Cytochrome P-450 1A1 Gene Polymorphisms and Risk of Breast Cancer: A HuGE Review


From the Epidemiology Group, Department of Public Health, University of Aberdeen, Aberdeen, Scotland.

Received for publication July 21, 2004; accepted for publication January 4, 2005.

Cytochrome P-450 (CYP) 1A1 plays a key role in phase I metabolism of polycyclic aromatic hydrocarbons and in estrogen metabolism. It is expressed predominantly in extrahepatic tissues, including the breast. Four CYP1A1 gene polymorphisms (3801T/C, Ile462Val, 3205T/C, and Thr461Asp) have been studied in relation to breast cancer. The 3801C variant is more common than the Val variant. Both variants occur more frequently in Asians than in White populations. The 3205T→C polymorphism has been observed in African Americans only. Little data are available on the geographic/ethnic distribution of the Thr461Asp polymorphism. The functional significance of the polymorphisms is unclear. In 17 studies, no consistent association between breast cancer and CYP1A1 genotype was found. Meta-analysis found no significant risk for the genotypes 1) 3801C/C (relative risk (RR) = 0.97, 95% confidence interval (CI): 0.52, 1.80) or 3801T/C (RR = 0.91, 95% CI: 0.70, 1.19) versus 3801T/T, 2) Val/Val (RR = 1.04, 95% CI: 0.63, 1.74) or Ile/Ile (RR = 0.92, 95% CI: 0.76, 1.10) versus Ile/Ile, or 3) Asp/Asp (RR = 0.95, 95% CI: 0.20, 4.49) or Thr/Thr (RR = 1.12, 95% CI: 0.87, 1.43) versus Thr/Thr. Future studies should explore possible interactions between CYP1A1 and sources of polycyclic aromatic hydrocarbons, markers of estrogen exposure, other lifestyle factors influencing hormonal levels, and other genes involved in polycyclic aromatic hydrocarbon metabolism or hormonal biosynthesis.
aryl hydrocarbon receptor, a transcription factor that regulates gene expression (6).

CYP1A1 expression occurs predominantly in extrahepatic tissue (7). CYP1A1 messenger RNA has been detected in normal and cancerous breast tissue (8, 9) and can be induced in human-breast-derived cell lines (6).

GENE VARIANTS

Several mutations in CYP1A1 have been described (for CYP1A1 allele nomenclature, refer to the following website: http://www.imm.ki.se/CYPalleles), and four polymorphisms have been studied in relation to breast cancer. Table 1 describes these four polymorphisms and the allele nomenclature system (10–13). The 3801T→C (14, 15) and 3205T→C (16) polymorphisms are located in the 3' non-coding region. The 2455A→G (17) and 2453C→A (13) polymorphisms arise close together in exon 7 and result in amino acid changes Ile462Val and Thr461Asp, respectively. Because the studies did not always include information on all polymorphisms, it was often not possible to identify which of the *2A, *2B, or *2C alleles were present. Therefore, the “3801T→C,” “Ile462Val,” “3205T→C,” and “Thr461Asp” nomenclature is used throughout this review.

Genotype frequencies

In 2001, Garte et al. (18) estimated CYP1A1*2A, CYP1A1*2B, CYP1A1*2C, and CYP1A1*3 genotype frequencies in Whites, Asians, and Africans by using data from 33 studies of Whites, nine studies of Asians, and five studies of Africans. In comparison, the present review includes data from 69 articles, including 20 studies published between 2002 and 2004, and also summarizes data for the Thr461Asp polymorphism.

Relevant papers were identified by searching MEDLINE and EMBASE from 1980 to week 4 of 2004 by using the MeSH heading “Cytochrome P-450 CYP1A1” or the text words “CYP1A1” or “P4501A1” combined with the MeSH headings “Polymorphism (Genetics),” “Mutation,” “Point mutation,” “Genotype,” or the text words “polymorphism,” “mutation$$,” “gene,” “genes,” “genetic$$,” “genotyp$$,” or “allele$.” Additional articles were identified from the Centers for Disease Control and Prevention Genomics and Disease Prevention Information System and by hand searching reference lists in published papers. Eligible studies presented frequencies for each genotype separately in nondiseased persons. Studies that did not include controls for breast cancer patients were excluded if there were fewer than 200 subjects in each ethnic group, which would limit precision of the estimates of the genotype frequencies. If there appeared to be an overlap in subjects between studies, only the largest study was reported. Hardy-Weinberg equilibrium was assessed by using the Pearson χ² test.

