http://www.expertreviews.org/

expert reviews in molecular medicine

Breast cancer gene discovery

Kornelia Polyak

Many important advances have been made in the past decade in understanding breast cancer at the molecular level, and two important high-penetrance breast cancer genes – *BRCA1* and *BRCA2* – have been identified. However, germline mutations in these two genes are responsible for only a minority (~5%) of all breast carcinomas, and the genes responsible for the majority of breast cancer cases remain to be identified. There is evidence that there are additional high-to-moderate-penetrance breast cancer susceptibility genes but, given the high degree of molecular heterogeneity in breast carcinomas, it is likely that each of these genes is responsible for only a subset of cases. There are also many candidate low-penetrance breast cancer genes and many more are likely to be identified. In addition to germline, and somatic, sequence alterations, epigenetic changes in many genes are likely to play an important role in the pathobiology of breast cancer. Recently developed genomic technologies and the completion of the human genome sequence provide us with powerful tools to identify novel candidate breast cancer genes that could play an important role in breast tumourigenesis.

What is a breast cancer gene? In a strict definition, a breast cancer gene is a gene in which germline mutations or polymorphisms confer increased susceptibility to breast cancer. In a broader sense, breast cancer genes are genes that play a causative role in the pathogenesis of breast cancer even if they do not have any germline or somatic sequence alterations. This latter group includes genes that are aberrantly expressed in breast tumours due to epigenetic changes (such as methylation), and genes that may be abnormally activated/inactivated at the protein level. Some of these genes are specific for breast cancer, whereas others are involved in other cancer types as well.

Breast cancer is one of the most common cancers diagnosed in women worldwide and is a leading cause of cancer-related death (Ref. 1). Therefore, the identification of breast cancer genes is a major scientific and social problem. The identification of such genes will not only enable the identification of individuals at high risk, but also aid in the design of more-effective therapies for breast cancer prevention and treatment. Pathological stages in breast tumourigenesis, along with an indication of underlying genetic alterations, are shown in Figure 1. Similar to the development of other cancer types, breast tumourigenesis is a multistep process; it starts with ductal hyperproliferation and progresses into in situ, then invasive, and finally metastatic carcinomas.

Breast cancer susceptibility is genetically determined, but the penetrance (i.e. the frequency of individuals with a specific genotype who manifest the trait in the phenotype) of known

Kornelia Polyak
Assistant Professor of Medicine, Department of Adult Oncology, Dana-Farber Cancer Institute and Harvard Medical School, 44 Binney St, Boston, MA 02115, USA. Tel: +1 617 632 2106; Fax: +1 617 632 4005; E-mail: Kornelia_Polyak@dfci.harvard.edu

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

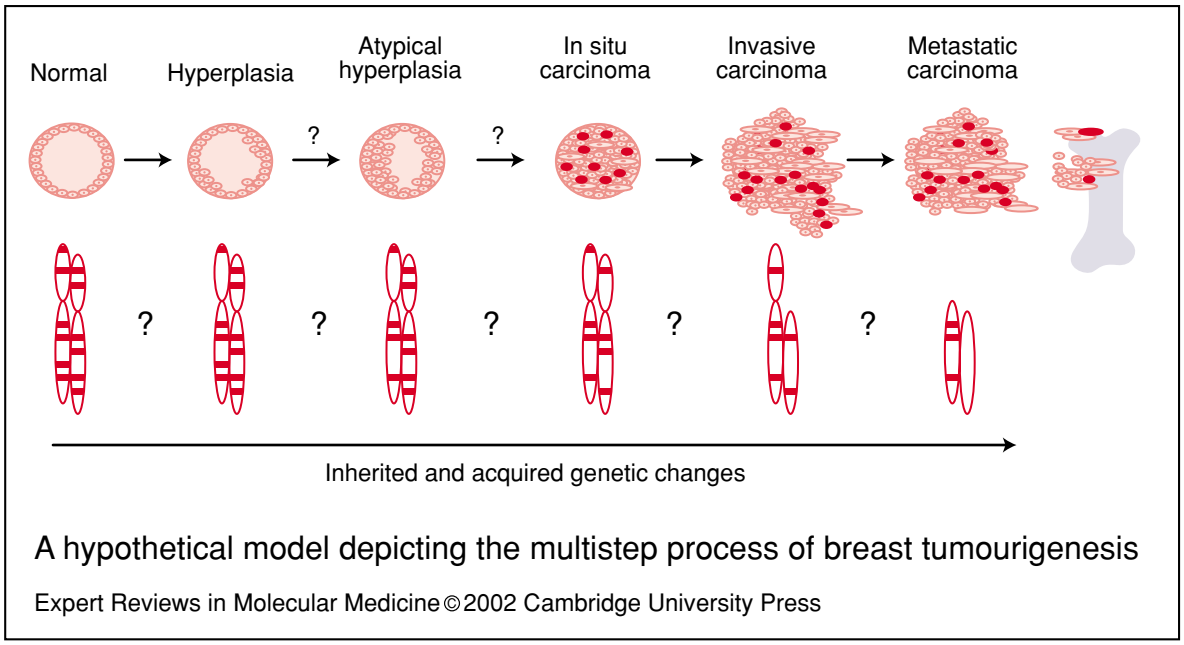

Figure 1. A hypothetical model depicting the multistep process of breast tumourigenesis. Clinical and pathological stages are indicated. The relationships among the earliest-stage tumours (hyperplasia, atypical hyperplasia, and in situ carcinoma) are not well understood and it is questionable if all these lesions represent obligate precursors of invasive carcinomas. Tumour progression is driven by inherited and acquired genetic changes, which are reflected in altered profiles of gene expression and of protein levels and modifications **(fig001kpb)**.

genes that predispose to breast cancer varies widely (Ref. 2). A minority of breast cancer cases are attributable to inherited mutations in rare high-penetrance breast cancer susceptibility genes such as *BRCA1* and *BRCA2* (i.e. individuals with mutations in these genes frequently develop breast cancer), but the majority of cases are probably due to multiple low-penetrance genes (Refs 2, 3, 4). This assumption is based on multiple studies suggesting that most of the excess familial risk of breast cancer is genetic, yet known genes explain fewer than 25% of this excess risk, indicating that other breast cancer genes remain to be identified (Refs 5, 6, 7, 8). Analyses of genetic models to explain familial breast cancer risk that is not due to *BRCA1* and *BRCA2* suggest that the most likely scenario is a polygenic model with many susceptibility alleles, although the existence of an autosomal recessive mendelian allele cannot be excluded (Refs 9, 10). It is the interaction between these genes and environmental factors (e.g. geographical location, smoking, fruit/vegetable consumption) that determines the individual's overall breast cancer risk. In this sense, breast cancer, similar to diabetes, asthma or heart disease, is a complex genetic disease (Ref. 11).

Family linkage studies are useful for the identification of high-penetrance, and perhaps even moderate-penetrance, breast cancer genes, whereas the identification of low-penetrance genes requires a combination of approaches including candidate-gene-based association studies in affected and non-affected populations, analysis of somatic mutations in tumours, and the use of novel, comprehensive and unbiased genomic techniques (Fig. 2) (Ref. 11). This article reviews the strategies and techniques that are available for the identification of additional breast cancer genes.

Linkage studies: high-penetrance genes

Families with multiple cases of cancer, frequently arising at an early age, have been useful for the identification of many cancer predisposition genes. Although inherited mutations in such genes are rare, these genes are frequently mutated in sporadic tumours as well. Breast cancer is particularly well suited for these types of studies, since one of the strongest risk factors for breast cancer is the occurrence of the disease in first-degree relatives, particularly if they are

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994 ©2002 Cambridge University Press

