A Systematic Review Of Genetic Polymorphisms and Breast Cancer Risk

Alison M. Dunning, Catherine S. Healey, Paul D. Pharoah, Bruce A. J. Ponder, and Douglas F. Easton

Cancer Research Campaign Human Cancer Genetics Group, University Department of Oncology [A. M. D., C. S. H., P. D. P. P., B. A. J. P.], and Cancer Research Campaign Genetic Epidemiology Group, University Department of Community Medicine [M. D. T., D. F. E.]; Strangeways Research Laboratories, Worts Causeway, Cambridge CB1 8RN, United Kingdom

Abstract

Studies investigating the relationship between common genetic variants and cancer risk are being reported with rapidly increasing frequency. We have identified 46 published case-control studies that have examined the effect of common alleles of 18 different genes on breast cancer risk. Of these, 12 report statistically significant associations, none of which were reported by more than one study. However, many of the studies were small: 10 of the 46 had 80% power or greater to detect a rare allele homozygote relative risk of 1.5 or greater. For the polymorphisms in TP53, the risk estimates, although nonsignificant, were of borderline statistical significance. For polymorphisms in BRCA1, COMT, CYP17, CYP1A1, NAT1, and NAT2, the best estimate of risk either from the individual studies or the meta-analyses was sufficiently precise to exclude a relative risk of 1.5 or greater. For the polymorphisms in EDH17B2, ER, CYP2D6, CYP2E1, GSTT1, GSTT2, and TNFRα, the risk estimates, although nonsignificant, were insufficiently precise to exclude a moderate risk (>1.5).

Precise estimation of the risks associated with these and other as yet untested genes, as well as investigation of more complex risks arising from gene-gene and gene-environment interactions, will require much larger studies.

Introduction

Breast cancer is a common disease in Western societies, with a lifetime prevalence of 1 in 12 in the United Kingdom (1) and 1 in 8 in the USA (2). Although 10–15% of breast cancer cases have some family history of the disease, only 5% can be explained by rare, highly penetrant mutations in genes such as BRCA1 and BRCA2 (3). First-degree relatives of breast cancer patients have a 2-fold increase in risk over the general population (4), most of which cannot be accounted for by BRCA1/2 (5). Although some of the familial risk may be due to shared environment, there may be other common, low-penetrance genetic variants which alter predisposition to breast cancer.

A frequently used experimental design for identifying common low-penetrance alleles is the association study (6). In this design, polymorphic genotype frequencies are compared between groups with different phenotypes. The aim is to study polymorphisms that may either be causally related to disease risk or are in strong linkage disequilibrium with disease-causing variants. Phenotypes investigated can be either continuously variable, such as serum lipid levels, or discrete, such as disease cases versus matched controls, in which instance the study design is a classic case-control study. The simplest polymorphisms to use are biallelic, which most commonly arise from a SNP, and give rise to three different genotype classes: the common allele homozygote, the heterozygote, and the rare allele homozygote. Multiallelic, repeat length polymorphisms (microsatellites), such as the (TTTA)n polymorphism in CYP19, can also be used, but these usually give rise to many genotypes. Genotypes often are grouped to simplify analysis, but this is valid only if a rational grouping strategy can be applied. Furthermore, the higher mutation rate in microsatellites is likely to lead to weaker associations unless the polymorphism itself is functional (e.g., the androgen receptor polyglutamine tract and prostate cancer risk). Genetic association studies using populations are more powerful than linkage studies within pedigrees for identifying low-penetrance alleles, which by definition may not be expressed in multiple members of a single family (7). However, at present, association studies can be carried out only on candidate genes.

There are a variety of ways of presenting gene polymorphism data in relation to breast cancer risk, depending on the nature of the polymorphism. In the case of simple biallelic polymorphisms, allele frequencies in cases and controls can be

Received 12/3/98; revised 6/22/99; accepted 7/12/99.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 This work was supported by the Cancer Research Campaign (CRC). B. A. J. P. is a Giff Fellow of the Cancer Research Campaign.

2 These authors contributed equally to this study.

3 To whom requests for reprints should be addressed, at CRC Human Cancer Genetics Group, University Department of Oncology, Strangeways Research Laboratories, Worts Causeway, Cambridge CB1 8RN, UK.

4 The abbreviations used are: SNP, single nucleotide polymorphism; OR, odds ratio; COMT, catechol-O-methyltransferase; GST, glutathione-S-transferase; NAT, N-acetyl transferase; UTR, untranslated region.
Review: Genetic Polymorphisms and Breast Cancer Risk

The first low-penetrance breast cancer susceptibility locus identified by the association approach was the HRAS1 minisatellite (8). This has more than 30 alleles, of which 4 are common, whereas the rest, which have a combined frequency of ~0.06, are classified as “rare.” These rare alleles have been found to be associated with a 1.9-fold increased relative risk of breast cancer (and increased risk of other cancers) and are estimated to account for 9% of all breast cancer incidence. However, the molecular mechanism underlying this association remains unclear. The HRAS1 locus has been reviewed extensively (8, 9) and so will not be considered further here.

Association studies have also been performed on other genes with putative involvement in breast cancer susceptibility. The candidate genes studies thus far can be divided into three main groups: genes for proteins with roles in steroid hormone metabolism; genes coding for carcinogen metabolism enzymes; and common alleles of genes that have been identified through family studies such as TP53 and BRCA1. The candidate gene polymorphisms reviewed here are listed in Table 1.

Steroid Hormone Metabolism Genes. Several factors alter exposure to endogenous hormones, and many of these, such as age at menarche, age at first pregnancy, number of pregnancies, and age at menopause, affect breast cancer risk (10). Hence, genes involved in the metabolism of sex hormones are strong

### Table 1 Genetic polymorphisms investigated in relation to breast cancer risk

<table>
<thead>
<tr>
<th>Gene</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Detection method</th>
<th>Functional effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steroid hormone metabolism genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMT</td>
<td>Exon 4 G → A</td>
<td>Val158Met</td>
<td>Creates Hpy99I and NcolI sites</td>
<td>Reduced activity</td>
</tr>
<tr>
<td>CYP17</td>
<td>Promoter T → C (T1931C)</td>
<td>None</td>
<td>Creates MspAI site</td>
<td>Creates a fifth SpI site and might increase transcription</td>
</tr>
<tr>
<td>CYP19</td>
<td>Intron 4 (TTTA) microsatellite</td>
<td>None</td>
<td>PCR fragment size</td>
<td>Unlikely</td>
</tr>
<tr>
<td>CYP2D6</td>
<td>Intron 3 G → A (G1934A) (B allele)</td>
<td>Frameshift</td>
<td>Creates BsrWI site</td>
<td>Nonfunctioning enzyme</td>
</tr>
<tr>
<td>Del Lys281 (C allele)</td>
<td></td>
<td>Premature stop at residue 544</td>
<td>Destroys BstNI site</td>
<td>Nonfunctioning enzyme</td>
</tr>
<tr>
<td>Other genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRCA1</td>
<td>C2731T</td>
<td>Pro871Leu</td>
<td>Allele-specific oligohybridization</td>
<td>Unknown</td>
</tr>
<tr>
<td>BRCA1</td>
<td>G1186A</td>
<td>Gln356Arg</td>
<td>Creates RsaI site</td>
<td>Unknown</td>
</tr>
<tr>
<td>HSP70-2</td>
<td>1267</td>
<td>Silent</td>
<td>Creates PstI site</td>
<td>Unknown</td>
</tr>
<tr>
<td>HSP70-hom</td>
<td>2437</td>
<td>Met493Thr</td>
<td>Creates Ncol site</td>
<td>Unknown</td>
</tr>
<tr>
<td>TNF-α</td>
<td>-308 G → A</td>
<td>None</td>
<td>Allele-specific PCR</td>
<td>Increased constitutive and inducible levels TNF-α</td>
</tr>
<tr>
<td>TP53</td>
<td>Exon 3 G → C</td>
<td>Arg72Pro</td>
<td>Allele-specific PCR</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>16-bp insertion in intron 3</td>
<td>None</td>
<td>Destroys TaqI site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td></td>
<td>Intron 6 G → A</td>
<td>None</td>
<td>Destroys BamHI site</td>
<td>Low activity allele</td>
</tr>
</tbody>
</table>