Web tables 1, 2, 3, and 4 show homozygous variant and heterozygous genotype frequencies for the 3801T→C, Ile462Val, 3205T→C, and Thr461Asp polymorphisms (13, 17, 19–85). (This information is described in the first four of eight supplementary tables; each is referred to as “Web table” in the text and is posted on the website of the Human Genome Epidemiology Network (http://www.cdc.gov/genomics/hugenet/reviews.htm) as well as on the Journal’s website (http://aje.oupjournals.org/).) The subjects in most studies are volunteers (with the sampling frame unspecified) or hospital or clinic patients. It is unclear whether genotype frequencies in such series will reflect those in the general population. Considerable data are available from Japanese, western European, and White American populations. Data are limited, or not available, for other populations.

**TABLE 1. CYP1A1 3801T→C, Ile462Val, 3205T→C, and Thr461Asp polymorphisms**

<table>
<thead>
<tr>
<th>Allele nomenclature (10, 11)</th>
<th>Nucleotide change</th>
<th>Amino acid change</th>
<th>Location</th>
<th>Proposed allele nomenclature</th>
<th>Reference to first report: study, year (reference no.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP1A1*1A</td>
<td>None</td>
<td>None</td>
<td>3' non-coding region (downstream of polyadenylation site)</td>
<td>CYP1A1*1</td>
<td>Garte (12)</td>
</tr>
<tr>
<td>CYP1A1*2A</td>
<td>3801T→C</td>
<td>None</td>
<td>3' non-coding region</td>
<td>CYP1A1*2</td>
<td>CYP1A1*2A</td>
</tr>
<tr>
<td>CYP1A1*2B</td>
<td>3801T→C</td>
<td>None</td>
<td>3' non-coding region</td>
<td>CYP1A1*2B</td>
<td>Bale et al., 1987 (14); Spurr et al., 1987 (15)</td>
</tr>
<tr>
<td>CYP1A1*2C</td>
<td>2455A→G</td>
<td>Isoleucine→valine</td>
<td>Exon 7, codon 462</td>
<td>CYP1A1*3</td>
<td>Hayashi et al., 1991 (17)</td>
</tr>
<tr>
<td>CYP1A1*3</td>
<td>3205T→C</td>
<td>None</td>
<td>3' non-coding region (upstream of polyadenylation site)</td>
<td>CYP1A1*4</td>
<td>CYP1A1*3</td>
</tr>
<tr>
<td>CYP1A1*4</td>
<td>2453C→A</td>
<td>Threonine→asparagine</td>
<td>Exon 7, codon 461 (heme binding region)</td>
<td>CYP1A1*5</td>
<td>Crofts et al., 1993 (16)</td>
</tr>
<tr>
<td>CYP1A1*4</td>
<td>2453C→A</td>
<td>Threonine→asparagine</td>
<td>Exon 7, codon 461 (heme binding region)</td>
<td>CYP1A1*4</td>
<td>Cascorbi et al., 1996 (13)</td>
</tr>
</tbody>
</table>

Am J Epidemiol 2005;161:901–915
Information is also lacking on genotype frequencies in different age groups. Most studies consider only the 3801T→C and/or Ile462Val polymorphisms, which has the potential to result in misclassification. When individual polymorphisms are assessed, those persons who do not carry the specific variant may not be true wild-type homozygotes; a proportion may carry another variant. Moreover, the presence of the Thr461Asp polymorphism may interfere with detection of the Ile462Val polymorphism, resulting in overestimation of the Val allele if the polymerase chain reaction product has not been digested with BsrD1 (13). Genotype frequencies were in Hardy-Weinberg equilibrium, except in two studies of the 3801T→C polymorphism (63, 65) and nine studies of the Ile462Val polymorphism (19, 41, 53, 64, 67, 70, 76, 78, 82).

3801T→C (CYP1A1*2A, CYP1A1*2B). The 3801C variant is most prevalent in Asian populations, where the frequency of the C/C genotype is 2–18 percent and that of the T/C genotype is 32–55 percent. In European and White American series, 0–5 percent are C/C and 9–28 percent are T/C. Frequencies in African Americans are intermediate between White and Asian populations (4–6 percent C/C, 35–39 percent T/C).

In our pooled analysis, the C/C genotype frequency was 13 percent (95 percent confidence interval [CI]: 12.0, 14.0) in Asians, 1 percent (95 percent CI: 0.9, 1.4) in Whites, and 6 percent (95 percent CI: 3.7, 9.1) in African Americans. The heterozygote frequency was 44 percent (95 percent CI: 42.6, 45.6) in Asians, 17 percent (95 percent CI: 16.5, 18.0) in Whites, and 36 percent (95 percent CI: 31.7, 40.6) in African Americans.