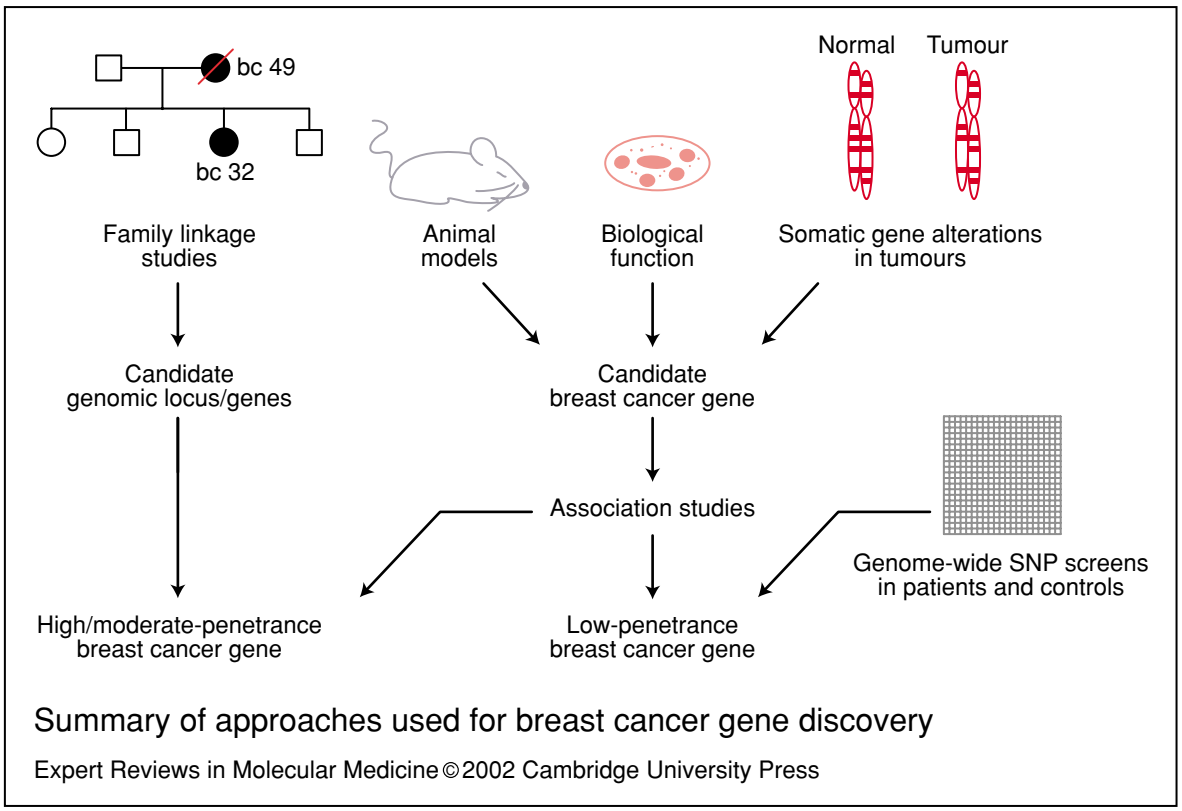

Figure 2. Summary of approaches used for breast cancer gene discovery. Family linkage studies are useful only for the identification of high-to-moderate-penetrance breast cancer genes. Association studies using unbiased genome-wide screens, or candidate genes selected on the basis of biological function, somatic gene alterations (genetic, epigenetic and gene expression changes) or animal models, usually identify low-penetrance breast cancer genes. bc 49, breast cancer arising at age 49 years; SNP, single-nucleotide polymorphism **(fig002kpb)**.

diagnosed at an early age. The identification of a susceptibility locus on chromosome 17 was, therefore, of great importance, not only because it proved that early-onset breast cancer is inherited but also because it demonstrated that a high-penetrance breast cancer susceptibility gene exists (Ref. 12). The hunt for *BRCA1*, and subsequently *BRCA2*, was a 'brute-force' linkage mapping approach in affected families and was one of the most competitive races in the history of gene cloning (Refs 13, 14); *BRCA1* was found on chromosome 17q21 and *BRCA2* on 13q12.3. However, in contrast to most cancer predisposition genes identified through family studies, neither of these two genes is mutated in sporadic breast carcinomas (Refs 15, 16).

Although there are some indications that somatic epigenetic inactivation of *BRCA1* and *BRCA2* occurs in some sporadic tumours, overwhelming evidence suggests the occurrence of differing tumourigenic pathways for inherited and sporadic breast carcinomas (Refs 17, 18). Somewhat disappointingly, follow-up large-scale mutation screening studies determined that not only are *BRCA1* and *BRCA2* not mutated in sporadic tumours, but mutations in these two genes are responsible for only a fraction of inherited cases (Ref. 2). Since family history still remains a strong predictive factor in patients with no detectable mutations in these genes, additional high-to-moderate breast cancer susceptibility genes remain to be identified (Refs 5, 6, 7, 19, 20). Germline mutations in several other genes (Table 1), including *TP53* ('tumour protein p53'), *PTEN* ('phosphatase and tensin homologue'), *LKB1*, *CHK2* ('checkpoint homologue 2') and *MLH1*/*MSH2* ('MutL homologue 1/MutS homologue 2'), also predispose individuals to breast cancer, but these cases are very rare and are usually part of a multicancer and/or developmental abnormality syndrome (Ref. 21). Since the discovery of *BRCA1*

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994 ©2002 Cambridge University Press

Table 1. Genes with germline mutations associated with breast cancer susceptibility (tab001kpb)

Gene name	Mode of discovery	Putative function	Refs
BRCA1	Linkage in families	DNA-damage checkpoint control, recombination, transcription	12, 13
BRCA2	Linkage in families	DNA-damage checkpoint control, recombination, transcription	14
BRCAx-Chr13q21	CGH plus linkage in families	?	24, 28, 29
BRCAx-Chr8p	Linkage in families	?	22
TP53	Candidate-gene testing in families	DNA-damage checkpoint control	92
LKB1	CGH plus linkage in families	Apoptosis/growth arrest?	30, 93, 94, 95
PTEN	RDA plus candidate-gene testing in families	Negative regulation of mitogenic/survival signals	51, 53
MLH1/*MSH2*	Linkage in families	Mismatch repair	96, 97
ATM	Linkage in families	DNA-damage checkpoint control	98
CHK2	Candidate-gene testing in families	DNA-damage checkpoint control	99
HRAS1-VNTR	Candidate-gene testing	Proto-oncogene: altered transcription/ linkage disequilibrium/genomic instability	100
CYP1A1	Candidate-gene testing	Metabolism of oestrogen and polycyclic aromatic hydrocarbons	101
CYP2D6	Candidate-gene testing	Metabolism of steroid hormones	102, 103
CYP17	Candidate-gene testing	Metabolism of steroid hormones	104
NAT1	Candidate-gene testing	Metabolism of aromatic and heterocyclic amines	105, 106
NAT2	Candidate-gene testing	Metabolism of aromatic and heterocyclic amines	105, 107
COMT	Candidate-gene testing	Metabolism of catechol oestrogens	108, 109
SOD2	Candidate-gene testing	Metabolism of superoxide anions	110, 111
SULT1A	Candidate-gene testing	Metabolism of carcinogens and endogenous hormones	112, 113
GSTM1	Candidate-gene testing	Altered detoxification of carcinogens	114, 115
GSTP1	Candidate-gene testing	Altered detoxification of carcinogens	114
GSTT1	Candidate-gene testing	Altered detoxification of carcinogens	114
XRCC1, 3, 5	Candidate-gene testing	Base-excision repair	116

(continued on next page)

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994 ©2002 Cambridge University Press

Table 1. Genes with germline mutations associated with breast cancer susceptibility (tab001kpb) *(continued)*

Gene name	Mode of discovery	Putative function	Refs
TNF-α	Candidate-gene testing	Altered TNF-α levels/signalling	117, 118
HSP70	Candidate-gene testing	Altered oestrogen signalling?	117, 118
ESR1 (*ERα*)	Candidate-gene testing	Altered oestrogen signalling	119
AR	Candidate-gene testing	Altered androgen signalling	36
PGR	Candidate-gene testing	Altered progesterone signalling	37
AIB1	Candidate-gene testing	Altered oestrogen signalling	38
HLA	Candidate-gene testing	Immune surveillance	120

Abbreviations: AIB, amplified in breast cancer 1; AR, androgen receptor; ATM, ataxia telangiectasia mutated; BRCA 1, 2, x: breast cancer 1, 2, x; CGH, comparative genomic hybridisation; CHK2, checkpoint homologue; COMT: cathecol-*O*-methyltransferase; CYP, cythochrome P 450; ER, estrogen (oestrogen) receptor; GST, glutathione transferase; HLA: histocompatibility antigen; HRAS1-VNTR, variable number of tandem repeats; HSP70, heat shock protein 70; MLH1, MutL homologue 1; MSH2, MutS homologue 2; NAT, *N*-acetyl transferase 2; PGR, progesterone receptor; PTEN, phosphatase and tensin homologue; RDA, representational difference analysis; SOD2, superoxide dismutase 2; SULT1A, sulphotransferase 1A; TP53, tumour protein p53; TNF-α, tumour necrosis factor α; XRCC 1, 3, 5, X-ray repair complementing defective repair in Chinese hamster cells 1, 3, 5.

and *BRCA2*, two additional loci, on chromosomes 8p and 13q, have been implicated in beast cancer susceptibility, but so far no new breast cancer susceptibility gene (referred to here as *BRCAx)* has been identified (Refs 22, 23, 24).