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Detection method</th>
<th>Functional effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edel (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP1A1</td>
<td>Exon 7 A → G (A4889G)</td>
<td>Ile462Val</td>
<td>Creates Ncol site</td>
<td>Uncertain, possible increase in enzyme activity</td>
</tr>
<tr>
<td>EDH17B2</td>
<td>Exon 6 A → G</td>
<td>Ser312Gly</td>
<td>Creates BspUII site</td>
<td>Unlikely</td>
</tr>
<tr>
<td>ER</td>
<td>CCC325CGG</td>
<td>Pro325Pro</td>
<td>Allele-specific oligohybridization</td>
<td>None</td>
</tr>
<tr>
<td>PR</td>
<td>Alu repeat insertion in intron G</td>
<td>None</td>
<td>PCR fragment size or creates TaqI site</td>
<td>Unlikely</td>
</tr>
</tbody>
</table>

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Detection method</th>
<th>Functional effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSTM1</td>
<td>Gene deletion</td>
<td></td>
<td>Allele-specific PCR</td>
<td>Null individuals have no enzyme</td>
</tr>
<tr>
<td>GSTP1</td>
<td>A313G</td>
<td>Ile105Val</td>
<td>Creates Aloc261 site</td>
<td>Reduced enzyme activity</td>
</tr>
<tr>
<td>GSTT1</td>
<td>Gene deletion</td>
<td></td>
<td>Allele-specific PCR</td>
<td>Null individuals have no enzyme activity</td>
</tr>
<tr>
<td>NAT1</td>
<td>A1088T</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>NAT2</td>
<td>G191A</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>CYP450A</td>
<td>G857A</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>GSTM1</td>
<td>Gene deletion</td>
<td></td>
<td>Allele-specific PCR</td>
<td>Null individuals have no enzyme activity</td>
</tr>
<tr>
<td>GSTP1</td>
<td>A313G</td>
<td>Ile105Val</td>
<td>Creates Aloc261 site</td>
<td>Reduced enzyme activity</td>
</tr>
<tr>
<td>GSTT1</td>
<td>Gene deletion</td>
<td></td>
<td>Allele-specific PCR</td>
<td>Null individuals have no enzyme activity</td>
</tr>
<tr>
<td>NAT1</td>
<td>A1088T</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>NAT2</td>
<td>G191A</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>CYP450A</td>
<td>G857A</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
</tbody>
</table>

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Detection method</th>
<th>Functional effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCRF1</td>
<td>C2731T</td>
<td>Pro871Leu</td>
<td>Allele-specific oligohybridization</td>
<td>Unknown</td>
</tr>
<tr>
<td>BRCA1</td>
<td>G1186A</td>
<td>Gln356Arg</td>
<td>Creates RsaI site</td>
<td>Unknown</td>
</tr>
<tr>
<td>HSP70-2</td>
<td>1267</td>
<td>Silent</td>
<td>Creates PstI site</td>
<td>Unknown</td>
</tr>
<tr>
<td>HSP70-hom</td>
<td>2437</td>
<td>Met493Thr</td>
<td>Creates Ncol site</td>
<td>Unknown</td>
</tr>
<tr>
<td>TNF-α</td>
<td>-308 G → A</td>
<td>None</td>
<td>Allele-specific PCR</td>
<td>Increased constitutive and inducible levels TNF-α</td>
</tr>
<tr>
<td>TP53</td>
<td>Exon 3 G → C</td>
<td>Arg72Pro</td>
<td>Allele-specific PCR</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>16-bp insertion in intron 3</td>
<td>None</td>
<td>Destroys TaqI site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td></td>
<td>Intron 6 G → A</td>
<td>None</td>
<td>Destroys BamHI site</td>
<td>Low activity allele</td>
</tr>
</tbody>
</table>

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Detection method</th>
<th>Functional effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edel (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP1A1</td>
<td>Exon 7 A → G (A4889G)</td>
<td>Ile462Val</td>
<td>Creates Ncol site</td>
<td>Uncertain, possible increase in enzyme activity</td>
</tr>
<tr>
<td>EDH17B2</td>
<td>Exon 6 A → G</td>
<td>Ser312Gly</td>
<td>Creates BspUII site</td>
<td>Unlikely</td>
</tr>
<tr>
<td>ER</td>
<td>CCC325CGG</td>
<td>Pro325Pro</td>
<td>Allele-specific oligohybridization</td>
<td>None</td>
</tr>
<tr>
<td>PR</td>
<td>Alu repeat insertion in intron G</td>
<td>None</td>
<td>PCR fragment size or creates TaqI site</td>
<td>Unlikely</td>
</tr>
</tbody>
</table>

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Detection method</th>
<th>Functional effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSTM1</td>
<td>Gene deletion</td>
<td></td>
<td>Allele-specific PCR</td>
<td>Null individuals have no enzyme</td>
</tr>
<tr>
<td>GSTP1</td>
<td>A313G</td>
<td>Ile105Val</td>
<td>Creates Aloc261 site</td>
<td>Reduced enzyme activity</td>
</tr>
<tr>
<td>GSTT1</td>
<td>Gene deletion</td>
<td></td>
<td>Allele-specific PCR</td>
<td>Null individuals have no enzyme activity</td>
</tr>
<tr>
<td>NAT1</td>
<td>A1088T</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>NAT2</td>
<td>G191A</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>CYP450A</td>
<td>G857A</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>GSTM1</td>
<td>Gene deletion</td>
<td></td>
<td>Allele-specific PCR</td>
<td>Null individuals have no enzyme activity</td>
</tr>
<tr>
<td>GSTP1</td>
<td>A313G</td>
<td>Ile105Val</td>
<td>Creates Aloc261 site</td>
<td>Reduced enzyme activity</td>
</tr>
<tr>
<td>GSTT1</td>
<td>Gene deletion</td>
<td></td>
<td>Allele-specific PCR</td>
<td>Null individuals have no enzyme activity</td>
</tr>
<tr>
<td>NAT1</td>
<td>A1088T</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>NAT2</td>
<td>G191A</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td>CYP450A</td>
<td>G857A</td>
<td>None</td>
<td>Creates Mpl site</td>
<td>Low activity allele</td>
</tr>
</tbody>
</table>