Ile462Val (CYP1A1*2B, CYP1A1*2C). In all ethnic groups, the Val variant occurs less frequently than the 3801C variant. Similar to the 3801C variant, it is most common among Asians, where 1–8 percent are Val/Val and 15–46 percent are Ile/Val. In Europeans and US Whites, at most 3 percent are Val/Val and as many as 15 percent are Ile/Val. The Val variant is less common among African Americans than Whites. In studies including African Americans, no subjects had the Val/Val genotype, but up to 6 percent were Ile/Val.

Our pooled estimate of Val/Val genotype frequency was 5 percent (95 percent CI: 4.0, 5.0) in Asians, 0.7 percent (95 percent CI: 0.5, 0.8) in Whites, and 0 percent in African Americans. The pooled estimate of Ile/Val genotype frequency was 31 percent (95 percent CI: 29.5, 31.7), 8 percent (95 percent CI: 7.8, 8.9), and 5 percent (95 percent CI: 3.2, 7.3), respectively.

3205T→C (CYP1A1*3). The 3205C variant was originally thought to occur in African Americans only. This view is supported by studies of Turkish (34), French (84), German (13), Polish (50), Russian (51), and US White subjects (63, 65) in whom the 3205C variant was not found. In four African-American series, less than 1 percent had the C/C genotype, while 14–24 percent were heterozygotes. In our pooled analysis, the C/C and T/C genotype frequencies were 0.1 percent (95 percent CI: 0.0, 0.8) and 15 percent (95 percent CI: 12.8, 18.3), respectively, in African Americans.

Thr461Asp (CYP1A1*4). Asp/Asp homozygotes are very rare (≤1 percent). The Thr/Asp genotype frequency is 4–12 percent in Turkish, European, and White North American populations. The pooled estimates of the Asp/Asp and Thr/Asp genotype frequencies in Whites were 0.2 percent (95 percent CI: 0.1, 0.4) and 8 percent (95 percent CI: 7.1, 8.8), respectively.

Associations between the CYP1A1 polymorphisms. Studies of linkage between the polymorphisms are limited by the relative rarity of the variants. From the comparatively few studies carried out, the 3801T→C and Ile462Val polymorphisms appear to be closely linked in Asians (17, 21, 25, 34), less closely linked in Europeans (41, 52), and not linked in African Americans (60). In 81 Africans and African Americans carrying the 3205C variant, 23 percent also carried the 3801C variant, and no subjects carried the Val variant (18). The Thr461Asp and 3801T→C polymorphisms were not linked in Turkish (34), German (13), or Polish (50) populations. No evidence for linkage between the Thr461Asp and Ile462Val polymorphisms was found in White American (64) or German series (13).

Functional effects

Because the 3801T→C polymorphism is located in the noncoding region, it was originally thought that any apparent functional consequences of the variant were due to linkage with another polymorphism in, for example, the coding region or the aryl hydrocarbon receptor. However, polymorphisms in noncoding sequences may influence gene function by altering the level, location, or timing of gene expression or messenger RNA stability (86).

Studies of the 3801T→C polymorphism and basal and/or induced CYP1A1 messenger RNA expression in lymphocytes and placenta have been inconsistent (87–91). For the Ile462Val polymorphism, one study found that mean messenger RNA (induced/basal) levels increased with number of Val variants (92). In another study, heterozygotes for both 3801C and Val variants had twofold increased basal CYP1A1 expression compared with homozygotes for the 3801T and Ile alleles (63). In one study, the 3205T→C and Thr461Asp polymorphisms were not associated with steady-state CYP1A1 messenger RNA levels (87).

The Val variant caused a twofold increase in complementary DNA–expressed activity in transformed yeast cells (93), but the kinetic properties of the two variants do not differ (94). In purified Escherichia coli, there was no difference between the allelic variants in benzo[a]pyrene bioactivation (95).

Studies of genotype and CYP1A1-dependent enzymatic activity in lymphocytes are inconsistent. Studies either suggested high activity associated with the 3801C and Val variants (89, 92, 96–98) or produced null findings (99–102). No significant effect of the Thr461Asp polymorphism has been found (101, 102).

The 3801C variant has been associated with higher levels of DNA adducts in breast tissue in some studies (103, 104), but not others (105–107). Findings from one study of Ile462Val and Thr461Asp polymorphisms and breast tissue adducts were null (106). Results of studies of 3801T→C, Ile462Val, and Thr461Asp polymorphisms in other tissues have been inconsistent (91, 108–112), as have those in white
blood cells (111–122). Findings of studies of 3801T→C and/or Ile462Val and levels of DNA damage, as assessed by 8-hydroxydeoxyguanosine in breast tissue (123), urine (124, 125) or leukocytes (126), or DNA-protein cross-links (33) have been null.