Both of these new loci were identified through initial molecular studies performed on the tumours of the affected patients. In the case of chromosome 8p, multiple independent studies have demonstrated a high fraction of loss of heterozygosity (LOH) in this region in sporadic tumours (Refs 25, 26). Many of these studies were performed in early-stage tumours, such as in ductal carcinoma in situ [DCIS; before the tumour grows through the breast duct wall into the surrounding tissues (Fig. 1)], and are therefore presumably devoid of random chromosomal alterations. Thus, the probability that these LOH events target a legitimate breast tumour suppressor gene is high. Initial linkage studies performed with chromosome 8p markers in a small set of German breast cancer families established a positive LOD score (this being the logarithm of the total relative probability that a linkage relationship exists among selected loci) for this locus (Refs 22, 23). However, in a larger set of families, no evidence for linkage was found (Ref. 27). Therefore, if a *BRCAx* gene exists on chromosome 8p, it is likely to be responsible for only a very small portion of inherited cases.

The candidate 13q21 *BRCAx* locus was identified on the basis of comparative genomic hybridisation (CGH) studies in tumours of patients with familial breast cancers (at least three cases per family, and without germline mutations in known high-penetrance breast cancer genes) (Ref. 24). The most frequently deleted locus in these tumours was chromosome 13q21, and this was then tested in linkage studies in the affected families where a positive LOD score was established. However, linkage analyses in two different sets of breast cancer families could not confirm positive linkage to 13q21 (Refs 28, 29). The discrepancy between these studies could be due either to the fact that the families were selected on the basis of different criteria or to the fact that the 13q21 *BRCAx* gene is responsible for only a minority of familial cases. Only the identification of the gene implicated will determine which of these possibilities is true. Nevertheless, a similar approach (CGH followed by targeted linkage) has been successfully used for the identification of the *LKB1* tumour suppressor gene (Ref. 30). Therefore,

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994 ©2002 Cambridge University Press

it is likely to be useful for the identification of *BRCA* genes as well. The best way to perform these types of studies would be to analyse multiple early-stage tumours from the same patient using a high-resolution genome-wide genetic screen, such as the bacterial artificial chromosome (BAC) or cDNA-array CGH, followed by high-throughput candidate-gene testing of the implicated loci.

The fact that despite all these efforts no *BRCAx* gene has been identified in the past six years indicates that the search for *BRCAx* is a difficult task. There are multiple possible reasons why *BRCAx* is more difficult to find than either *BRCA1* or *BRCA2* was. First, *BRCA1* and *BRCA2* predispose to other relatively rare cancer types – ovarian (both genes) and male breast cancer (*BRCA2*) – in addition to early-onset breast cancer, making the identification of affected families easier. Second, and somewhat related to the first problem, breast cancer is a very common cancer in women; therefore, it is possible that, by selecting families solely based on the occurrence of breast cancer, sporadic cases (i.e. not due to inherited mutation in a high-penetrance gene) will be inadvertently included, hampering the power of the linkage studies. Third, the penetrance of *BRCAx* might be much lower than that of *BRCA1* or *BRCA2*, and observing the phenotype could require cooperation with multiple low-penetrance modifier genes. Finally, it is possible that there are multiple high-to-moderate-penetrance *BRCA* genes to be identified, each being responsible for only a small number of inherited cases.

Association studies: low-penetrance genes

Candidate genes based on function

Whereas linkage studies in families are useful for the identification of high-penetrance, and to a lesser extent even moderate-penetrance, genes, low-penetrance genes are impossible to identify using this approach (Ref. 11). Although low-penetrance genes confer only a small-to-moderate breast cancer risk, variants of these genes might occur at a high frequency in the population. Therefore, combinations of low-penetrance genes are likely to be responsible for a much larger fraction of breast cancer cases than are high-penetrance genes.

To date, most low-penetrance genes have been identified on the basis of candidate selection approaches followed by association studies in affected and control populations (Table 1) (Ref. 3). The simplest and most efficient version of these association studies is a case–control approach that compares allele frequencies in breast cancer and in control, unaffected patients. In order to be confident that the observed association of allele frequencies with breast cancer is not due to the biased selection of patients, the affected and the unaffected populations have to be well matched (with respect to ethnicity, age, gender and so on). Genes to be tested have usually been selected if their biochemical function is thought to play a role in breast tumourigenesis (Table 1). Therefore, several of these low-penetrance breast cancer susceptibility genes are involved in DNA-damage signalling and repair, hormone and xenobiotic metabolism, anti-oxidant defence, and immune surveillance (see Refs in Table 1).

In addition, polymorphisms in high-penetrance genes have been investigated to determine if any of these have an effect on breast cancer risk. For example, several common polymorphisms in *TP53*, *ATM* ('ataxia telangiectasia mutated'), *BRCA1* and *BRCA2* have been extensively studied in several different populations (Refs 3, 31, 32, 33, 34, 35). Although the increased risk of breast cancer associated with the R841W *BRCA1* variant remains to be confirmed in larger populations, several *BRCA2* variants and the Arg72Pro polymorphism in *TP53* do appear to confer a slightly elevated risk of breast cancer (Refs 3, 33).

Modifiers of high-penetrance genes

Another potential way to identify low-penetrance breast cancer susceptibility genes is to search for genes that modify the penetrance of high-penetrance genes, although many investigators believe that modifier genes probably account for only a tiny fraction of the total variance in these families. According to some investigators, the power of this approach might be that because there are many breast cancer cases in the affected cohort, even genes with slight effects on penetrance can be revealed. However, there are potential problems as well, including the possibility that modifiers of high-penetrance genes might not have an effect in the general population, and difficulties associated with selecting enough unaffected controls from the mutation carriers. In the case of *BRCA1* and *BRCA2*, several studies have been performed to identify genes that influence the age of onset and the occurrence of associated ovarian cancer

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

(Ref. 20). Among others, genes involved in oestrogen and androgen signalling (oestrogen, androgen and progesterone receptors, and an oestrogen receptor co-activator, AIB1), as well as detoxifying enzymes [glutathione *S*-transferase T1 (GSTT1)], might affect cancer penetrance in mutation carriers (Refs 36, 37, 38). In addition, targeted genotyping approaches indicate the presence of a modifier locus at 5q33-4, but the gene responsible has not been identified (Ref. 2). An interesting finding of an unselected case–control association study was that one of the alleles of *BRCA2* (N372H) had a different distribution among females and males, with females showing an excess of heterozygotes but no homozygotes, whereas the reverse was found in males (Ref. 34). This might indicate that there is a sex-specific selection for certain alleles either during the generation of the germ cells or in utero during embryonic development, owing to the fact that BRCA2 function could be important for these processes. This result also indicates that breast cancer patients should be analysed for associated phenotypes unrelated to the cancer, such as infertility, since these might aid the discovery of new modifier alleles.

In addition to polymorphisms in the coding region of the candidate breast cancer susceptibility genes, recent results indicate that sequence variations in the non-coding regions such as promoters and introns could also have an effect on breast cancer risk. The molecular effects of these alterations are not always clear, but one possibility is that they influence the mRNA/protein levels of the genes. The importance of this type of effect is highlighted by a recent analysis of the effect of a sequence change in the *APC* ('adenomatous polypsosis coli') gene in colon cancer (Ref. 39). An unidentified sequence alteration outside of the coding exons led to a twofold decrease in APC protein levels that was sufficient to initiate colonic polyp formation and increase colon cancer risk. Similarly, a polymorphism in the promoter region in the gene encoding insulin-like growth factor I (IGF-I) appears to be associated with increased breast cancer risk (Ref. 40).

SNP-ing the genome

Although the candidate-gene testing approach has been successful in identifying several putative new low-penetrance breast cancer genes, and will continue to be used in the future, it is inherently limited to characterised genes (which currently constitute the minority of genes encoded by the human genome). Unbiased comprehensive studies are needed to uncover additional, uncharacterised breast cancer genes using polymorphic markers. So far, these studies have been technically challenging to complete in the general population. However, recent advances in many different fields will probably make these studies possible. The completion of the human genome project and subsequent re-sequencing of parts of the genome in different individuals has led to the discovery of a large number of single-nucleotide polymorphisms (SNPs) (Ref. 41). The identification of these SNPs, in combination with recently developed high-throughput SNP-genotyping methods, will enable us to perform genome-wide association studies in large populations (Refs 42, 43, 44). The recent finding that the human genome is probably composed of blocks (tens to hundreds of kilobases) of haplotypes might make the execution of genome-wide association studies easier (Ref. 45), since instead of individual SNPs of unknown relevance one can look at the frequency of haplotypes (containing multiple SNPs) in the affected and control populations. Although no study of this type has been performed in breast cancer, it is likely that such studies will be conducted in the future once the haplotype map of the human genome is defined. In addition to the technical challenges involved, there are several other difficulties that might make these types of studies impossible to perform or interpret. For example, the extent of linkage disequilibrium in surrounding common gene variants in various human populations is unknown. Similarly, the number of SNPs required to perform genome-wide studies is still subject to debate. Several recent reviews discuss the current status of SNP genotyping and future challenges (Refs 46, 47, 48, 49, 50).