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Detection method</th>
<th>Functional effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCRF1</td>
<td>C2731T</td>
<td>Pro871Leu</td>
<td>Allele-specific oligohybridization</td>
<td>Unknown</td>
</tr>
<tr>
<td>BRCA1</td>
<td>G1186A</td>
<td>Gln356Arg</td>
<td>Creates RsaI site</td>
<td>Unknown</td>
</tr>
<tr>
<td>HSP70-2</td>
<td>1267</td>
<td>Silent</td>
<td>Creates PstI site</td>
<td>Unknown</td>
</tr>
<tr>
<td>HSP70-hom</td>
<td>2437</td>
<td>Met493Thr</td>
<td>Creates Ncol site</td>
<td>Unknown</td>
</tr>
<tr>
<td>TNF-α</td>
<td>-308 G → A</td>
<td>None</td>
<td>Allele-specific PCR</td>
<td>Increased constitutive and inducible levels TNF-α</td>
</tr>
<tr>
<td>TP53</td>
<td>Exon 3 G → C</td>
<td>Arg72Pro</td>
<td>Allele-specific PCR</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>16-bp insertion in intron 3</td>
<td>None</td>
<td>Destroys TaqI site</td>
<td>Low activity allele</td>
</tr>
<tr>
<td></td>
<td>Intron 6 G → A</td>
<td>None</td>
<td>Destroys BamHI site</td>
<td>Low activity allele</td>
</tr>
</tbody>
</table>
candidates for breast cancer susceptibility genes. Those in the sex hormone biosynthesis pathway may affect production of, and thus exposure to, the most active estrogen; estradiol. Genes in this pathway include CYP17, CYP19, and the gene for 17β-hydroxysteroid dehydrogenase type 2.

The bioavailability of hormones is partially controlled by catabolism, and catechol estrogens (2-hydroxy-estrogens) are the major breakdown products of estrogens. COMT is a phase II enzyme that methylates catechol-estrogens during their conjugation and inactivation. It has two forms: one membrane-bound and the other cytosolic; both are expressed in breast tissue and share a polymorphism associated with differences in methylation activity.

The sex hormones control the activation of responsive genes by first binding to specific receptors and forming complexes that can in turn bind to sequences in the promoters of downstream, hormone-responsive genes. Thus, steroid hormone receptor genes, such as ER, PR, and AR, are candidates for breast cancer susceptibility genes.

Carcinogen Metabolism Genes. Several enzymes have evolved for the detoxification of xenobiotic compounds, and their gene expression is induced in response to the presence of the compound (e.g., polycyclic aromatic hydrocarbons found in tobacco smoke). The actions of phase I and phase II enzymes render susceptible compounds more soluble and more readily excreted and ought to reduce cancer risk. However, the more soluble products of some compounds are even more potent carcinogens than the less soluble form. Hence, a genetic change that increases the expression of the gene or the activity of the protein produced may increase the amount of reactive carcinogen formed and, thus, increase the risk of cancer.

Two phase I enzymes, CYP1A1 and CYP2D6, are induced by, and act on, carcinogens found in tobacco smoke. Both have polymorphic differences in either inducibility or activity. CYP2E1, an enzyme that metabolizes ethanol, is also a candidate because epidemiological studies suggest that breast cancer risk is increased with alcohol consumption (11).

The GST family are phase II enzymes that detoxify carcinogens and their reactive intermediates, such as those produced by CYP1A1, by facilitating their conjugation to glutathione and subsequent excretion. For both GSTM1 and GSTT1 [reviewed by Rebbeck (12)], a high percentage of the Caucasian population are homozygous for null alleles (up to 60% and 20%, respectively) and have no detoxifying GST activity. Levels of DNA adducts, sister-chromatid-exchange, and somatic genetic mutations may be increased in carriers of GSTM1 and GSTT1 null genotypes (12).

The N-acetyl transferases, NAT1 and NAT2, are also phase II enzymes, and they participate in the detoxification of the arylamines, some of the main carcinogenic components of tobacco smoke, and also the amines produced during the cooking of meat (13, 14). However, the action of NATs on these carcinogens can produce electrophilic ions that may induce point mutations in DNA. Polymorphism in both genes results in two phenotypes: slow acetylators who are homozygous for low-activity alleles, and fast acetylators who carry one or more high-activity alleles.

Common Alleles of High-Penetration Genes. Mutations in the TP53 and BRCA1 genes are associated with a high lifetime risk of breast and other cancers (15, 16). TP53 is a tumor suppressor gene whose protein is produced in response to DNA damage through radiation or genotoxic agents, resulting in cell cycle arrest in G1 and induction of pathways leading to DNA repair or apoptosis. Mutation in the TP53 gene results in decreased p53 activity, which may lead to failure of cells with DNA damage to arrest and thus to continue to replicate with damaged DNA. In the case of BRCA1, where the protein function is still uncertain, the majority of confirmed mutations generate truncated proteins that are likely to have severely reduced activity. It has been hypothesized that amino acid substitutions outside the major functional domains may confer more moderate breast cancer risks. The majority of these substitutions are rare (17), and putative functional effects remain unconfirmed. However, it has been possible to test empirically the risks associated with the common polymorphisms.

Some of the genes discussed above have been assessed in multiple studies, whereas attempts to replicate the findings of other studies have yet to be reported. Here we have attempted to identify all of the published reports. Because most of these are based on small sample sizes, we have combined the results where possible in meta-analyses to obtain more precise estimates of risk. We discuss the implications of these findings and highlight areas where further work could be carried out profitably.

Materials and Methods

Literature Search. Published studies were identified using the Medline (National Library of Medicine, Washington DC) and BIDS databases for 1983–July 1998, using the search terms “breast-neoplasms” and “polymorphism(s).” We also searched for studies on each specific candidate gene: for example, for CYP17 the terms “CYP17” and “breast neoplasm” were used. In addition, the bibliographies of studies identified by the electronic searches were hand searched. Eligible studies were those that compared genotypic or allele frequencies of candidate genes in a series of breast cancer cases with a series of non-breast cancer controls using genomic DNA. Studies in which tumor DNA was used from cases with genomic DNA from controls were excluded. Studies of HRAS alleles and breast cancer risk were also excluded because these have been reviewed extensively by other authors (8, 9).