The 3801T→C polymorphism does not appear to be associated with serum estrone or estradiol levels (127) or the ratio of baseline urinary estrogen metabolites (2-hydroxyestrone/16-hydroxyestrone) (63). However, after indole-3-carbinol was ingested, the 2-hydroxyestrone/16-hydroxyestrone ratio increased significantly for persons with the 3801ITT genotype; heterozygotes showed no significant increase (63).

In some (74, 128–132), but not all (75, 111, 115, 116, 133–135), studies, the 3801C and Val variants are associated with higher urinary levels of 1-hydroxypyrene, a biomarker for PAH exposure. Results of studies of Ile462Val and urinary levels of 2-naphthol, another PAH biomarker, are inconsistent (74, 135). Findings of studies of the CYP1A1 genotype and urinary levels of cotinine (136), malondialdehyde (124), and biomarkers for organic solvent exposure (137) have been null.

Mammographic breast density is positively related to breast cancer risk (138, 139). In one study, neither the 3801T→C nor the Ile462Val polymorphisms were associated with breast density (140).

**DISEASE**

In 2002, over 1 million new cases of breast cancer were diagnosed worldwide (141). In both developed and developing countries, it is the most common cancer in women (142). In developed countries, incidence increases rapidly with age to about age 50 years; thereafter, rates rise less rapidly with age (143). There is a 16-fold variation in incidence between the population with the highest rate (Montevideo, Uruguay; world age-standardized incidence 114.9 per 100,000 in 1993–1995) and that with the lowest (The Gambia, 7.0 per 100,000 in 1997–1998) (144). In many populations, there has been a consistent long-term rise in incidence, which cannot be entirely attributed to the introduction of mammographic screening (145).

The autosomal dominant susceptibility genes, BRCA1 and BRCA2, account for about 5 percent of breast cancers (146, 147). Familial aggregation, which confers increased risk for first- and second-degree relatives (148, 149), does not appear to be entirely due to BRCA1 and BRCA2 (150), suggesting that other aspects of genetic susceptibility are important.

The products of CYP1A1 are involved in estrogen and PAH metabolism. The most firmly established risk factors for breast cancer relate to cumulative exposure of the breast to endogenous hormones, particularly estrogen (143). Risk is increased for women with longer cumulative exposure, that is, for those experiencing early menarche, late menopause, late first full-term pregnancy, or no pregnancies (151).

Exogenous hormones have also been associated with increased risk of breast cancer. Risk is increased among current users of hormone replacement therapy (152, 153) and current users of oral contraceptives (154, 155). Other lifestyle risk factors, such as postmenopausal obesity, lack of physical activity (156), and alcohol intake (157), may influence risk via effects on estrogen levels.

PAHs may be involved in breast cancer etiology. These substances are lipophilic and are stored in adipose tissue, including that of the breast (158), and they are activated and metabolized by breast epithelial cells (159). Adduct levels are higher in normal breast tissue of breast cancer cases than in that of healthy controls (160), although it is unclear whether this is a cause or effect of disease. PAHs also affect estrogen production and metabolism, thereby acting as xenoestrogens; many xenoestrogenic compounds induce mammary carcinogenesis in experimental animals (161). PAHs themselves are powerful mammary carcinogens in mice (162).

Tobacco smoke is a major environmental source of PAH exposure (163). Most studies of breast cancer and smoking show a weak positive or null association (164–170), although the association may be stronger for premenopausal women or for those who started smoking at an early age (171, 172) or smoked before their first full-term pregnancy (173). Although positive associations with passive smoking have been reported (168, 174, 175), a recent review concluded that this factor was unlikely to increase risk (176).

PAHs (and heterocyclic amines) are formed when meats are exposed to temperatures that cause pyrolysis (177). An expert review of observational evidence and a recent meta-analysis suggested that high-meat diets increase breast cancer risk (178, 179), and, whereas most investigators have not considered cooking methods, some studies found raised risk with increased consumption of fried, broiled, and/or well-done meat (180–183).

**ASSOCIATIONS**

Web tables 5, 6, 7, and 8 summarize 17 studies of CYP1A1 and breast cancer risk (25, 32, 33, 38, 45, 56, 63, 65, 66, 73, 79, 81, 107, 127, 128, 184, 185) identified by using the search strategy described earlier, with the addition of the MeSH heading “Breast neoplasms” or the text word “breast.” The subjects included in the studies of Huang et al. (32, 186), Taioli et al. (63, 187), Li et al. (105) and Zhu et al. (107), and Ritchie et al. (188) and Bailey et al. (65) may overlap. Therefore, only the largest of each set was included in Web tables 5, 6, 7, and 8 and in our meta-analyses.