Somatic genetic alterations in tumours

Homozygous deletions and mutations

In addition to analysing germline mutations, studying somatic genetic alterations in breast tumours might also lead to the identification of candidate breast cancer susceptibility genes. The identification of the *PTEN* and *DPC4*/*Smad4* ('deleted in pancreatic carcinoma 4'/'SMA and MAD related protein 4') tumour suppressor genes through mapping homozygous deletions in breast

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

and pancreatic tumours, respectively, proves the validity of this approach (Refs 51, 52). Similarly, the locus responsible for Peutz–Jeghers syndrome (the *LKB1* gene), which is an autosomal dominant condition associated with multiple gastrointestinal hamartomatous polyps and melanocytic spots of the lips, buccal mucosa and digits, was initially localised by CGH studies in tumours of affected patients (Ref. 30).

There are many different techniques that could be used for the identification of genes deleted or amplified in tumours. The *PTEN* tumour suppressor gene was discovered based on representational difference analysis (RDA) of a breast tumour followed by mapping of the deleted fragment and candidate-gene testing (Ref. 51). In the case of *DPC4*, detailed LOH mapping in multiple tumours led to the definition of a locus that corresponded to a region of consensus LOH and a homozygous deletion (Ref. 52). Although PTEN is almost never mutated in sporadic breast carcinomas, germline mutations within the gene are responsible for two inherited cancer predisposition syndromes: Cowden's disease and Bannayan–Riley–Ruvalcaba syndrome (Refs 53, 54, 55).

Candidate-gene approaches are also applied in searches for genes somatically inactivated in tumours, and almost all newly identified candidate tumour suppressor genes tested for mutations in breast carcinomas show varying results. Examples for these genes are those encoding 'Patched' (*PTCH*), mitogen-activated protein kinase kinase 4 (*MKK4*) and E-cadherin (*CDH1*) (Refs 56, 57, 58, 59, 60) (Table 2). PTCH is a receptor for 'Sonic hedgehog' (SHH) and is mutated in a hereditary cancer predisposition known as Gorlin syndrome. In one study, analysis of PTCH in seven breast tumours identified two missense mutations of unknown relevance, whereas a larger study analysing 45 breast tumours found no evidence of genetic alterations (Refs 56, 57). Thus, alterations in PTCH/SHH signalling are unlikely to play an important role in breast tumourigenesis. The *MKK4* gene was implicated as a tumour suppressor gene since it was found to be homozygously deleted in two human tumour cell lines, derived from pancreatic and lung carcinomas (Ref. 58). E-cadherin is a cell adhesion protein with growth/invasion suppressor function and shows germline mutations in hereditary diffuse gastric carcinoma patients (Ref. 61). In addition, the *CDH1* gene encoding E-cadherin is located on chromosome 16q, a region demonstrating frequent LOH in sporadic breast carcinomas (Ref. 62). Analysis of 42 medullary and infiltrating ductal breast carcinoma samples revealed no alterations in *CDH1*, but four out of seven lobular invasive carcinomas harboured a truncating mutation (Ref. 59). Since this initial study, many other groups have independently confirmed the presence of frequent *CDH1* mutations in lobular breast carcinomas, indicating that mutations in this gene play an important role in the development of lobular carcinomas. Despite the fact that *CDH1* mutations have been identified very infrequently in ductal carcinomas (Ref. 63), the expression of *CDH1* is frequently abolished by promoter methylation, and LOH at 16q is also frequently observed in these tumour types (Ref. 64). This might indicate either that elimination of CDH1 protein function is also important in ductal carcinomas or that LOH at 16q targets an as-yet-unidentified tumour suppressor gene.

High-throughput genotyping and mutation screens

Recent technical advances make the search for somatic genetic alterations easier to perform and allow large-scale 'brute-force' approaches. The above-described genome-wide SNP arrays can be used for the localisation of regions deleted or amplified in the tumours. Similarly, a genome-wide BAC or cDNA-array CGH can facilitate the identification of new breast cancer genes (Refs 65, 66, 67). The simultaneous comprehensive genotype and gene expression analysis of tumours – using CGH or SNP arrays for genotyping, and cDNA/oligo arrays or serial analysis of gene expression (SAGE) for the analysis of gene expression patterns – promises to be especially powerful for the uncovering of new candidate genes.

A massive 'brute-force' project for new cancer gene discovery was recently initiated by the Wellcome Trust Fund at the Sanger Institute (Cambridge, UK) (see http://www.sanger.ac.uk/CGP/ for details). One of the goals of this project is to conduct a genome-wide homozygous deletion screen in over 1200 human cancer cell lines using polymorphic CA/GT repeats followed by STS screens ('STS' stands for sequence-tagged site; short stretches of unique DNA sequence that can be specifically detected by a PCR assay), with the aim of identifying new tumour suppressor

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

Table 2. Genes with somatic genetic and epigenetic alterations in breast cancer (tab002kpb)

Gene	Alteration	Putative function	Refs
MKK4	Mutation	Signal transduction	58, 72
PTCH	Mutation	Hedgehog signalling	56, 57
CDH1	Mutation/methylation	Cell adhesion	60, 64
17q- erbB2	Amplification	Receptor tyrosine kinase	73, 121
8q-c-myc	Amplification	Transcription factor	73, 122
11q-CCND1	Amplification	Cell cycle/transcription	73, 123
11q-FGF3	Amplification	Growth factor	73
11q-FGF4	Amplification	Growth factor	73
20q-AIB1	Amplification	Transcription co-activator	124
20q-ZNF217	Amplification	Transcription factor	125
CDKN2A	Methylation	Cell cycle	126, 127
CCND2	Methylation	Cell cycle	128
SFN	Methylation	Cell cycle	77, 129
RARβ2	Methylation	Transcription factor	130
HIN-1	Methylation	Secreted growth inhibitor	76
BRCA1	Methylation	Recombination/transcription	131
GSTP1	Methylation	Metabolism	132
FABP3	Methylation	Secreted growth inhibitor	133
HOXA5	Methylation	Transcription factor	134
ARHI	Imprinting/LOH	Ras/Rap signalling	135

Abbreviations: AIB, amplified in breast cancer 1; ARHI, Ras homologue gene family member I; BRCA1, breast cancer 1; CCND1, cyclin D1; CCND2, cyclin D2; CDH1, E-cadherin; CDKN2A, cyclin-dependent kinase inhibitor 2A (p16); erbB2, erythroblastic leukemia viral oncogene homologue 2; FABP3, fatty-acid-binding protein 3; FGF 3, 4: fibroblast growth factor 3, 4; GST, glutathione transferase; HIN-1, 'high in normal' 1; HOXA5, homeo box A5; MKK4, mitogen-activated protein kinase kinase 4; myc, v-myc myelocytomatosis viral oncogene homologue; PTCH, patched; RARβ2, retinoic acid receptor β2; SFN, stratifin (14-3-3σ); ZNF217, zinc finger protein 217.

genes. In addition to the deletion mapping in cell lines, the investigators are also planning to sequence systematically all the coding exons and flanking splice junctions of every transcribed or predicted human gene in multiple tumour types including 16 primary breast carcinomas and matched normal genomic DNA (for details, see http://www.sanger.ac.uk/CGP/). The completion of this project could lead to the identification of many new breast cancer genes.

The use of somatic genetic changes to uncover new breast cancer genes is limited by the fact that, owing to genomic instability, tumours can have very high rates of somatic genetic events. Many

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

of these somatic changes might be random and not related to the tumourigenic process. The use of early-stage tumours, possibly multiple tumours from the same individual, can overcome this problem. The first recognised obligate precursor of invasive tumours is DCIS, as described above, but unfortunately even DCIS tumours have already accumulated multiple genetic alterations and some are aneuploid (Refs 68, 69). Earlier-stage tumours, such as usual or atypical ductal hyperplasia, are not easy to diagnose and their malignant potential and relationship to invasive carcinomas is unclear at this point.

Another complicating problem with searching for breast cancer genes based on LOH studies in tumours is the fact that certain chromosomal regions show frequent alterations that are not due to the genes located in the region but instead to fragile chromosome structure. The most common fragile sites in the human genome are 3p, 6q, 7q, 16q and 17q (Ref. 70). Even homozygous deletions, which are relatively rare and thought to be the clear indication of a tumour suppressor gene, can be misleading, since two recent studies reported polymorphic homozygous deletions without any specific gene targets (Refs 71, 72).