Meta-analysis. Whenever possible, raw data from comparable studies were analyzed jointly, using likelihood methods. Analyses were based on logistic regression and were carried out using the program S Plus. Each study was treated as a separate stratum, and the (control) allele or carrier frequencies in each study were permitted to be distinct, enabling studies with very different population frequencies to be considered jointly. For those metabolic polymorphisms associated with a specific phenotype, the principal analyses combined genotypes by phenotype classes (for example, poor metabolizer, slow acetylator). For all other polymorphisms, ORs were compared for each phenotype. Genotype-specific ORs were estimated assuming that the genotype frequencies in controls were consistent with Hardy-Weinberg equilibrium (18). Likelihood ratio tests (with degrees of freedom equal to number of genotypes − 1) were then used to test whether the ORs differed significantly from 1. If significant evidence of an association was found, a test for heterogeneity was performed. Ninety-five percent confidence intervals were calculated by direct examination of the likelihood surface. Joint analysis by subgroup (e.g., menopausal status) could only be performed if studies reported results for each subgroup. Specific analysis considering confounding factors such as age was not possible because raw data were not available.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Country</th>
<th>Cases</th>
<th>Controls</th>
<th>Ethnic group</th>
<th>Comments</th>
<th>Genes studied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agundez et al. 1995</td>
<td>Spain</td>
<td>160 NS&lt;sup&gt;19&lt;/sup&gt;</td>
<td>132 NS&lt;sup&gt;19&lt;/sup&gt;</td>
<td>White Spanish</td>
<td></td>
<td>NAT2</td>
</tr>
<tr>
<td>Ambrosone et al. 1995</td>
<td>USA</td>
<td>177 P</td>
<td>233 P</td>
<td>Post-men</td>
<td></td>
<td>CYP1A1</td>
</tr>
<tr>
<td>Ambrosone et al. 1996</td>
<td>USA</td>
<td>119 H</td>
<td>114 P</td>
<td>Pre-men</td>
<td></td>
<td>NAT2</td>
</tr>
<tr>
<td>Andersen et al. 1994</td>
<td>Norway</td>
<td>191 H</td>
<td>204 P</td>
<td>NS</td>
<td>Various sources of male and female controls used</td>
<td>ER</td>
</tr>
<tr>
<td>Bailey et al. 1998</td>
<td>USA</td>
<td>164 H</td>
<td>162 NS</td>
<td>Caucasian</td>
<td>Controls matched on age</td>
<td>CYP1A1, GSTM1</td>
</tr>
<tr>
<td>Ambrosone et al. 1996</td>
<td>USA</td>
<td>59 H</td>
<td>59 NS</td>
<td>African-American</td>
<td></td>
<td>GSTP1</td>
</tr>
<tr>
<td>Buchert et al. 1993</td>
<td>USA</td>
<td>167 H</td>
<td>114 H</td>
<td>25-61</td>
<td>Caucasian</td>
<td>CYP2D6</td>
</tr>
<tr>
<td>Campbell et al. 1996</td>
<td>UK</td>
<td>208 NS</td>
<td>113 NS</td>
<td>NS</td>
<td></td>
<td>P53</td>
</tr>
<tr>
<td>Charrier et al. 1999</td>
<td>France</td>
<td>361 P</td>
<td>437 P</td>
<td>18-101</td>
<td>Caucasian</td>
<td>GSTM1</td>
</tr>
<tr>
<td>Choussane et al. 1997</td>
<td>Tunisia</td>
<td>40 H</td>
<td>106 P</td>
<td>Mean = 43</td>
<td>Controls were healthy blood donors</td>
<td>TNF-α, HSπ-70</td>
</tr>
<tr>
<td>Dunning et al. 1997</td>
<td>UK</td>
<td>212 H</td>
<td>277 P</td>
<td>45-70</td>
<td>East-anglian case-control series</td>
<td>BRCA1</td>
</tr>
<tr>
<td>Dunning et al. 1998</td>
<td>UK</td>
<td>268 P&lt;sup&gt;36&lt;/sup&gt;</td>
<td>353 P</td>
<td>36-45</td>
<td>UK case-control study</td>
<td>CYP17</td>
</tr>
<tr>
<td>Feigelson et al. 1997</td>
<td>USA</td>
<td>174 P&lt;sup&gt;50&lt;/sup&gt;</td>
<td>285 P</td>
<td>45-75</td>
<td>Nested case-control study</td>
<td>CYP17</td>
</tr>
<tr>
<td>Garrett et al. 1995</td>
<td>Ireland</td>
<td>187 NS</td>
<td>90 NS</td>
<td>NS</td>
<td>Asian, Latino, African American</td>
<td>PR</td>
</tr>
<tr>
<td>Haiman et al. 1999</td>
<td>USA</td>
<td>464 P&lt;sup&gt;51&lt;/sup&gt;</td>
<td>619 P</td>
<td>NS</td>
<td>Nested case-control study within the Nurses’ Health Study</td>
<td>CYP19</td>
</tr>
<tr>
<td>Harries et al. 1997</td>
<td>Scotland</td>
<td>62 H</td>
<td>155 H</td>
<td>NS</td>
<td>Caucasian</td>
<td>GSTP1</td>
</tr>
<tr>
<td>Helzlsouer et al. 1998</td>
<td>USA</td>
<td>110 P&lt;sup&gt;52&lt;/sup&gt;</td>
<td>113 P</td>
<td>Mean = 60</td>
<td>Nested case-control study, controls matched for age,</td>
<td>GSTM1, GSTP1, GSTT1</td>
</tr>
<tr>
<td>Helzlsouer et al. 1998</td>
<td>USA</td>
<td>466 P&lt;sup&gt;53&lt;/sup&gt;</td>
<td>466 P</td>
<td>Mean = 61</td>
<td>race, and menopausal status</td>
<td></td>
</tr>
<tr>
<td>Hunter et al. 1997</td>
<td>USA</td>
<td>466 P&lt;sup&gt;54&lt;/sup&gt;</td>
<td>466 P</td>
<td>Mean = 61</td>
<td>Nested case-control study using incidental cases from the Nurses’ Health Study, Controls matched on age, menopausal status, and postmenopausal hormone use Phenotype study. Controls were patients with benign breast disease</td>
<td>NAT2, CYP1A1</td>
</tr>
<tr>
<td>Hunter et al. 1997</td>
<td>USA</td>
<td>466 P&lt;sup&gt;54&lt;/sup&gt;</td>
<td>466 P</td>
<td>Mean = 61</td>
<td>Nested case-control study using incidental cases from the Nurses’ Health Study, Controls matched on age, menopausal status, and postmenopausal hormone use Phenotype study. Controls were patients with benign breast disease</td>
<td>NAT2, CYP1A1</td>
</tr>
<tr>
<td>Hunter et al. 1997</td>
<td>USA</td>
<td>108 H&lt;sup&gt;56&lt;/sup&gt;</td>
<td>123 H</td>
<td>20-80</td>
<td>NS</td>
<td>CYP19</td>
</tr>
<tr>
<td>Iwase et al. 1996</td>
<td>Japan</td>
<td>93 H</td>
<td>347 NS</td>
<td>Japanese</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kelsey et al. 1997</td>
<td>USA</td>
<td>245 P&lt;sup&gt;57&lt;/sup&gt;</td>
<td>245 P</td>
<td>&gt;95% Caucasian</td>
<td>Prevalent cases</td>
<td>GSTM1</td>
</tr>
<tr>
<td>Kristensen et al. 1998</td>
<td>Norway/</td>
<td>366 P&lt;sup&gt;58&lt;/sup&gt;</td>
<td>252 PH&lt;sup&gt;PH&lt;/sup&gt;</td>
<td>NS</td>
<td>Incident cases</td>
<td></td>
</tr>
<tr>
<td>Lancaster et al. 1998</td>
<td>USA</td>
<td>98 H&lt;sup&gt;59&lt;/sup&gt;</td>
<td>446 NS</td>
<td>Mean = 25</td>
<td>Nested case-control study within Nurses’ Health Study, Controls matched on age and menopausal status</td>
<td>CYP2D6</td>
</tr>
<tr>
<td>Ladero et al. 1991</td>
<td>Spain</td>
<td>187 H&lt;sup&gt;60&lt;/sup&gt;</td>
<td>151 NS</td>
<td>18-83</td>
<td>Phenotype study</td>
<td>CYP2D6</td>
</tr>
<tr>
<td>Ladona et al. 1996</td>
<td>Spain</td>
<td>68 H&lt;sup&gt;61&lt;/sup&gt;</td>
<td>101 H</td>
<td>NS</td>
<td></td>
<td>PR</td>
</tr>
</tbody>
</table>
The Steroid Hormone Metabolism Genes. The COMT Val158Met polymorphism has been investigated in three studies. Thompson et al. (60), found a significantly increased risk of premenopausal breast cancer in COMT<sup>Met</sup> carrier women, with a nonsignificantly reduced risk of postmenopausal cancer. However, an earlier study had reported the opposite (although nonsignificant) effects for the same genotype, i.e., a reduced risk of premenopausal cancer and an increased risk of postmenopausal breast cancer (39), and a third study found no effect in unscreened cases (51). Conflicting results have also been reported in four studies of the 1931T>C polymorphism in the CYP17 promoter. Fiegelson et al. (30) found an increased risk of breast cancer for the CYP17<sup>C</sup> carrier in a subgroup analysis of 40 advanced cases, a finding not confirmed in three other studies (29, 35, 62). A possible role for CYP19 in breast cancer has been suggested by Haiman et al. (32) and Kristensen et al. (37) reported an increased risk for carriers of the (TTTA)<sub>12</sub> alleles; however, another group reported a statistically significant inverse association for this allele (55), and our own unpublished data have failed to confirm a significant risk (65). Haiman et al. also found a significantly increased risk for carriers of the rare (TTTA)<sub>10</sub> allele (32), but again this finding has not been confirmed by others (55, 65). Of two studies of the estrogen receptor gene polymorphism, CCC325CCG, one found a significantly elevated risk for the G carrier (42), but a subsequent, larger study failed to confirm this result (58).

