Pairwise Combinations of Estrogen Metabolism Genotypes in Postmenopausal Breast Cancer Etiology

Timothy R. Rebbeck,1,2 Andrea B. Troxel,1,2 Amy H. Walker,1 Saarene Panossian,1 Stephen Gallagher,1 Ekaterina G. Shatalova,5 Rebecca Blanchard,5 Sandra Norman,1,2 Greta Bunin,1,2,4 Angela DeMichele,1,2,3 Michelle Berlin,6 Rita Schinnar,1 Jesse A. Berlin,1,2 and Brian L. Strom1,2,3

1Center for Clinical Epidemiology and Biostatistics and Department of Biostatistics and Epidemiology; 2Abramson Cancer Center; 3Department of Medicine; 4Division of Oncology, Children’s Hospital of Philadelphia; 5Fox Chase Cancer Center, Philadelphia, Pennsylvania; and 6Department of Obstetrics and Gynecology Oregon Health and Science University, Portland, Oregon

Abstract

Estrogen exposures have been associated with breast cancer risk, and genes involved in estrogen metabolism have been reported to mediate that risk. Our goal was to better understand whether combinations of candidate estrogen metabolism genotypes are associated with breast cancer etiology. A population-based case-control study in three counties of the Philadelphia Metropolitan area was undertaken. We evaluated seven main effects and 21 first-order interactions in African Americans and European Americans for genotypes at COMT, CYP1A1, CYP1A2, CYP1B1, CYP3A4, SULT1A1, and SULT1E1 in 878 breast cancer cases and 1,409 matched random digit–dialed controls. In European Americans, we observed main effect associations of genotypes containing any CYP1A1*2C (odds ratio, 1.71; 95% confidence interval, 1.09-2.67) and breast cancer. No significant main effects were observed in African Americans. Three significant first-order interactions were observed. In European Americans, interactions between SULT1A1*2 and CYP1A1*2C genotypes (P_interaction < 0.001) and between SULT1E1 and CYP1A2*1F genotypes were observed (P_interaction = 0.006). In African Americans, an interaction between SULT1A1*2 and CYP1B1*4 was observed (P_interaction = 0.041). We applied the false-positive report probability approach, which suggested that these associations were noteworthy; however, we cannot rule out the possibility that chance led to these associations. Pending future confirmation of these results, our data suggest that breast cancer etiology in both European American and African American postmenopausal women may involve the interaction of a gene responsible for the generation of catecholestrogens with a gene involved in estrogen and catecholestrogen sulfation. (Cancer Epidemiol Biomarkers Prev 2007;16(3):444–50)

Introduction

There is substantial evidence that estrogen exposure is associated with breast cancer risk. This exposure can be in the form of endogenous exposures related to obesity, lifestyle, or reproductive history (1-3) or exogenous exposures such as use of hormone replacement therapy. It has been well established that hormone replacement therapy containing unopposed estrogen increases breast cancer risk and that combined estrogen-progestin hormone replacement therapy is associated with even higher risks (4).

It can be hypothesized that the metabolism of estrogens, mediated by inherited genotypes, influences the estrogen-breast cancer relationship. The genes involved in the disposition of estrogen are well known, and include catechol-O-methyltransferase (COMT), progesterone receptor (PGR), the sulfotransferases SULT1A1 and SULT1E1, and members of the cytochrome P50 family including CYP1B1, CYP1A2, CYP1A1, and CYP3A4 (Fig. 1; Table 1; ref. 5). There is also support from many studies for the hypothesis that these genes influence breast cancer risk (6-13). Therefore, the combined knowledge of the function of these genes in a relevant biological pathway, evidence for the functional significance of most of the variants in these genes, and prior evidence from the literature regarding associations with breast cancer provide a relatively high prior probability that these genes may be involved in breast cancer etiology. However, studies to date have focused on the main effects of these genes and have not comprehensively evaluated whether these genes interact with one another to influence breast cancer etiology. Therefore, we evaluated whether first-order interaction of genes involved in the downstream metabolism of estrogens were associated with breast cancer risk in a population-based sample of postmenopausal women.

Materials and Methods

Study Design and Data Collection. The Women’s Insights and Shared Experiences study is a population-based case-control study in which incident breast cancer cases and endometrial cancer cases were identified through hospitals and the Pennsylvania State Cancer Registry, and frequency-matched controls were identified from the community using random-digit dialing. Thus, both a breast cancer case-control study and an endometrial cancer case-control study were conducted using similar methods. Additional details of our study design can be found in studies by Bunin et al. (14) and Strom et al. (15) The source population for this study was the three counties of Philadelphia (PA), Delaware (PA), and Camden (NJ). Potentially eligible cases were women residing in these counties at the time of diagnosis who were ages 50 to 79 years old and were newly diagnosed with breast cancer between July 1, 1999, and June 30, 2002. The cases were
identified through active surveillance at 61 of 62 hospitals in these counties (one small hospital was unable to meet the regulatory requirements to obtain human subjects approval). The median duration from reference date to ascertainment was 52 days, and from ascertainment to interview was 226 days.