Despite all the potential problems associated with the analysis of somatic genetic alterations, one clear advantage of this approach is that it allows the identification of genes involved in tumour progression, including those required for invasion, angiogenesis and metastasis. In summary, although analyses of somatic genetic alterations have led to the identification of several breast cancer genes, it has thus far proven to be a fairly inefficient method. However, the use of new high-throughput sequencing approaches, like the one currently applied by the Cancer Genome Project at the Sanger Institute, will probably yield many new putative breast cancer genes in the near future.

Amplifications and epigenetic changes

Certain genes and chromosomal regions are amplified in breast tumours (Table 2). Increased gene copy number changes can be detected by CGH or as homogeneously staining (HS) regions (chromosomal regions that show unusual homogenous cytogenetic staining) or double minutes (DMs; very small accessory chromosomes) and they usually lead to the overexpression of genes included in the region of amplification (Ref. 73). Contrary to deletions, where homozygous inactivation of a gene is a definitive proof that the gene is the target of the deletion, in the case of amplifications it is frequently difficult to determine which gene is the critical target of this genetic event. This is due to multiple reasons: (1) amplifications are frequently large and include many genes; (2) the amplified gene might not always be overexpressed in the tumour; and (3) there might be a cluster of genes targeted by the event. Because of these facts, the above-described combined genotype/gene expression pattern analysis of breast tumours could be particularly useful for the identification of new genes targeted by amplifications. The best-characterised amplicons in breast cancer are on chromosome 8q, 11q, 17q and 20q (Ref. 73). For all of these amplicons, there are many candidate genes identified, such as *c-myc* (8q), *CCDN1* (11q; encoding cyclin D1), *FGF3* and *FGF4* (11q; encoding fibroblast growth factor 3 and 4), *HER2/erbB2* (17q; encoding erythroblastic leukemia viral oncogene homologue 2), *AIB1* (20q) and *ZNF217* (20q; encoding zinc finger protein 217).

In addition to genetic changes, epigenetic alterations including methylation or imprinting could also lead to the somatic inactivation of candidate tumour suppressor genes. Examples of genes showing epigenetic changes are listed in Table 2. Genes methylated in breast cancer could also be identified using comprehensive gene expression profiling (if the gene is consistently downregulated in tumours) or by arrays or subtractive hybridisation approaches specifically designed for the detection of methylated genes (Refs 74, 75). Examples of genes identified using the former approach – in this case, SAGE analysis of normal and cancerous mammary epithelial cells – are *HIN-1* ('high in normal 1') and *14-3-3σ* (encoding stratifin) (Refs 76, 77).

As demonstrated by the effect of *AIB1* on *BRCA1* penetrance, polymorphisms in amplified or epigenetically altered genes might influence breast cancer susceptibility, but not many such studies have been performed to date. In the case of amplified genes, sequence variants of the promoter region that alter expression, or in the coding region that alter activity, could modify breast cancer risk. Similarly, polymorphisms of CpG sites in the promoter of methylated genes could confer increased resistance to breast cancer. In addition, genes targeted by these events are excellent therapeutic or diagnostic targets, since

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

amplifications or overexpression usually result in gain-of-function phenotypes that are necessary for the tumour cells, whereas genes specifically methylated in cancers could be used for breast cancer diagnosis.

Animal models of breast cancer

The nearly completed mouse and rat genome sequences and the almost complete synteny between these rodent genomes and the human genome makes possible the use of animal models of breast cancer for the identification of new candidate cancer susceptibility genes. This is especially true in the mouse where genetic crosses can be performed to investigate complex genetic diseases. In the mouse, there are already several mammary tumour models that to some extent reflect the human disease (Ref. 78). The most promising approach for the generation of these models is the somatic random inactivation of tumour suppressor genes or the activation of oncogenes only in the desired cell type or tissue. An example of this is the generation of mammary-gland-specific knock-outs of *BRCA1* and *BRCA2* using cre recombinase under the control of the whey acidic protein (WAP) promoter (Refs 79, 80). These mouse models then can be used to search for modifiers of the penetrance of the phenotype.

However, a major problem with the use of mice as models of human breast cancer is that mouse strains are very inbred, and are therefore unlikely to reflect the highly variable human populations. In addition, the differing reproductive and environmental factors and lifespans of rodents and humans are problems that severely limit the usefulness of this approach for the identification of genes relevant to human disease. This is demonstrated by the use of mice that carry a mutated multiple intestinal neoplasia (*Min*) gene, which is the mouse homologue of the human *APC* gene, resulting in the development of intestinal tumours. The Min mouse is a fairly faithful model of human colon cancer and has therefore been used to identify *APC* modifier genes. However, although one of the putative modifiers identified ['modifier of Min' (MOM1), a secretory phospholipase Pla2g2a] strongly inhibited intestinal polyp formation in mice, it does not seem to play a role in human colorectal tumours (Refs 81, 82, 83).

Although rat strains are less inbred than mice, and rat mammary carcinomas are more similar to human breast cancers both in terms of histopathological progression and hormone responsiveness, many of the problems associated with the use of mice also hold true in rats (Ref. 84). In addition, rats are less amenable to genetic manipulation, and so most of the rat breast cancer models are based on chemical-induced carcinogenesis that is unlikely to reflect the human disease.

Thus, although rodent models of breast cancer have the potential to aid in the identification of new breast cancer genes, their usefulness remains to be proven. However, one definite advantage of animal models of cancer is that they can be used for the identification of new cancer preventive and therapeutic strategies that can be translated into human clinical trials.

Clinical implications and applications

What is the clinical importance of identifying breast cancer genes? In the case of inherited cancer susceptibility, identifying and testing the genes responsible is useful if the result of the test would change the clinical management of the patient (Ref. 85). However, this is true only if there are effective early diagnostic and intervention strategies available for the particular cancer type the patient is predisposed to. In the case of the known high-penetrance breast cancer susceptibility genes *BRCA1* and *BRCA2*, the value of genetic testing is still only presumed and not yet established (Ref. 86). Currently, the most effective preventive therapy of choice is bilateral mastectomy, but many women with familial risk of breast cancer and even proven mutation carriers do not want to undergo this treatment (Ref. 87). Therefore, new technologies for the early detection of, and new preventive therapies for, breast and ovarian cancer are needed.

A recently described proteomic pattern technology appears to be very promising since, in a preliminary study, it demonstrated 100% sensitivity and 95% specificity for the diagnosis of ovarian cancer (Ref. 88). The approach is based on surface-enhanced laser desorption/ionisation time-of-flight mass spectrometry (SELDI-TOF MS), which combines chromatography and mass spectrometry, together with bioinformatic analysis of serum from affected and unaffected patients, resulting in the definition of a 'normal' and 'cancer' pattern. Similar approaches are currently being developed for the diagnosis of breast and other cancer types.

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

Polymorphism screens for candidate low-penetrance breast cancer genes are currently performed only for research purposes. Genetic testing for multiple low-penetrance genes might eventually become reality, but only if accurate prediction is possible. The complexity of polygenic susceptibility and the possible combination of genotype with environmental and behavioural factors make the feasibility and usefulness of this type of test somewhat questionable (Ref. 89).

Ideally, understanding the genetic basis of breast cancer should lead to more-effective cancer prevention and treatment, since genetic defects are potential therapeutic targets. However, in many cases, genetic abnormalities are not tractable pharmaceutically. The analysis of tumour samples with associated clinical data using genomic technologies promises to be useful for the identification of new molecular targets that determine the clinical behaviour of the tumours. Many such studies on many different cancer types have been performed in the past few years, and many candidate targets have been identified. However, future studies are required to evaluate the clinical potential of these newly identified genes. The success of anti-oestrogen therapy (e.g. tamoxifen) for the prevention and treatment of oestrogen-receptor-positive breast tumours and the use of the recombinant humanised anti-Her2/ErbB2 monoclonal antibody trastuzumab (Herceptin®) for the treatment of Her2/ErbB2-positive tumours suggest that further molecular-based therapies will probably be identified and successfully used in the clinical management of breast cancer patients (Ref. 90).

Outstanding research questions and future prospects

Despite all of the efforts, the genetic basis of the majority of breast carcinomas is still poorly defined. However, it is clear that there is not a single breast cancer gene, an equivalent of *APC* in colorectal tumours, and that hereditary and sporadic breast tumours might have their own tumourigenic pathways. Therefore, the search must continue to find additional breast cancer genes using all the available traditional and new genomic techniques. The major challenge of identifying other breast cancer genes might be solved by using the new technical tools becoming available. However, the major obstacle has been, and continues to be, the translation of these findings into clinical practice. As demonstrated by the success of Herceptin® and STI571 (Gleevec™), an inhibitor of the Bcr-Abl tyrosine kinase for the treatment of chronic myelogenous leukaemia (Ref. 91), molecularly targeted cancer treatment is much more effective, with fewer side effects, than the other currently used anti-cancer therapies. Many new putative breast cancer therapeutic and diagnostic targets are being investigated in the clinic and many more are likely to follow. The identification of additional breast cancer susceptibility genes is sure to aid the identification of individuals at risk and the design of targeted cancer-preventive therapies.