**Results**

We identified 46 eligible studies (19–64), which are listed in Table 2. Although the basic study designs were all the same, a wide variety of sources of both cases and controls were used. Most reported studies were quite small: the median number of cases and controls combined was 391 (range, 58–1431).

The results of the individual studies are given in Table 3, which shows the relative risk of breast cancer associated with the different polymorphisms as estimated by the OR and 95% confidence interval. To present these data in a coherent and easily comparable manner, we have given the genotypic risks, where possible, we have estimated the appropriate ORs and confidence intervals from the published raw data.

Although 18 genes have been studied, few investigators have reported significant risks, and where significant results have been reported, these have often not been confirmed by other studies of the same polymorphism. The associations that have been reported as significant by at least one study are discussed below.

**Steroid Hormone Metabolism Genes.** The COMT Val158Met polymorphism has been investigated in three studies. Thompson et al. (60), found a significantly increased risk of premenopausal breast cancer in COMT<sup>Met</sup> carrier women, with a nonsignificantly reduced risk of postmenopausal cancer. However, an earlier study had reported the opposite (although nonsignificant) effects for the same genotype, i.e., a reduced risk of premenopausal cancer and an increased risk of postmenopausal breast cancer (39), and a third study found no effect in unscreened cases (51). Conflicting results have also been reported in four studies of the 1931T>C polymorphism in the CYP17 promoter. Fiegelson et al. (30) found an increased risk of breast cancer for the CYP17<sup>C</sup> carrier in a subgroup analysis of 40 advanced cases, a finding not confirmed in three other studies (29, 35, 62). A possible role for CYP19 in breast cancer has been suggested by Haiman et al. (32) and Kristensen et al. (37) reported an increased risk for carriers of the (TTTA)<sub>12</sub> alleles; however, another group reported a statistically significant inverse association for this allele (55), and our own unpublished data have failed to confirm a significant risk (65). Haiman et al. also found a significantly increased risk for carriers of the rare (TTTA)<sub>10</sub> allele (32), but again this finding has not been confirmed by others (55, 65). Of two studies of the estrogen receptor gene polymorphism, CCC325CCG, one found a significantly elevated risk for the G carrier (42), but a subsequent, larger study failed to confirm this result (58).
<table>
<thead>
<tr>
<th>Gene/polymorphism</th>
<th>“Rare” allele frequency in controls</th>
<th>Risk group</th>
<th>OR</th>
<th>95% CI</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Steroid hormone metabolism genes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMT Val158Met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lavigne et al. 1997 (39)</td>
<td>0.52</td>
<td>Val/Met het</td>
<td>1.30</td>
<td>0.69–3.07</td>
<td>ORs adjusted for age</td>
</tr>
<tr>
<td>Post-men cases</td>
<td>Val/Met het</td>
<td>0.57</td>
<td>0.14–2.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Met/Met het</td>
<td>0.24</td>
<td>0.04–1.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-men cases</td>
<td>Met/Met het</td>
<td>1.70</td>
<td>0.77–3.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Met/Met hom</td>
<td>2.18</td>
<td>0.93–5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thompson et al. 1998 (60)</td>
<td>0.43</td>
<td>Val/Met het</td>
<td><strong>2.5</strong></td>
<td><strong>1.4–4.6</strong></td>
<td></td>
</tr>
<tr>
<td>Met/Met hom</td>
<td><strong>2.2</strong></td>
<td><strong>1.3–3.7</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-men cases</td>
<td>Met carrier</td>
<td>0.3</td>
<td>0.3–1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-men cases</td>
<td>Met carrier</td>
<td>0.6</td>
<td>0.3–1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Millikan et al. 1998 (51)</td>
<td>0.35</td>
<td>Val/Met het</td>
<td>0.8</td>
<td>0.5–1.2</td>
<td></td>
</tr>
<tr>
<td>Met/Met hom</td>
<td>0.8</td>
<td>0.4–1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP17 promoter T → C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feigelson et al. 1997 (30)</td>
<td>0.41</td>
<td>C carrier</td>
<td>2.52</td>
<td><strong>1.07–5.94</strong></td>
<td></td>
</tr>
<tr>
<td>Advanced cases</td>
<td>C carrier</td>
<td>2.52</td>
<td><strong>1.07–5.94</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dunning et al. 1998 (29)</td>
<td>0.38</td>
<td>C carrier</td>
<td>1.57</td>
<td>0.85–2.91</td>
<td></td>
</tr>
<tr>
<td>Advanced cases</td>
<td>C carrier</td>
<td>1.57</td>
<td>0.85–2.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helalsouer et al. 1998 (35)</td>
<td>0.42</td>
<td>C carrier</td>
<td>1.76</td>
<td>0.99–3.14</td>
<td></td>
</tr>
<tr>
<td>Advanced cases</td>
<td>C carrier</td>
<td>1.76</td>
<td>0.99–3.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weston et al. 1998 (62)</td>
<td>0.38</td>
<td>C carrier</td>
<td>1.06</td>
<td>0.3–3.58</td>
<td></td>
</tr>
<tr>
<td>Aggressive disease</td>
<td>C carrier</td>
<td>1.06</td>
<td>0.3–3.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP19 (TTTA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healey et al. 1999 (55)</td>
<td>0.021</td>
<td>(TTTA)_{12} carrier</td>
<td>1.76</td>
<td>0.99–3.14</td>
<td></td>
</tr>
<tr>
<td>Kristensen et al. 1998 (37)</td>
<td>0.016</td>
<td>(TTTA)_{12} carrier</td>
<td>1.76</td>
<td>0.99–3.14</td>
<td></td>
</tr>
<tr>
<td>Siegelmann-Danieli et al. 1999 (55)</td>
<td>0.016</td>
<td>(TTTA)_{12} carrier</td>
<td>1.76</td>
<td>0.99–3.14</td>
<td></td>
</tr>
<tr>
<td>Haiman et al. 1999 (32)</td>
<td>0.020</td>
<td>(TTTA)_{12} carrier</td>
<td>1.76</td>
<td>0.99–3.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(TTTA)_{10} carrier</td>
<td>1.07</td>
<td>0.35–3.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Carrier frequency estimated from published allele frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDH17B2 Ser312Gly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mannermaa et al. 1994 (47)</td>
<td>0.47</td>
<td>Ser/Gly</td>
<td>0.62</td>
<td>0.36–1.07</td>
<td></td>
</tr>
<tr>
<td>Familial cases</td>
<td>Gly/Gly</td>
<td>0.68</td>
<td>0.4–1.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gly carrier</td>
<td>0.56</td>
<td>0.35–0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ser/Gly</td>
<td>0.44</td>
<td>0.11–1.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gly/Gly</td>
<td>1.02</td>
<td>0.19–5.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDH17B2 Ser312Gly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER XbaI RFLP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andersen et al. 1994 (22)</td>
<td>0.68</td>
<td>Hom</td>
<td>2.02</td>
<td>0.96–4.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Het</td>
<td>2.02</td>
<td>0.96–4.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XbaI carrier</td>
<td>2.02</td>
<td>0.96–4.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER CCC325CCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iwase et al. 1996 (42)</td>
<td>0.133</td>
<td>G allele carriers</td>
<td><strong>2.91</strong></td>
<td><strong>1.05–8.28</strong></td>
<td></td>
</tr>
<tr>
<td>Southey et al. 1998 (58)</td>
<td>0.209</td>
<td>GG het</td>
<td>1.09</td>
<td>0.79–1.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG hom</td>
<td>1.59</td>
<td>0.70–3.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G allele carrier</td>
<td>1.13</td>
<td>0.77–1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No G homozygotes in controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PR PROGINS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garrett et al. 1995 (31)</td>
<td>0.18</td>
<td>T1/T2</td>
<td>1.55</td>
<td>0.86–2.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T2/T2</td>
<td>0.40</td>
<td>0.06–2.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T2 carrier</td>
<td>1.39</td>
<td>0.70–2.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manolitsas et al. 1997 (48)</td>
<td>0.14</td>
<td>T2 carrier OR</td>
<td>0.77</td>
<td>0.50–1.18</td>
<td></td>
</tr>
<tr>
<td>Lancaster et al. 1998 (38)</td>
<td>0.13</td>
<td>T2 carrier OR</td>
<td>0.90</td>
<td>0.38–2.09</td>
<td></td>
</tr>
<tr>
<td>Carcinogen metabolism genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP1A1 Ile462Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambrosone et al. 1995 (20)</td>
<td>0.077</td>
<td>Ile/Val het</td>
<td>1.53</td>
<td>0.88–2.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ile/Val hom</td>
<td>2.85</td>
<td>0.49–16.56</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3  
Continued