Pathology reports and other medical records were reviewed to validate cancer diagnoses and obtain information about tumor type, size, grade, stage, and hormone receptor activity. Breast cancer eligibility was confirmed if a pathology report was compatible with a first primary, invasive, breast cancer of any stage (I, II, III), any grade, and any tissue type (ductal, lobular, mucinous, papillary, mixed). Women with ductal carcinoma in situ, lobular carcinoma in situ, and other nonmalignant tumor types were excluded.

Controls were selected by random-digit dialing from the same geographic regions as the cases, and frequency-matched to the cases on age (in 5-year age groups) and calendar date of diagnosis to the cases on age (in 5-year age groups) and race. Additional eligibility criteria for both cases and controls included living in a noninstitutional setting; age 20-84 years; having no severe cognitive, language, or speech impairment; and ability to speak English; and having a household telephone; ability to speak English; and not speak English, 5 were deceased, 199 could not be recontacted because they moved or changed their phone number, and 439 refused. Of the 1,963 controls who were eligible and accessible, 1,524 (78%) completed the interview (64% of those referred). The analysis was restricted to the 627 European American and 251 African American breast cancer cases who were postmenopausal at the time of their diagnosis, and 838 European American and 571 African American age- and race-matched postmenopausal controls.

The study was approved by the University of Pennsylvania Committee on Studies Involving Human Beings and by the institutional review boards of all the participating hospitals. Subjects gave verbal informed consent for the interview and written informed consent for the buccal samples.

**Laboratory Methods.** Buccal swabs were obtained by mail from each participant. Extraction of genomic DNA was done using the QIAamp 96 DNA Buccal Swab Biorobot kit and done on a 9604 Biorobot (Qiagen, Inc., Valencia, CA). No whole-genome amplification was done before genotype analysis. To limit the number of hypotheses tested (and thereby the potential for false-positive associations), we identified the seven genes involved in the downstream metabolism of estrogen. We chose one single nucleotide polymorphism (SNP) in each gene that represented a functionally relevant SNP with a sufficiently high allele frequency to provide

---

### Table 1. Description of genes and variants studied

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP nucleotide designation</th>
<th>Rs no.</th>
<th>SNP aliases</th>
<th>Primary variant function</th>
<th>Hypothesized estrogenicity of variant</th>
<th>Hypothesized mutagenicity of variant</th>
<th>Hypothesized breast cancer risk of variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMT</td>
<td>1947G&gt;A</td>
<td>rs4680</td>
<td>Val^{150}Met, Val^{108}Met, Val^{60}Met</td>
<td>Methionine allostere has decreased activity (42)</td>
<td>—</td>
<td>Increased</td>
<td>Increased</td>
</tr>
<tr>
<td>CYP1A1</td>
<td>6750A&gt;G</td>
<td>rs1048943</td>
<td>*2C, *2m2</td>
<td>Valine allostere has increased inducibility to 2-, 4-, and 16a-catecholestrogens and the main source of extrahepatic 16α-catecholestrogen (43)</td>
<td>Weak</td>
<td>Increased</td>
<td>Protective</td>
</tr>
<tr>
<td>CYPIA2</td>
<td>734C&gt;A, 163C&gt;A</td>
<td>rs762551</td>
<td>*1F</td>
<td>*1F has increased inducibility and ultrarapid activity to produce 2-OH- and 4-OH-catecholestrogens (44)</td>
<td>Weak</td>
<td>Increased</td>
<td>Protective</td>
</tr>
<tr>
<td>CYPIB1</td>
<td>1358A&gt;G or 3290A&gt;G</td>
<td>rs1800440</td>
<td>Asn^{62}Ser, *4</td>
<td>Serine allostere associated with higher catalytic efficiency toward 4-OH-estradiol (45)</td>
<td>Potent</td>
<td>Genotoxic, Free radical generator</td>
<td>Increased</td>
</tr>
<tr>
<td>CYP3A4</td>
<td>729A&gt;G</td>
<td>rs2740574</td>
<td>*1B</td>
<td>*1B allele may confer increased CYPIA4 expression (46)</td>
<td>Potent</td>
<td>Genotoxic, Free radical generator</td>
<td>Increased</td>
</tr>
<tr>
<td>SULT1A1</td>
<td>638G&gt;A</td>
<td>rs9282861</td>
<td>Arg^{213}His, *2</td>
<td>2 allostere has lower thermostability, lower enzyme activity, and lower estrogen sulfation ability (36, 37, 47)</td>
<td>Increased</td>
<td>Increased</td>
<td>Increased</td>
</tr>
<tr>
<td>SULT1E1</td>
<td>–64G&gt;A</td>
<td>rs3736599</td>
<td>Promoter variant</td>
<td>Not determined</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