Acknowledgements and funding

The author is funded by the National Cancer Institute SPORE (Specialized Program of Research Excellence) in Breast Cancer and Cancer Genome Anatomy Projects, the V Foundation, and the Dana-Farber Cancer Institute. I am grateful to Drs Ian Krop and Dale Porter (Dana-Farber Cancer Institute, Boston, MA, USA), as well as Annika Lindblom (Karolinska Institute, Sweden) and an anonymous reviewer, for their critical reading of this manuscript and for their helpful suggestions.

References

1 Greenlee, R.T. et al. (2000) Cancer statistics, 2000. CA Cancer J Clin 50, 7-33, PubMed ID: 20199188
2 Nathanson, K.L. et al. (2001) Breast cancer genetics: what we know and what we need. Nat Med 7, 552-556, PubMed ID: 21227090
3 Dunning, A.M. et al. (1999) A systematic review of genetic polymorphisms and breast cancer risk. Cancer Epidemiol Biomarkers Prev 8, 843-854, PubMed ID: 20014347
4 Weber, B.L. and Nathanson, K.L. (2000) Low penetrance genes associated with increased risk for breast cancer. Eur J Cancer 36, 1193-1199, PubMed ID: 20340770
5 Easton, D.F. (1999) How many more breast cancer predisposition genes are there? Breast Cancer Res 1, 14-17, PubMed ID: 21148258
6 Cui, J. et al. (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, 420-431, PubMed ID: 21090507
7 Antoniou, A.C. et al. (2002) A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. Br J Cancer 86, 76-83, PubMed ID: 21845539

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

8 Pharoah, P.D. et al. (1997) Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer 71, 800-809, PubMed ID: 97323953
9 Antoniou, A.C. et al. (2001) Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. Genet Epidemiol 21, 1-18, PubMed ID: 21336299
10 Pharoah, P.D. et al. (2002) Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31, 33-36, PubMed ID: 21980215
11 Houlston, R.S. and Tomlinson, I.P. (2000) Detecting low penetrance genes in cancer: the way ahead. J Med Genet 37, 161-167, PubMed ID: 20164776
12 Hall, J.M. et al. (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250, 1684-1689, PubMed ID: 91102542
13 Miki, Y. et al. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266, 66-71, PubMed ID: 95025896
14 Wooster, R. et al. (1995) Identification of the breast cancer susceptibility gene BRCA2. Nature 378, 789-792, PubMed ID: 96112016
15 Futreal, P.A. et al. (1994) BRCA1 mutations in primary breast and ovarian carcinomas. Science 266, 120-122, PubMed ID: 95025878
16 Lancaster, J.M. et al. (1996) BRCA2 mutations in primary breast and ovarian cancers. Nat Genet 13, 238-240, PubMed ID: 96225456
17 (1997) Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. Breast Cancer Linkage Consortium. Lancet 349, 1505-1510, PubMed ID: 97310292
18 Hedenfalk, I. et al. (2001) Gene-expression profiles in hereditary breast cancer. N Engl J Med 344, 539-548, PubMed ID: 21079642
19 Claus, E.B. et al. (1998) Effect of BRCA1 and BRCA2 on the association between breast cancer risk and family history. J Natl Cancer Inst 90, 1824-1829, PubMed ID: 99053258
20 Nathanson, K.L. and Weber, B.L. (2001) "Other" breast cancer susceptibility genes: searching for more holy grail. Hum Mol Genet 10, 715-720, PubMed ID: 21157257
21 Iau, P.T., Macmillan, R.D. and Blamey, R.W. (2001) Germ line mutations associated with breast cancer susceptibility. Eur J Cancer 37, 300-321, PubMed ID: 21134758
22 Seitz, S. et al. (1997) Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: linkage analysis in German breast cancer families. Oncogene 14, 741-743, PubMed ID: 97190227
23 Seitz, S. et al. (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, 983-991, PubMed ID: 98026370
24 Kainu, T. et al. (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, 9603-9608, PubMed ID: 20402581
25 Anbazhagan, R., Fujii, H. and Gabrielson, E. (1998) Allelic loss of chromosomal arm 8p in breast cancer progression. Am J Pathol 152, 815-819, PubMed ID: 98161528
26 Yaremko, M.L., Recant, W.M. and Westbrook, C.A. (1995) Loss of heterozygosity from the short arm of chromosome 8 is an early event in breast cancers. Genes Chromosomes Cancer 13, 186-191, PubMed ID: 95399365
27 Rahman, N. et al. (2000) Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. Oncogene 19, 4170-4173, PubMed ID: 20420243
28 Thompson, D. et al. (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, 827-831, PubMed ID: 21664364
29 Du, Q. et al. (2002) No evidence for a familial breast cancer susceptibility gene at chromosome 13q21 in Swedish breast cancer families. Int J Cancer 98, 799-800, PubMed ID: 21917739
30 Hemminki, A. et al. (1997) Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. Nat Genet 15, 87-90, PubMed ID: 97141925
31 Sjalander, A. et al. (1996) p53 polymorphisms and haplotypes in breast cancer. Carcinogenesis 17, 1313-1316, PubMed ID: 96281691
32 Durocher, F. et al. (1996) Comparison of BRCA1 polymorphisms, rare sequence variants and/or missense mutations in unaffected and breast/ovarian cancer populations. Hum Mol Genet 5, 835-842, PubMed ID: 96372821
33 Barker, D.F. et al. (1996) BRCA1 R841W: a strong candidate for a common mutation with moderate

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994 ©2002 Cambridge University Press

phenotype. Genet Epidemiol 13, 595-604, PubMed ID: 97123469
34 Healey, C.S. et al. (2000) A common variant in BRCA2 is associated with both breast cancer risk and prenatal viability. Nat Genet 26, 362-364, PubMed ID: 20517350
35 Bebb, G. et al. (1997) "AT risk" for breast cancer. Lancet 349, 1784-1785, PubMed ID: 97414515
36 Rebbeck, T.R. et al. (1999) Modification of BRCA1-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat. Am J Hum Genet 64, 1371-1377, PubMed ID: 99223430
37 Wang-Gohrke, S. et al. (2000) Progesterone receptor gene polymorphism is associated with decreased risk for breast cancer by age 50. Cancer Res 60, 2348-2350, PubMed ID: 20269243
38 Rebbeck, T.R. et al. (2001) Modification of BRCA1- and BRCA2-associated breast cancer risk by AIB1 genotype and reproductive history. Cancer Res 61, 5420-5424, PubMed ID: 21347215
39 Yan, H. et al. (2002) Small changes in expression affect predisposition to tumorigenesis. Nat Genet 30, 25-26, PubMed ID: 21624593
40 Yu, H. et al. (2001) Polymorphic CA repeats in the IGF-I gene and breast cancer. Breast Cancer Res Treat 70, 117-122, PubMed ID: 21619116
41 Altshuler, D. et al. (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. Nature 407, 513-516, PubMed ID: 20479973
42 Kwok, P.Y. (2001) Methods for genotyping single nucleotide polymorphisms. Annu Rev Genomics Hum Genet 2, 235-258, PubMed ID: 21557855
43 Hsu, T.M. et al. (2001) Universal SNP genotyping assay with fluorescence polarization detection. Biotechniques 31, 560, 562, 564-568, passim, PubMed ID: 21454231
44 Hsu, T.M. et al. (2001) Genotyping single-nucleotide polymorphisms by the invader assay with dual-color fluorescence polarization detection. Clin Chem 47, 1373-1377, PubMed ID: 21361255
45 Daly, M.J. et al. (2001) High-resolution haplotype structure in the human genome. Nat Genet 29, 229-232, PubMed ID: 21470358
46 Kruglyak, L. (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 22, 139-144, PubMed ID: 99295928
47 Thomson, G. (2001) Significance levels in genome scans. Adv Genet 42, 475-486, PubMed ID: 20488807
48 Lai, E. (2001) Application of SNP technologies in medicine: lessons learned and future challenges. Genome Res 11, 927-929, PubMed ID: 21275474
49 Miller, R.D. and Kwok, P.Y. (2001) The birth and death of human single-nucleotide polymorphisms: new experimental evidence and implications for human history and medicine. Hum Mol Genet 10, 2195-2198, PubMed ID: 21527005
50 Gray, I.C., Campbell, D.A. and Spurr, N.K. (2000) Single nucleotide polymorphisms as tools in human genetics. Hum Mol Genet 9, 2403-2408, PubMed ID: 20461130
51 Li, J. et al. (1997) PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science 275, 1943-1947, PubMed ID: 97228181
52 Hahn, S.A. et al. (1996) DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 271, 350-353, PubMed ID: 96144684
53 Liaw, D. et al. (1997) Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nat Genet 16, 64-67, PubMed ID: 97285123
54 Rhei, E. et al. (1997) Mutation analysis of the putative tumor suppressor gene PTEN/MMAC1 in primary breast carcinomas. Cancer Res 57, 3657-3659, PubMed ID: 97433097
55 Marsh, D.J. et al. (1997) Germline mutations in PTEN are present in Bannayan-Zonana syndrome. Nat Genet 16, 333-334, PubMed ID: 97385233
56 Vorechovsky, I., Benediktsson, K.P. and Toftgard, R. (1999) The patched/hedgehog/smoothened signalling pathway in human breast cancer: no evidence for H133Y SHH, PTCH and SMO mutations. Eur J Cancer 35, 711-713, PubMed ID: 99434653
57 Xie, J. et al. (1997) Mutations of the PATCHED gene in several types of sporadic extracutaneous tumors. Cancer Res 57, 2369-2372, PubMed ID: 97336016
58 Teng, D.H. et al. (1997) Human mitogen-activated protein kinase kinase 4 as a candidate tumor suppressor. Cancer Res 57, 4177-4182, PubMed ID: 97470610
59 Berx, G. et al. (1995) E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. Embo J 14, 6107-6115, PubMed ID: 96134909
60 Berx, G. and Van Roy, F. (2001) The E-cadherin/catenin complex: an important gatekeeper in