<table>
<thead>
<tr>
<th>Gene/polymorphism</th>
<th>Case details</th>
<th>“Rare” allele frequency in controls</th>
<th>Risk group</th>
<th>OR</th>
<th>95% CI</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taioli et al. 1995 (59)</td>
<td>All cases</td>
<td>0.091</td>
<td>Ile/Val het</td>
<td>1.1</td>
<td>0.3–4.0</td>
<td>No Val/Val homozygous cases</td>
</tr>
<tr>
<td>Ishibe et al. 1998 (41)</td>
<td>All cases</td>
<td>NS</td>
<td>Val carrier</td>
<td>0.88</td>
<td>0.58–1.33</td>
<td>Multivariate-adjusted OR</td>
</tr>
<tr>
<td>Bailey et al. 1998 (23)</td>
<td>All Caucasian</td>
<td>0.037</td>
<td>Ile/Val het</td>
<td>1.26</td>
<td>0.57–2.78</td>
<td>No Val/Val homozygotes</td>
</tr>
<tr>
<td>CYP1A1 3' UTR T6235C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taioli et al. 1995 (59)</td>
<td>Caucasian</td>
<td>0.115</td>
<td>TC het</td>
<td>1.7</td>
<td>0.6–4.9</td>
<td></td>
</tr>
<tr>
<td>Ishibe et al. 1998 (41)</td>
<td>All cases</td>
<td>NS</td>
<td>C carrier</td>
<td>1.05</td>
<td>0.74–1.50</td>
<td>Multivariate-adjusted OR</td>
</tr>
<tr>
<td>Bailey et al. 1998 (23)</td>
<td>Caucasian</td>
<td>0.096</td>
<td>TC het</td>
<td>1.49</td>
<td>0.81–2.74</td>
<td></td>
</tr>
<tr>
<td>CYP1A1 3' UTR T5639C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taioli et al. 1995 (59)</td>
<td>African American</td>
<td>0.081</td>
<td>TC het</td>
<td>1.2</td>
<td>0.3–5.3</td>
<td>No CC homozygotes</td>
</tr>
<tr>
<td>Bailey et al. 1998 (23)</td>
<td>Caucasian</td>
<td>0.119</td>
<td>TC het</td>
<td>0.75</td>
<td>0.31–1.83</td>
<td>No CC homozygotes</td>
</tr>
<tr>
<td>CYP2E1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shields et al. 1996 (54)</td>
<td>Pre-men</td>
<td>0.09</td>
<td>C carrier</td>
<td>1.04</td>
<td>0.48–2.24</td>
<td></td>
</tr>
<tr>
<td>Ambrosone et al. 1995 (20)</td>
<td>Post-men cases</td>
<td>0.502</td>
<td>Null</td>
<td>1.10</td>
<td>0.73–1.64</td>
<td>Adjusted for age, age at menarche/first pregnancy/ menopause, BMI, and family history</td>
</tr>
<tr>
<td>GSTM1 deletion</td>
<td>Zhong et al. 1993 (64)</td>
<td>All cases</td>
<td>0.418</td>
<td>Null</td>
<td>1.27</td>
<td>0.87–1.87</td>
</tr>
<tr>
<td></td>
<td>Ambrosone et al. 1995 (20)</td>
<td>Post-men cases</td>
<td>0.502</td>
<td>Null</td>
<td>1.10</td>
<td>0.73–1.64</td>
</tr>
<tr>
<td></td>
<td>Kelsey et al. 1997 (44)</td>
<td>Prevalent cases</td>
<td>0.51</td>
<td>Null</td>
<td>1.30</td>
<td>0.91–1.86</td>
</tr>
<tr>
<td></td>
<td>Bailey et al. 1998 (23)</td>
<td>Caucasian</td>
<td>0.383</td>
<td>Null</td>
<td>1.08</td>
<td>0.74–1.57</td>
</tr>
<tr>
<td></td>
<td>African-American</td>
<td>0.407</td>
<td>Null</td>
<td>0.77</td>
<td>0.50–1.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Helzlsouer et al. 1998 (34)</td>
<td>All cases</td>
<td>0.464</td>
<td>Null</td>
<td>2.10</td>
<td>1.22–3.64</td>
</tr>
<tr>
<td></td>
<td>Pre-men cases</td>
<td>1.0</td>
<td></td>
<td></td>
<td>0.28–3.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-men cases</td>
<td>2.50</td>
<td></td>
<td></td>
<td>1.34–4.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Charrier et al. 1999 (26)</td>
<td>All cases</td>
<td>0.51</td>
<td>Null</td>
<td>1.19</td>
<td>0.90–1.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;50 years</td>
<td>0.51</td>
<td>Null</td>
<td>0.97</td>
<td>0.69–1.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥50 years</td>
<td></td>
<td></td>
<td>1.99</td>
<td>1.19–3.32</td>
</tr>
<tr>
<td>GSTT1 deletion</td>
<td>Helzlsouer et al. 1998 (34)</td>
<td>All</td>
<td>0.214</td>
<td>Null</td>
<td>1.50</td>
<td>0.76–2.95</td>
</tr>
<tr>
<td></td>
<td>Pre-men</td>
<td>1.50</td>
<td></td>
<td></td>
<td>0.42–5.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-men</td>
<td>1.50</td>
<td></td>
<td></td>
<td>0.67–3.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bailey et al. 1998 (23)</td>
<td>Caucasian</td>
<td>0.272</td>
<td>Null</td>
<td>1.08</td>
<td>0.66–1.75</td>
</tr>
<tr>
<td></td>
<td>African-American</td>
<td>0.288</td>
<td>Null</td>
<td></td>
<td>0.63</td>
<td>0.27–1.47</td>
</tr>
<tr>
<td>GSTP1 Ile105Val</td>
<td>Helzlsouer et al. 1998 (34)</td>
<td>All</td>
<td>0.29</td>
<td>Ile/Val het</td>
<td>1.48</td>
<td>0.81–2.73</td>
</tr>
<tr>
<td></td>
<td>Pre-men</td>
<td>1.98</td>
<td></td>
<td></td>
<td>0.44–8.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-men</td>
<td>1.33</td>
<td></td>
<td></td>
<td>0.68–2.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>All</td>
<td>Val/Val hom</td>
<td>1.97</td>
<td>0.77–5.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pre-men</td>
<td>0.54</td>
<td></td>
<td></td>
<td>0.04–6.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-men</td>
<td>2.71</td>
<td></td>
<td></td>
<td>0.91–8.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Harries et al. 1997 (33)</td>
<td>All</td>
<td>0.28</td>
<td>Ile/Val het</td>
<td>1.53</td>
<td>0.83–2.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Val/Val hom</td>
<td>1.58</td>
<td>0.49–5.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAT2 A1088T</td>
<td>Millikan et al. 1998 (50)</td>
<td>Pre-men cases</td>
<td>0.47</td>
<td>T carrier</td>
<td>1.2</td>
<td>0.8–1.8</td>
</tr>
<tr>
<td></td>
<td>Post-men cases</td>
<td>0.46</td>
<td>T carrier</td>
<td>1.0</td>
<td>0.7–1.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zheng et al. 1999 (63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
et al. (59), who found an increased breast cancer risk for African-American women who were CC homozygote for the 3'UTR T→C (T6235C). In contrast, the other study of African-American women found a nonsignificantly reduced breast cancer risk associated with this genotype (23). Of five studies of the GSTM1 gene deletion, only one found that null individuals were at significantly increased risk (34), and in subgroup analysis, this effect was restricted to postmenopausal breast cancer. Charrier et al. (26) also found a significantly increased risk of breast cancer in the subgroup of patients diagnosed over the age of 50.