Meta-analyses of studies of 3801T→C, Ile462Val, and Thr461Asp were carried out. From the papers, we abstracted the odds ratios or relative risks for homozygous variants (3801TC/C, Val/Val, or Asp/Asp) and heterozygotes (3801TT/C, Ile/Val, or Thr/Asp) versus homozygous wild types (3801TT/T, Ile/Ile, or Thr/Thr). When reported, the adjusted effect estimate was included in the analysis in preference to the unadjusted one. If odds ratios were not reported, we computed unadjusted odds ratios from the data presented. Analyses were conducted by using Stata statistical software, release 7.0 (189). Heterogeneity was assessed by the Q test, with a fixed-effects model used if p ≥ 0.1 and a random-effects model used if p < 0.1. The I² statistic was also calculated as a measure of consistency between studies.
(190). Except for the association between breast cancer and the 3801T→C polymorphism, the estimates of effect in the first published study were similar to those for the cumulative meta-analyses.

**Study characteristics**

Four studies took place in Japan, two in Taiwan, six in the United States, and one each in Canada, Brazil, France, Greece, and the United Kingdom. Thirteen studies analyzed the 3801T→C polymorphism (2,484 cases), 10 analyzed the Ile462Val polymorphism (3,535 cases), two analyzed the 3205T→C polymorphism (280 cases), and three analyzed the Thr461Asp polymorphism (2,245 cases).

In one study, case DNA was derived from tumor specimens (38); in the remainder, and for all control series, DNA came from blood samples. Of the US studies, two involved subjects of whom the majority (or all) were White, three included more than one ethnic group (analyzed separately in two studies), and, in one, ethnicity was not reported. One study included postmenopausal women only; all others either consisted of both pre- and postmenopausal women (n = 6) or did not describe the subjects’ menopausal status (n = 10). Eleven studies included fewer than 200 breast cancer cases.

In 15 studies, cases were recruited from clinics or hospital series; in one study, cases were identified from a cancer registry; and one study was nested within the Nurses’ Health Study. Without information on all potentially eligible cases in the population, it is difficult to assess the generalizability of the results. At least four control series included “volunteers” from either an unspecified source or a convenient population such as medical workers—a potential source of bias. Seven studies presented estimates adjusted for potential confounding factors.

In general, the studies considered the polymorphisms separately. Therefore, the effect of one polymorphism may have been overshadowed by the effects of others, whereas construction of haplotypes may have revealed effects that were not apparent by analyzing single polymorphisms. Studies of the Ile462Val polymorphism, with the exception of those by Bailey et al. (65), Krajinovic et al. (56), and Basham et al. (79), may have suffered from some minor misclassification due to the undetected presence of the Thr461Asp polymorphism.

**3801T→C (CYP1A1*2A, CYP1A1*2B)**

Most studies found no evidence of an association between the 3801T→C polymorphism and breast cancer risk (33, 38, 45, 56, 65, 107, 127, 184, 185) (Web table 5). In Taiwan, women with the C/C genotype had a raised risk compared with other genotypes combined (32). African-American women with the C/C genotype also had an increased risk compared with those with the T/T genotype (63), but this study included only 25 cases. The 3801C variant was associated with reduced risk for Japanese and non-White Brazilian women (25, 66). However, in both studies, the cases were surgical series, and controls were not population based.

Our meta-analysis included eight studies for which data were available for all three genotypes separately (25, 32, 38, 45, 56, 63, 65, 66). Breast cancer risk did not differ from unity for C/C versus T/T (random-effects relative risk (RR) = 0.97, 95 percent CI: 0.52, 1.80; Q = 15.26, p = 0.08) or for T/T versus T/T (random-effects RR = 0.91, 95 percent CI: 0.70, 1.2, Q = 17.34, p = 0.07). The I² statistics for these analyses were 41 percent and 42 percent, respectively, indicating moderate heterogeneity across studies.

**Ile462Val (CYP1A1*2B, CYP1A1*2C)**

A Japanese study found a significantly reduced risk for women with the Ile/Val genotype compared with the Ile/Ile genotype (RR = 0.66, 95 percent CI: 0.44, 0.99) (25). However, meta-analysis found no association between breast cancer risk and the Val/Val (fixed-effects RR = 1.04, 95 percent CI: 0.63, 1.74; Q = 4.59, p = 0.33, I² = 13 percent) (25, 32, 73, 79, 81) or Ile/Val (fixed-effects RR = 0.92, 95 percent CI: 0.76, 1.10; Q = 11.57, p = 0.17, I² = 31 percent) (25, 32, 56, 63, 65, 73, 79, 81) genotypes versus the Ile/Ile genotype.

