarian cancer and is, for practical purposes, formally excluded if the individual has transmitted the trait to offspring. Second, a double recombination event might have occurred between adjacent markers so that the disease gene becomes associated with different marker alleles to those segregating in other affected pedigree members. The chance that a double recombination event might occur between adjacent markers used in the consortium analysis is low,
however, since the distance between them is relatively small. Third, a single meiotic cross-over may have occurred between the gene and one of its flanking markers. If one assumes a high probability that the cancer case is truly genetic and that a double recombination event has not occurred, then this phenomenon would exclude the chromosome arm beyond the marker ‘abandoned’ by the disease as a possible site for the susceptibility gene. Such a case is described as a critical recombinant.

Performing this kind of analysis for each linked pedigree allows a recombination exclusion map to be produced. Pedigrees which are likely to be linked to the chromosome 17q gene can be identified on the basis of prior and conditional probability in accordance with Bayesian principles (Emery, 1986). The prior probability is the likelihood that the family is linked to a 17q gene simply on the basis of pedigree structure. International Linkage Consortium data suggests that this is likely to be about 50 per cent for breast cancer pedigrees and over 75 per cent for breast/ovarian cancer pedigrees. The conditional probability of linkage is determined from marker allele data using 17q probes. The posterior or final probability of linkage is calculated from these figures. We compute that 7 out of 15 Edinburgh pedigrees register posterior odds of linkage to chromosome 17q greater than 85 per cent (range 85.5 per cent to 99.9 per cent). Four of these families contain a total of five members with critical recombinants involving 42D6 and MFD188. Two critical recombinants exclude regions telomeric to 42D6 and three exclude regions centromeric to MFD188. These are therefore considered to be the most probable flanking loci of the susceptibility gene in Edinburgh families (Porter et al., 1993a) although some other groups in the International Linkage Consortium have critical recombinants which place the gene just centromeric to MFD188.

Additional evidence that this region contains an important tumour suppressor gene is accumulating from loss of heterozygosity studies co-ordinated by the International Linkage Consortium. These involve the same markers utilized in family linkage analysis. Although this work is ongoing, if preliminary results from Edinburgh are confirmed, it appears that the degree of loss of heterozygosity in breast tumours is particularly high at 42D6 and MFD188 on chromosome 17q12-21. Furthermore, work on breast tumours from BRCA1-linked breast-ovarian cancer pedigrees reveals that in each of nine tumours which showed loss of allele heterozygosity, the losses were from the wild-type chromosome (Smith et al., 1992). This suggests that BRCA1 is indeed a tumour suppressor gene.

**IMPLICATIONS**

The localization of a major susceptibility gene which is responsible for perhaps 50 per cent of the genetic component of breast cancer is a major breakthrough. At present it allows identification of certain individuals at particularly high risk through 17q marker typing, although this is only possible in pedigrees with several cases of breast cancer in which many individuals have already been typed. Only extensive pedigree ascertainment is able to produce LOD scores of sufficient magnitude to significantly alter risk assessment for unaffected individuals. Examples of such pedigrees are shown in Figure 1. Only diagnoses of breast and ovarian cancer are shown, together with allele typing at 42D6 and MFD188. Posterior probability of linkage is greater than 90 per cent in these families. The unaffected male and
female relatives marked with an asterix have a high probability that they are gene carriers. The likelihood that female carriers will eventually develop malignant disease is determined not only by carrier status but also by age at which risk assessment is performed (age-specific penetrance curves for carriers have been determined by the International Linkage Consortium) and by non-genetic factors which are known to affect the chance of malignant transformation. Computer programs which assess risk of breast cancer for individuals on the basis of genetic marker and non-genetic data include modifications of the program LIPID (Ott 1974, 1976) by Winter (1985).

Just as certain individuals will be at greater known risk of disease following allele typing, so will certain others be at lower than previously estimated risk. The discovery of polymorphic markers which flank BRCA1 even more tightly will enable LOD scores to provide more accurate risk assessment for family members. The importance of good genealogy should not be overlooked as linkage analysis can be wildly misleading for families in which these basic data are erroneous.