Other Genes. A significant effect of the Arg72Pro polymorphism in TP53 has been reported by Sjalander et al. (56), who found that TP53 Arg/Pro heterozygotes were at increased risk, as
were TP53<sup>Pro</sup> carriers. Pro homozygotes were also at increased risk, although the confidence intervals were wide and the effect was not statistically significant. Two other studies of the same polymorphism had found no significant effect (43, 61). The genes <em>HSP-70</em> and <em>TNF-α</em> have been investigated by one group, which found significantly increased risk for rare allele carriers of both genes (see Table 3). However, this study was small, and their results have not been confirmed or refuted by other groups.

**Meta-analysis.** Comparable data from more than one study that could be combined were available for 25 different combinations of polymorphism and patient subgroups. The results of these analyses are given in Table 4. Statistically significant differences in genotype frequencies were found in three case-control comparisons of unselected cases. These were for CYP19 (TTTA)<sub>n</sub> polymorphism ([TTTA]<sub>10</sub> carrier OR = 2.33; <em>P</em> = 0.002), the GSTP1 Ile105Val polymorphism (Val carrier OR = 1.60; <em>P</em> = 0.02), and the TP53 Arg72Pro (Pro carrier OR = 1.27; <em>P</em> = 0.03). In addition, the GSTM1 gene deletion was found to be significant in postmenopausal breast cancer (null homozygote OR = 1.33; <em>P</em> = 0.04). There was also some evidence that homozygotes for the <em>PR PROGINS</em> allele are protected against breast cancer, although this result was of borderline statistical significance.

**Gene-Environment Interaction.** Although there is little epidemiological evidence that cigarette smoking is associated with breast cancer, studies have assessed the possible interaction between polymorphisms in the genes that encode carcinogen-metabolizing enzymes, CYP1A1, NAT2, CYP2E1, and GSTM1, with smoking (20, 21, 40, 41, 44, 50, 54). None of these studies found an overall increase in breast cancer risk with the variant metabolizing enzymes, but some reported interactions with smoking. Ambroson et al. (20) found an increase in risk in “light” smokers who carried the Val462 allele. In contrast, Ishibe et al. (41) found no interaction with the Ile462Val polymorphism but did find an increased risk in current smokers who were C carriers of the T6235C polymorphism. For NAT2, Ambroson et al.
(21) found an association between smoking >15 cigarettes per day and risk that was limited to slow acetylators, but two subsequent larger studies found that cigarette smoking was not appreciably associated with breast cancer among either fast or slow acetylators (40, 50). CYP2E1 genotype and smoking has been investigated in one study (54), which found a risk of premenopausal breast cancer in smokers who were also carriers of the DraI cutting site in intron 6. However, this result was only statistically significant after adjustment for other risk factors, and the confidence intervals were wide. A lack of interaction between GSTM1 genotype and cigarette smoking has been reported by Kelsey et al. (44).