**3205T→C (CYP1A1*3)**

There was no association between the 3205C variant and breast cancer in the two available studies (Web table 7). However, these studies each included small series (n = 27 and n = 59) of African-American breast cancer cases (63, 65).

**Thr461Asp (CYP1A1*4)**

In a Canadian study, carriers of the Asp variant had an increased breast cancer risk (adjusted RR = 3.3, 95 percent CI: 1.1, 9.7) (56) (Web table 8). Results of the other studies, in White American women and African-American women, and in White women in England, were null. Meta-analysis found no association between disease risk and the Asp/Asp (fixed-effects RR = 0.95, 95 percent CI: 0.2, 4.49; Q = 0.52, p = 0.77, I² = 0 percent) or Thr/Asp (fixed-effects RR = 1.12, 95 percent CI: 0.87, 1.43; Q = 0.89, p = 0.64, I² = 0 percent) genotypes versus the Thr/Thr genotype (56, 65, 79).

**Combinations of genotypes**

Taioli et al. (187) assessed the impact of combinations of 3801T→C, Ile462Val, and 3205T→C genotypes on breast cancer risk. Compared with homozygotes for the 3801T, Ile, and 3205T alleles, only the 3801/C/C genotype was associated with increased risk for African-American women (RR = 5.8, 95 percent CI: 1.0, 36.0), but the effect estimate was imprecise.

One study combined Ile462Val and Thr461Asp genotypes and found no significant effect in any of the three combined genotype groups relative to the group with the Ile/Ile and Thr/Thr genotypes (79). Another study assessed disease risk for subjects with either the Val or Asp variant; no significant association was found in White women or African-American women (65).
Subgroup analyses

Menopausal status, age at menarche, and estrogen and progesterone receptor status. In a Taiwanese study, the association of the 3801C/C genotype with raised disease risk was evident in postmenopausal, but not premenopausal, women, and further analysis suggested that the relation might be more pronounced for women experiencing early menarche (32). Other studies found no association between the 3801C variant and breast cancer when subjects were stratified by menopausal status (66, 185) or age at menarche (66). There was no evidence for an association with the Ile462Val polymorphism for either pre- or postmenopausal women (32, 65, 79, 185). In a Canadian study, the increased risk associated with the Thr461Asp polymorphism was evident for postmenopausal women only (56); however, a study in the United Kingdom found no difference in Thr461Asp genotypic risks by menopausal status (79). There were no significant associations between CYP1A1 polymorphisms and estrogen or progesterone receptor status (25, 65, 127, 185).

Age at diagnosis and clinical characteristics. Studies investigating CYP1A1 genotype and age at diagnosis of breast cancer have produced inconsistent results (25, 32, 38, 65, 81). The 3801C variant has been significantly associated with a higher frequency of lymph-node metastasis and the Val variant with a higher frequency of small tumors (<2 cm), while neither variant was associated with histology or histologic grade (25). Other studies found no association between the four polymorphisms and tumor size, stage, type, grade, or nodal status (65) or between 3801T→C or Ile462Val and tumor type or stage of disease (185).

Survival

In a British study of 1,793 incident or prevalent breast cancer cases, the Ile462Val polymorphism was not related to survival (191). The hazard ratio was reduced for Thr/Asp heterozygotes compared with Thr/Thr homozygotes, but not significantly (hazard ratio = 0.67, 95 percent CI: 0.33, 1.37) (191).

Other diseases

CYP1A1 has been explored in relation to several cancers, particularly those in which smoking is implicated. In pooled and meta-analyses, the 3801C and Val variants were associated with increased lung cancer risk in Whites, but not Asians (192–195). Neither the 3205C (196) nor the Asp (13, 197) variants were associated with lung cancer risk. Also investigated, with mainly either inconsistent or unconfirmed results, have been tumors of the head and neck (29, 31, 40, 43, 48, 82, 198–215), large bowel (24, 53, 58, 216–221), prostate (72, 222–224), female gynecologic sites (62, 76, 225–230), skin (231, 232), and kidney (84) and liver (28), as well as leukemias and lymphomas (36, 57, 78, 233–238). Results of studies of bladder (44), brain (239), and pancreatic (64, 108, 240–242) cancer have been null.

Associations have been found between CYP1A1 and other diseases, including male infertility (243), systemic lupus erythematosus (244), type II porphyria cutanea tarda (245), psoriasis (39), ankylosing spondylitis (246), and rheumatoid arthritis (247). Findings from studies of endometriosis (46, 248, 249) and Parkinson’s disease (54, 68, 250, 251) have been inconsistent, while those for asthma (252), atherosclerosis (253), cirrhosis (37), Crohn’s disease (254), age-related macular degeneration (255), leukoplakia (20), early pregnancy loss (256), acne (41), and oral clefting (257) have been null.