What can be done to prevent or detect disease early in suspected high risk individuals? This question has been posed in the past with reference to all female members of breast cancer pedigrees but now carries greater urgency as a result of advances in genetic screening. There are as yet few data to assist us in defining beneficial protocols and we have to draw inferences from our knowledge of screening and treatment of large population groups. In Edinburgh families linked to
chromosome 17q with a posterior probability greater than 85 per cent we have not yet given patients a detailed interpretation of their likely genetic status as revealed by linkage analysis. We believe that the application of research findings in clinical practice should always proceed with caution, particularly where risk assessment may be inaccurate. For example, a woman from a family with 90 per cent probability of BRCA1 linkage and whose personal lifetime risk is elevated from 35 to 70 per cent as a result of marker data will have a 10 per cent chance that the assessment of family BRCA1 linkage status is erroneous. We consider that from an ethical standpoint it is quite acceptable to await a definitive BRCA1 gene identification in order to confirm genetic status prior to offering very precise estimates of risk based on specific and personal linkage results. The complex psychosocial aspects of counselling women with a positive family history of breast cancer and who often have inaccurate perceptions of their own risk should not be overlooked (Kelly, 1980; Evans et al., 1991) and give further cause for a cautious clinical approach. At many centres, women with a positive family history are invited to report at high-risk clinics for annual examination. Some groups recommend that this should commence at age 30 or 5 years younger than the earliest diagnosis of breast cancer in the family. Ultrasonography and mammography may also be used in these clinics, although there is concern that gene-carriers may be more sensitive to ionizing radiation (Kovacs and Almendral, 1987; Lemon, 1991). Surgical prophylaxis in the form of bilateral subcutaneous mastectomy may occasionally be considered and there are at least theoretical grounds for believing that the anti-oestrogen Tamoxifen may provide useful prophylaxis. Evaluation of this drug in high risk women is currently underway as part of an international collaboration.

With respect to ovarian cancer, women at risk may have an age-incidence curve shifted to the right of that for breast cancer. Organ targeted screening may again involve annual examination and ultrasonography, with the possibility that colour doppler ultrasound, in skilled hands, may detect occult ovarian tumour blood-flow. It is unfortunate that in women with a positive family history, bilateral oophorectomy does not always abolish risk of non-ovarian intra-abdominal malignancy which on microscopy is histologically indistinguishable from epithelial ovarian carcinoma (Tobacman et al., 1982). It is hypothesized that this may be because the susceptibility gene produces a tissue-specific field change which affects the entire coelomic endothelium from which ovaries, omentum, and peritoneal lining derive.

The range of organs targeted by the BRCA1 gene is perhaps less wide than in Li-Fraumeni syndrome. Studies to determine whether malignancies such as prostate, CNS, endometrial and colorectal tumours occur in excess in gene carriers are ongoing (Tulinius et al., 1992b). The possibility that a breast cancer gene may confer protection from other diseases such as ischaemic heart disease (Porter et al., 1991), as appears to be the case in relatives of colon cancer patients, is a compensatory silver lining to the cloud of anxiety which surrounds breast cancer susceptible families.

Although immediate clinical application of linkage analysis data is limited, the major benefit which localization of BRCA1 will eventually bring is its definitive identification. Only then will it be possible to screen individuals for relevant mutations without reference to laborious family studies in order to determine carrier
The search for the cause of the remaining 50 per cent or so of familial breast cancer continues. Some non 17q-linked pedigrees will certainly be chance clusters while others will certainly have a separate genetic aetiology. The identification of a germ-line p53 gene mutation which precipitates perhaps 0.5 per cent to 1 per cent of breast cancer overall appears to provide explanation for only a small proportion of familial breast cancer. It may be that modifications of the existing segregation analysis model will be necessary to adjust for the removal of BRCA1-linked families when LOD score analysis of non 17q-linked families is performed. The genetic aetiology of such remaining families may well be heterogenous. Karyotypic studies of pedigree members will continue to be worthwhile and may yet provide the key to the identification of further genes, as in retinoblastoma and FAP. There is some evidence, for example, in one of the Edinburgh breast/ovarian cancer pedigrees not linked to BRCA1 and exhibiting several different tumour-types that a constitutional chromosome 15p duplication segregates with malignancy in the karyotypically complex family (Porter et al., 1993b).

The field of breast cancer genetics is advancing rapidly. Within the next 5 years the identification of BRCA1 (and perhaps other genes) is likely to impose itself powerfully on clinical practice.

REFERENCES