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

breast cancer tumorigenesis and malignant progression. Breast Cancer Res 3, 289-293, PubMed ID: 21481462
61 Guilford, P. et al. (1998) E-cadherin germline mutations in familial gastric cancer. Nature 392, 402-405, PubMed ID: 98196671
62 Chen, T., Sahin, A. and Aldaz, C.M. (1996) Deletion map of chromosome 16q in ductal carcinoma in situ of the breast: refining a putative tumor suppressor gene region. Cancer Res 56, 5605-5609, PubMed ID: 97126239
63 Salahshor, S. et al. (2001) Low frequency of E-cadherin alterations in familial breast cancer. Breast Cancer Res 3, 199-207, PubMed ID: 21204176
64 Graff, J.R. et al. (1995) E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. Cancer Res 55, 5195-5199, PubMed ID: 96062166
65 Pollack, J.R. et al. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 23, 41-46, PubMed ID: 99400550
66 Pinkel, D. et al. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20, 207-211, PubMed ID: 98442661
67 Snijders, A.M. et al. (2001) Assembly of microarrays for genome-wide measurement of DNA copy number. Nat Genet 29, 263-264, PubMed ID: 21547521
68 Polyak, K. (2001) On the birth of breast cancer. Biochim Biophys Acta 1552, 1-13, PubMed ID: 21639992
69 Polyak, K. (2002) Molecular alterations in ductal carcinoma in situ of the breast. Curr Opin Oncol 14, 92-96, PubMed ID: 21650473
70 Richards, R.I. (2001) Fragile and unstable chromosomes in cancer: causes and consequences. Trends Genet 17, 339-345, PubMed ID: 21273579
71 Lin, H, Pizer, E.S. and Morin, P.J. (2000) A frequent deletion polymorphism on chromosome 22q13 identified by representational difference analysis of ovarian cancer. Genomics 69, 391-394, PubMed ID: 20512045
72 Su, G.H. et al. (2002) Mutation rate of MAP2K4/MKK4 in breast carcinoma. Hum Mutat 19, 81, PubMed ID: 21623225
73 Schwab, M. (1998) Amplification of oncogenes in human cancer cells. Bioessays 20, 473-479, PubMed ID: 98364615
74 Dai, Z. et al. (2001) Global methylation profiling of lung cancer identifies novel methylated genes. Neoplasia 3, 314-323, PubMed ID: 21455584
75 Gitan, R.S. et al. (2002) Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. Genome Res 12, 158-164, PubMed ID: 21638322
76 Krop, I.E. et al. (2001) HIN-1, a putative cytokine highly expressed in normal but not cancerous mammary epithelial cells. Proc Natl Acad Sci U S A 98, 9796-9801, PubMed ID: 21396515
77 Ferguson, A.T. et al. (2000) High frequency of hypermethylation at the 14-3-3 sigma locus leads to gene silencing in breast cancer. Proc Natl Acad Sci U S A 97, 6049-6054, PubMed ID: 20283948
78 Deng, C.X. and Brodie, S.G. (2001) Knockout mouse models and mammary tumorigenesis. Semin Cancer Biol 11, 387-394, PubMed ID: 21446966
79 Xu, X. et al. (1999) Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. Nat Genet 22, 37-43, PubMed ID: 99251577
80 Ludwig, T. et al. (2001) Development of mammary adenocarcinomas by tissue-specific knockout of Brca2 in mice. Oncogene 20, 3937-3948, PubMed ID: 21385695
81 Riggins, G.J. et al. (1995) Absence of secretory phospholipase A2 gene alterations in human colorectal cancer. Cancer Res 55, 5184-5186, PubMed ID: 96062163
82 Cormier, R.T. et al. (1997) Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis. Nat Genet 17, 88-91, PubMed ID: 97434218
83 Kahlenberg, M.S. et al. (1998) Absence of frequent involvement of modifier of Min(APC) in sporadic colorectal cancer. Ann Surg Oncol 5, 181-185, PubMed ID: 98186502
84 Wood, G.A., Korkola, J.E. and Archer, M.C. (2002) Tissue-specific resistance to cancer development in the rat: phenotypes of tumor-modifier genes. Carcinogenesis 23, 1-9, PubMed ID: 21628851
85 Knudson, A.G., Jr. (1995) Hereditary cancers: from discovery to intervention. J Natl Cancer Inst Monogr 17, 5-7, PubMed ID: 96117091
86 (1996) Statement of the American Society of Clinical Oncology: genetic testing for cancer susceptibility, Adopted on February 20, 1996. J Clin Oncol 14, 1730-1736; discussion 1737-1740, PubMed ID: 96208855

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994 ©2002 Cambridge University Press

87 Eisinger, F. et al. (1999) Cultural basis for differences between US and French clinical recommendations for women at increased risk of breast and ovarian cancer. Lancet 353, 919-920, PubMed ID: 99191788
88 Petricoin, E.F. et al. (2002) Use of proteomic patterns in serum to identify ovarian cancer. Lancet 359, 572-577, PubMed ID: 21856877
89 Holtzman, N.A. and Marteau, T.M. (2000) Will genetics revolutionize medicine? N Engl J Med 343, 141-144, PubMed ID: 20331430
90 Gibbs, J.B. (2000) Mechanism-based target identification and drug discovery in cancer research. Science 287, 1969-1973, PubMed ID: 20185695
91 Druker, B. (2001) Signal transduction inhibition: results from phase I clinical trials in chronic myeloid leukemia. Semin Hematol 38, 9-14, PubMed ID: 21417855
92 Malkin, D. et al. (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250, 1233-1238, PubMed ID: 91057657
93 Hemminki, A. et al. (1998) A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. Nature 391, 184-187, PubMed ID: 98089041
94 Chen, J. and Lindblom, A. (2000) Germline mutation screening of the STK11/LKB1 gene in familial breast cancer with LOH on 19p. Clin Genet 57, 394-397, PubMed ID: 20309215
95 Forster, L.F. et al. (2000) An investigation of the Peutz-Jeghers gene (LKB1) in sporadic breast and colon cancers. J Clin Pathol 53, 791-793, PubMed ID: 20518279
96 Risinger, J.I. et al. (1996) Molecular genetic evidence of the occurrence of breast cancer as an integral tumor in patients with the hereditary nonpolyposis colorectal carcinoma syndrome. Cancer 77, 1836-1843, PubMed ID: 96208580
97 Borg, A. et al. (2000) Germline BRCA1 and HMLH1 mutations in a family with male and female breast carcinoma. Int J Cancer 85, 796-800, PubMed ID: 20175473
98 Swift, M. et al. (1991) Incidence of cancer in 161 families affected by ataxia-telangiectasia. N Engl J Med 325, 1831-1836, PubMed ID: 92072632
99 Bell, D.W. et al. (1999) Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. Science 286, 2528-2531, PubMed ID: 20085462
100 Krontiris, T.G. et al. (1993) An association between the risk of cancer and mutations in the HRAS1 minisatellite locus. N Engl J Med 329, 517-523, PubMed ID: 93330317
101 Ishibe, N. et al. (1998) Cigarette smoking, cytochrome P450 1A1 polymorphisms, and breast cancer risk in the Nurses' Health Study. Cancer Res 58, 667-671, PubMed ID: 98143550
102 Buchert, E.T. et al. (1993) Relationship of CYP2D6 (debrisoquine hydroxylase) genotype to breast cancer susceptibility. Pharmacogenetics 3, 322-327, PubMed ID: 94198889
103 Ladona, M.G. et al. (1996) CYP2D6 genotypes in Spanish women with breast cancer. Cancer Lett 99, 23-28, PubMed ID: 96152601
104 Feigelson, H.S. et al. (1997) A polymorphism in the CYP17 gene increases the risk of breast cancer. Cancer Res 57, 1063-1065, PubMed ID: 97219897
105 Millikan, R.C. et al. (1998) Cigarette smoking, N-acetyltransferases 1 and 2, and breast cancer risk. Cancer Epidemiol Biomarkers Prev 7, 371-378, PubMed ID: 98272316
106 Zheng, W. et al. (1999) N-acetyltransferase 1 genetic polymorphism, cigarette smoking, well-done meat intake, and breast cancer risk. Cancer Epidemiol Biomarkers Prev 8, 233-239, PubMed ID: 99188687
107 Hunter, D.J. et al. (1997) A prospective study of NAT2 acetylation genotype, cigarette smoking, and risk of breast cancer. Carcinogenesis 18, 2127-2132, PubMed ID: 98055590
108 Lavigne, J.A. et al. (1997) An association between the allele coding for a low activity variant of catechol-O-methyltransferase and the risk for breast cancer. Cancer Res 57, 5493-5497, PubMed ID: 98069846
109 Thompson, P.A. et al. (1998) Genetic polymorphisms in catechol-O-methyltransferase, menopausal status, and breast cancer risk. Cancer Res 58, 2107-2110, PubMed ID: 98266790
110 Ambrosone, C.B. et al. (1999) Manganese superoxide dismutase (MnSOD) genetic polymorphisms, dietary antioxidants, and risk of breast cancer. Cancer Res 59, 602-606, PubMed ID: 99137492
111 Mitrunen, K. et al. (2001) Association between manganese superoxide dismutase (MnSOD) gene polymorphism and breast cancer risk. Carcinogenesis 22, 827-829, PubMed ID: 21222981
112 Seth, P. et al. (2000) Phenol sulfotransferases: hormonal regulation, polymorphism, and age of onset of breast cancer. Cancer Res 60, 6859-6863, PubMed ID: 21028076