Lavigne et al. (39) looked for interactions between the Val158 Met polymorphism in the COMT gene and several factors including family history, menopausal status, smoking, alcohol, oral contraceptive use, and hormone replacement therapy. Helzlsouer et al. (34) also explored the possibility of interactions between GSTM1, GSTP1, and GSTTI genotypes with family history, hormone replacement therapy, cigarette smoking, alcohol consumption, and body mass index. Although no overall effect was found in either study, both reported interactions with menopausal status and body mass index, and Helzlsouer et al. (34) additionally noted an interaction between GSTTI and alcohol consumption. However, no studies attempting to confirm these quite complex interactions have yet been published.

**Gene-Gene Interaction.** The possibility of gene-gene interactions have rarely been explored. One small study carried out genotyping in three genes: CYP1A1, GSTM1, and GSTTI (23). No interaction between these genes was found. The studies of Helzlsouer et al. (34) on the GST genes and Lavigne et al. (39) on COMT were carried out on the same populations. The increased relative risk of postmenopausal breast cancer for GSTM1<sup>Null</sup> and GSTP1<sup>Pro/Pro</sup>(GSTP1<sup>NVal</sup>) women of around 2 was found to be increased to a 3–4-fold increased risk if they were also COMT<sup>Met/Met</sup>. Again, no confirmatory studies have yet been published.

**Discussion**

We have found a substantial number of reports of studies that have investigated candidate genes for low-penetrance breast cancer susceptibility alleles. Despite this research effort, there is no clear evidence that any of these polymorphisms are strongly associated with breast cancer risk. Of the individual studies, few of the reported associations have been statistically significant, and no significant association has been reported by more than one study. In some cases, this might be due to a lack of statistical power in individual studies. Even among the significant associations, the magnitude of the effect found is rarely greater than 2.5-fold increased risk. If the rare allele frequency is 0.2, 315 cases and 315 controls would be required to detect this magnitude of risk for a rare allele homozygote with 90% power at the 5% significance level: 10 of the 46 studies reported were larger than this.

We have attempted to reduce this lack of power by combining results in meta-analyses, but even the results of these need to be interpreted with some caution. Of the 25 associations tested in at least two studies, 4 were significant at the 5% level. One of these was significant only in a subgroup of cases (postmenopausal) and was not significant overall. Such sub-group-specific associations must always be treated with caution given that there is no clear a priori reason to suspect a subgroup effect. Moreover, only one of the associations has a significance levels of <1%. Given the number of polymorphisms being tested, a significance level of $10^{-4}$ or smaller would be required to provide strong evidence.

In any systematic review, a major cause for concern is the potential for publication bias. The most common scenario is the nonpublication of negative studies (i.e., those finding no significant association), resulting in bias of any meta-analysis away from the null. We cannot exclude this possibility for the three polymorphisms with significant results in the meta-analyses. For the polymorphisms for which we have found no evidence for an association between genotype and breast cancer risk, it is possible, but unlikely, that bias toward the null has occurred. There is an urgent need for databases into which the results of all association studies (positive and negative) with candidate genes can be entered to minimize the effects of publication bias.

Consistency of reporting added further complications to the meta-analysis: many studies do not describe the ethnicity of their study populations; therefore, we have combined samples on the assumption that any variant studied will have the same effect in all populations. This assumption may only be valid if the variant studied is truly functional with respect to breast cancer risk. If the variant is simply a neutral marker for some other functional variant, the assumption may be invalid because linkage disequilibrium relationships often differ between populations. In addition, there is no consensus nomenclature for SNPs: early reports often described the restriction enzyme site involved, but there are instances of multiple polymorphisms detected by the same restriction enzyme within a single gene (e.g., there are two polymorphisms creatingMspI sites in CYP1A1). More recent studies usually identify the base substitution, but in introns and UTRs of genes, this can also be arbitrary. In Table 1, we presented the SNP and its general position (e.g., exon 3, intron 6, 3'UTR) for clarity, but workers will need to refer to the original articles for sufficient detail to replicate the DNA assays. The requirement for public databases of SNPs has recently been recognized (66), and it is hoped these will then provide a standardized description. Furthermore, methodological difficulties have been caused by genes having several different polymorphisms in linkage disequilibrium with one another. For example, CYP1A1 has two SNPs in strong linkage disequilibrium with one another; these are the exon 7 Ile462Val and the 3'UTR T→C (T6253C) SNPs. Both have been shown to be significantly associated with risk of some neoplasms (although not breast cancer), but combining them in a meta-analysis would require a somewhat more complicated analytical approach.

Several studies have reported on putative gene-gene and gene-environment interactions. The results of these analyses should, however, be treated with caution. The problem of post hoc subgroup analyses and multiple hypothesis testing renders the interpretation of positive results difficult. Negative results also should be treated with caution because few studies will have had sufficient power to detect moderate interaction effects. For example, 1500 cases and 1500 controls would be required to detect an interaction between two genotypes each with a frequency of 0.1 and a relative risk of 1.1 but with a relative risk of 2.5 when combined. These analyses should, therefore, be treated as hypothesis generating rather than hypothesis testing.

Despite these concerns, we believe some firm conclusions can be drawn. For several polymorphisms, the best estimate of risk either from the individual studies or the combined meta-analyses is sufficiently precise to exclude a relative risk of 1.5. These include the polymorphisms in BRCA1, COMT, CYP17, CYP1A1, NATH1, and NATH2. This does not necessarily imply...
that a negligible fraction of breast cancer incidence is attributable to such genes. For example, the upper confidence interval for the CYP17 effect (1.39) would still correspond to a population-attributable fraction of ~20%. We have, however, concluded that these polymorphisms contribute little to the familial aggregation of breast cancer: the same effect would correspond to a relative risk of siblings of cases of ~1.01 [for details of calculation, see Easton (67)], whereas epidemiological studies have found the relative risk to be ~2-fold.

For other polymorphisms, the risk estimates, although nonsignificant, are insufficiently precise to exclude a moderate risk (>1.5), and larger studies are needed to obtain more precise risk estimates. These include polymorphisms in EDH17B2, ER, CYP2D6, CYP2E1, GSTT1, HSP70, PR, and TNFα. Finally, the polymorphisms in CYP19, GSTM1, GSTP1, and TP53 appear to be stronger candidates for low-penetration breast cancer susceptibility genes, although they too need to be confirmed in larger studies. It is more likely, however, that the majority in variation in susceptibility to breast cancer is due to genes that have yet to be identified or tested. Candidate genes include those involved in DNA repair and micro- and macro-nutrient metabolism. In addition, gene-gene and gene-environment interactions may be important determinants of breast cancer risk. Such interactions would produce substantially increased risks in individuals with the right combination of factors, but large studies would be required to elucidate these effects. Further work is clearly needed to address these issues, but these studies will need to be substantially larger than the association studies published to date.

Acknowledgments

We thank Katie Blyth (Department of Pathology, University of Cambridge) for help in obtaining references.

References