INTERACTIONS

If CYP1A1 is involved in breast cancer, it may influence disease risk by interacting with exposure (or indicators of exposure) to PAHs or estrogen, for example, or with other genes involved in the metabolism of carcinogens, estrogens, or other hormones. Sample size is particularly important in this context. For instance, to detect a multiplicative interaction, very large sample sizes are required for adequate power (258). Although the sample size needed to detect other types of interactions may be smaller (259), a priori it is not usually clear what model of interaction would be predicted.

Gene-environment interactions

Smoking. In five studies investigating genotype-smoking interactions (63, 65, 79, 81, 185), two found evidence of an interaction (81, 185). In Ambrosone et al.’s study (81), adjusted relative risks for Val carriers versus Ile/Ile homozygotes among nonsmokers, light smokers (<29 pack-years of exposure), and heavy smokers (≥29 pack-years) were 1.3 (95 percent CI: 0.62, 2.70), 5.2 (95 percent CI: 1.16, 23.56), and 0.9 (95 percent CI: 0.24, 3.09), respectively, but no formal test of interaction was conducted. Ishibe et al. (185) found no interaction between pack-years of smoking and either 3801T→C or Ile462Val polymorphisms, but they observed effect modification for smoking status at diagnosis and age at which smoking started. Risk was significantly raised for current smokers carrying the 3801C variant versus 3801T/T nonsmokers (p for interaction = 0.06) and for women with either variant who started smoking before age 18 years versus 3801T/T nonsmokers (p for interaction = 0.04) and Ile/Ile nonsmokers (p for interaction = 0.08).

The numbers analyzed in the studies of genotype-smoking interactions were small, and interpretation is difficult because of differences in the way in which interactions were assessed (stratifying by smoking status (63, 81) or genotype (65), or using a single reference group of smoking status and genotype combined (185)) and in categorization of smoking status. For example, the interaction patterns observed by Ambrosone et al. (81) or Ishibe et al. (185) would not be detectable by using an ever/never smoking categorization, as has been used in other studies (63, 65).

Polychlorinated biphenyls (PCBs). PCBs have been linked to breast cancer risk because of their estrogenic (260) and tumor-promoting (261) properties. In the Nurses’ Health Study, a modest interaction between the Ile462Val polymorphism and plasma PCBs was found for postmenopausal.
but not premenopausal breast cancer (262). Among postmenopausal subjects, the adjusted relative risk for Val carriers in the upper tertile of plasma PCB levels, compared with Ile/Ile homozygotes in the lowest PCB tertile, was 2.78 (95 percent CI: 0.99, 7.82, p for interaction = 0.05). There was no interaction between PCBs and 3801T → C (262). In a subset of a study in western New York (81), Val carriers with an above-median PCB body burden had an increased risk compared with Ile/Ile homozygotes with a below-median PCB burden (adjusted RR = 2.9, 95 percent CI: 1.18, 7.45; p for interaction = 0.13) (263).

Gene-gene interactions

CYP1A1 and glutathione S-transferase (GST) gene activities may be interrelated. The GST genes belong to the Ah gene battery, since GST is one of six enzymes regulated by the aryl hydrocarbon receptor (1). In human B-cell lines, absence of GSTM1 was associated with induction of high levels of CYP1A1 messenger RNA by 2,3,7,8-tetrachlorodibenzop-dioxin, and presence of GSTM1 was associated with induction of low levels (264). Four breast cancer studies found no evidence of a CYP1A1-GSTM1 interaction (56, 65, 66, 81), although CYP1A1-GSTM1 genotype combinations have been associated with age at presentation (38). In two relatively small studies assessing CYP1A1-GSTTI genotype combinations, risk estimates were not significant, and tests for interaction were not reported (65, 66). Another study (45) found no significant differences in combined genotype frequencies between breast cancer patients and controls.

CYP17, CYP19, and catechol-O-methyltransferase (COMT) are involved in steroid hormone metabolism (265, 266). One study found an increased breast cancer risk associated with the presence of two “high-risk” genotypes, defined as homozygosity for the CYP1A1 3801C, CYP17 A2, or COMT low-activity alleles (RR = 3.5, 95 percent CI: 1.06, 12.04), but no test for interaction was conducted (186). In another study, 3801T/T homozygous carriers of the CYP19 (TTTA)7–(3bp) allele had increased risk of estrogen-receptor-positive breast cancer (adjusted RR = 3.00, 95 percent CI: 1.56, 5.74) compared with women carrying the 3801C variant but not the CYP19 (TTTA)7–(3bp) allele (127).