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994

113 Zheng, W. et al. (2001) Sulfotransferase 1A1 polymorphism, endogenous estrogen exposure, well- done meat intake, and breast cancer risk. Cancer Epidemiol Biomarkers Prev 10, 89-94, PubMed ID: 21087758
114 Helzlsouer, K.J. et al. (1998) Association between glutathione S-transferase M1, P1, and T1 genetic polymorphisms and development of breast cancer. J Natl Cancer Inst 90, 512-518, PubMed ID: 98198651
115 Kelsey, K.T. et al. (1997) Glutathione S-transferase class mu deletion polymorphism and breast cancer: results from prevalent versus incident cases. Cancer Epidemiol Biomarkers Prev 6, 511-515, PubMed ID: 97376166
116 Price, E.A. et al. (1997) Rare microsatellite polymorphisms in the DNA repair genes XRCC1, XRCC3 and XRCC5 associated with cancer in patients of varying radiosensitivity. Somat Cell Mol Genet 23, 237-247, PubMed ID: 98203653
117 Chouchane, L. et al. (1997) Polymorphism in the tumor necrosis factor-alpha promotor region and in the heat shock protein 70 genes associated with malignant tumors. Cancer 80, 1489-1496, PubMed ID: 97478179
118 Mestiri, S. et al. (2001) Genetic variation in the tumor necrosis factor-alpha promoter region and in the stress protein hsp70-2: susceptibility and prognostic implications in breast carcinoma. Cancer 91, 672-678, PubMed ID: 21135209
119 Andersen, T.I. et al. (1994) Oestrogen receptor (ESR) polymorphisms and breast cancer susceptibility. Hum Genet 94, 665-670, PubMed ID: 95080776
120 Chaudhuri, S. et al. (2000) Genetic susceptibility to breast cancer: HLA DQB*03032 and HLA DRB1*11 may represent protective alleles. Proc Natl Acad Sci U S A 97, 11451-11454, PubMed ID: 20481928
121 King, C.R., Kraus, M.H. and Aaronson, S.A. (1985) Amplification of a novel v-erbB-related gene in a human mammary carcinoma. Science 229, 974-976, PubMed ID: 85272597
122 Escot, C. et al. (1986) Genetic alteration of the c-myc protooncogene (MYC) in human primary breast carcinomas. Proc Natl Acad Sci U S A 83, 4834-4838, PubMed ID: 86259690
123 Schuuring, E. et al. (1992) Identification and cloning of two overexpressed genes, U21B31/PRAD1 and EMS1, within the amplified chromosome 11q13 region in human carcinomas. Oncogene 7, 355-361, PubMed ID: 92195667
124 Anzick, S.L. et al. (1997) AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. Science 277, 965-968, PubMed ID: 97400625
125 Collins, C. et al. (1998) Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. Proc Natl Acad Sci U S A 95, 8703-8708, PubMed ID: 98337979
126 Foster, S.A. et al. (1998) Inactivation of p16 in human mammary epithelial cells by CpG island methylation. Mol Cell Biol 18, 1793-1801, PubMed ID: 98187600
127 Herman, J.G. et al. (1995) Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. Cancer Res 55, 4525-4530, PubMed ID: 96006192
128 Evron, E. et al. (2001) Loss of cyclin D2 expression in the majority of breast cancers is associated with promoter hypermethylation. Cancer Res 61, 2782-2787, PubMed ID: 21182795
129 Umbricht, C.B. et al. (2001) Hypermethylation of 14-3-3 sigma (stratifin) is an early event in breast cancer. Oncogene 20, 3348-3353, PubMed ID: 21317138
130 Yang, Q. et al. (2001) Biallelic inactivation of retinoic acid receptor beta2 gene by epigenetic change in breast cancer. Am J Pathol 158, 299-303, PubMed ID: 20580917
131 Esteller, M. et al. (2000) Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. J Natl Cancer Inst 92, 564-569, PubMed ID: 20215032
132 Esteller, M. et al. (1998) Inactivation of glutathione S-transferase P1 gene by promoter hypermethylation in human neoplasia. Cancer Res 58, 4515-4518, PubMed ID: 99002641
133 Huynh, H., Alpert, L. and Pollak, M. (1996) Silencing of the mammary-derived growth inhibitor (MDGI) gene in breast neoplasms is associated with epigenetic changes. Cancer Res 56, 4865-4870, PubMed ID: 97050996
134 Raman, V. et al. (2000) Compromised HOXA5 function can limit p53 expression in human breast tumours. Nature 405, 974-978, PubMed ID: 20335889
135 Yu, Y. et al. (1999) NOEY2 (ARHI), an imprinted putative tumor suppressor gene in ovarian and breast carcinomas. Proc Natl Acad Sci U S A 96, 214-219, PubMed ID: 99093516

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994 ©2002 Cambridge University Press

Further reading, resources and contacts

Jong-Soo Lee and Jay H. Chung (2001) Diverse functions of BRCA1 in the DNA damage response. Exp. Rev. Mol. Med. 18 May, http://www.expertreviews.org/01003131h.htm

Meghan A. Arnold and Michael Goggins (2001) BRCA2 and predisposition to pancreatic and other cancers. Exp. Rev. Mol. Med. 14 May, http://www.expertreviews.org/0100309Xh.htm

The Wellcome Trust Fund Sanger Institute's website has details on the Cancer Genome Project:

http://www.sanger.ac.uk/CGP/

Comparative genomic hybridisation data on 38 human breast cancer cell lines are collated at the US National Human Genome Research Institute website:

http://www.nhgri.nih.gov/DIR/CGB/CR2000

Loss of heterozygosity studies in breast cancer are summarised at the University of Nottingham, UK, website:

http://www.nottingham.ac.uk/~pdzmgh/loh

Features associated with this article

Figures

Figure 1. A hypothetical model depicting the multistep process of breast tumourigenesis (fig001kpb).
Figure 2. Summary of approaches used for breast cancer gene discovery (fig002kpb).

Tables

Table 1. Genes with germline mutations associated with breast cancer susceptibility (tab001kpb).
Table 2. Genes with somatic genetic and epigenetic alterations in breast cancer (tab002kpb).

Citation details for this article

Kornelia Polyak (2002) Breast cancer gene discovery. Exp. Rev. Mol. Med. 15 August, http://www.expertreviews.org/0200491Xh.htm

Accession information: (02)00491-Xa.pdf (short code: txt001kpb); 15 August 2002
ISSN 1462-3994