CONCLUSIONS AND RECOMMENDATIONS FOR RESEARCH

The CYP1A1 3801C variant is more common than the Val variant. Both variants occur more frequently in Asian than in White populations. The 3205C variant has been observed in African Americans only, and little data are available on the geographic or ethnic distribution of the Thr461Asp polymorphism. The 3205T → C and Thr461Asp polymorphisms should be investigated in African, Asian, and Hispanic populations. The functional significance of all four polymorphisms is unclear, which could be due to the small sample sizes of most studies. Further investigation is warranted.

No consistent associations between breast cancer and CYP1A1 polymorphisms were found. While meta-analyses have greater power and precision for detecting gene-disease associations, our meta-analyses were limited by different genotype categorizations between studies. For the purposes of future meta-analyses, authors should provide results for all genotypes separately.

The 3801T → C and Ile462Val polymorphisms may modify the smoking-disease association, although the evidence is limited and inconsistent. A similar “inverse dose effect” has been observed in studies of CYP1A1, smoking, and lung cancer (60, 184, 268), and it has been suggested that the genetic variant might confer increased sensitivity to lower levels of exposure (269). Additional investigation is needed. The Val variant may interact with PCB levels to affect breast cancer risk, but confirmation is necessary. There was no evidence that GSTM1 or GSTTI and CYP1A1 polymorphisms have a joint effect on disease risk. Studies suggesting interactions with CYP17, CYP19, and COMT should be replicated.

INTERPRETING THE STUDIES OF INTERACTION WAS DIFFICULT BECAUSE OF THE DIFFERENT APPROACHES USED; ADOPTING A MORE UNIFIED APPROACH (e.g., Botto and Khoury (270)) in future studies would aid interpretation and synthesis of evidence.
Studies are needed to explore joint effects on breast cancer risk of the CYP1A1 genotype and 1) sources of PAH exposure other than tobacco, 2) markers of exposure to endogenous estrogens, 3) exposure to exogenous estrogens, 4) other lifestyle factors that influence hormone levels, 5) other genes encoding enzymes involved in PAH metabolism, and 6) other genes involved in hormonal biosynthesis. To detect gene-environment or gene-gene interactions, future studies must be large, and pooled analyses should be considered.

REFERENCES


Am J Epidemiol 2005;161:901–915
77. Roddams PL, Rollinson S, Kane E, et al. Poor metabolizers at the cytochrome P450 2D6 and 2C19 loci are at increased risk of developing adult acute leukaemia. Pharmacogenetics 2000;10:605–15.
102. Smart J, Daly AK. Variation in induced CYP1A1 levels: relationship to CYP1A1, Ah receptor and GSTM1 polymorphisms. Pharmacogenetics 2000;10:11–24.


112. Rojas M, Alexandrov K, Cascorbi I, et al. High benzo(a)pyrene diol-epoxide DNA adduct levels in lung and blood cells from subjects with combined CYP1A1 MspI/MspI and GSTM1*0/*0 genotypes. Pharmacogenetics 1998;8:19–18.


Am J Epidemiol 2005;161:901–915

CYP1A1 and Breast Cancer 913

Downloaded from https://academic.oup.com/aje/article-abstract/161/10/901/104834 by guest on 03 February 2019


Downloaded from https://academic.oup.com/aje/article-abstract/161/10/901/104834 by guest on 03 February 2019

APPENDIX

Internet Sites

Data on cancer incidence, survival, and mortality

International Agency for Research on Cancer—Cancer Mondial: http://www-dep.iarc.fr/
Surveillance, Epidemiology, and End Results Program: http://www.seer.cancer.gov/publicdata/
National Programme of Cancer Registries (NPCR): http://www.cdc.gov/cancer/npcr

Information on cancer

Cancer Research UK: http://www.cancerresearchuk.org
American Association of Cancer Research: http://www.aacr.org/
American Cancer Society: http://www.cancer.org/docroot/home/index.asp

Genetic information

Human Genome Epidemiology Network (HuGENet): http://www.cdc.gov/genomics/hugenet/default.htm
Public Health Genetics Unit: http://www.phgu.org.uk/index.php
Human Gene Mutation Database: http://archive.uwcm.ac.uk/ucwm/mg/hgmd0.html
GeneCards: http://www.cgal.icnet.uk/gene_cards/
Human cytochrome P-450 allele nomenclature: http://www.imm.ki.se/CYPalleles/
MRC Rosalind Franklin Centre for Genomics Research (includes links to other sites via The Genome Web): http://www.hgmp.mrc.ac.uk/
Links to other sites: http://cedar.genetics.soton.ac.uk/public_html/links.html