We ascertained 1,890 incident cases with breast cancer who met the age, county, diagnostic, diagnosis date, and race criteria. Of these, 8 were living in a nursing home, 44 did not speak English, 25 were not mentally or physically able to participate, 416 did not have physician consent, 125 were without correct address and/or phone number, and 58 died before we could contact to interview them. Another 234 refused, and 31 could not be interviewed before the study ended. Of the 1,214 cases who were eligible and accessible, 949 (78%) were interviewed (50% of ascertained). The survey research firm provided the names, addresses, and telephone numbers for 2,381 potential random-digit dialing controls. Of these, 181 were ineligible because of age, gender, county, race, or history of breast cancer. Of those remaining, 22 could not participate because of physical or mental impairments, 11 did not speak English, 5 were deceased, 199 could not be recontacted because they moved or changed their phone number, and 439 refused. Of the 1,963 controls who were eligible and accessible, 1,524 (78%) completed the interview (64% of those referred). The analysis was restricted to the 627 European American and 251 African American breast cancer cases who were postmenopausal at the time of their diagnosis, and 838 European American and 571 African American age- and race-matched postmenopausal controls.

Pathology reports and other medical records were reviewed to validate cancer diagnoses and obtain information about tumor type, size, grade, stage, and hormone receptor activity. Breast cancer eligibility was confirmed if a pathology report was compatible with a first primary, invasive, breast cancer of any stage (I, II, III), any grade, and any tissue type (ductal, lobular, mucinous, papillary, mixed). Women with ductal carcinoma in situ, lobular carcinoma in situ, and other nonmalignant tumor types were excluded.

Controls were selected by random-digit dialing from the same geographic regions as the cases, and frequency-matched to the cases on age (in 5-year age groups) and race. Additional eligibility criteria for both cases and controls included living in a noninstitutional setting; having a household telephone; ability to speak English; and having no severe cognitive, language, or speech impairment. The median duration from reference date to interview was 112 days.
adequate power for testing first-order interactions (Fig. 1; Table 1). The variants selected were as follows: COMT Val158Met (rs4680), CYP1A1 Ile62Val (rs2740574), CYP1A2*2 (rs762551), CYP1B1 (Asn452Ser, *4; rs1800440), CYP3A4*1B (rs2740574), SULT1A1 Arg122His (rs29928261), and the SULT1E1 −64G→A promoter variant (rs3736599). COMT, CYP1A1, CYP1A2, CYP1B1, CYP3A4, and SULT1E1 were assayed using previously described methods (5, 16). Additional assay validations were undertaken using direct sequencing (CYP1A1*2C), pyrosequencing (SULT1E1 −64G→A), and RFLP (COMT, CYP1B1*4, CYP3A4*1B). SULT1A1 genotypes were determined as previously reported (5, 16). The control frequency distributions for these variants by race has also been previously reported (5).

Statistical Methods. Odds ratio (OR) estimates and 95% confidence intervals (95% CI) were calculated to evaluate the relationship of hormone metabolism genes and breast cancer risk. Genotype coding was based on knowledge of the functional effects of the variants as well as the frequency of genotypes of interest (Table 1). Therefore, the genotypes considered here reflect functionally relevant classes (when known) that provided optimal statistical power.

Multiple conditional logistic regression was used to simultaneously account for the matching variables (defined by combinations of age group, race, and date of interview) and known risk factors for breast cancer. Because of prior reports that some hormone metabolism genes are associated with both breast cancer and with breast cancer risk factors (17, 18), we also considered the potential for confounding of genotype with hormone-related breast cancer risk factors in our analyses. All models considered the same set of potential confounders: (a) body mass index as a continuous variable during the age decade of the 1940s; (b) number of full-term pregnancies (never pregnant, 1, 2, and 3+); (c) age at menarche as a continuous variable; (d) age at menopause; (e) family history of breast cancer (any in first-degree relatives versus none in first-degree relatives); and (f) oral contraceptive use (never, <3 years, ≥3 years). A variable was included in the linear model if it was a statistically significant predictor of breast cancer or if changed the point estimate of any genotype effect by 10% or more. If a variable ever met this criterion, it was included as a covariate in all models. Other variables that were not statistically significant in any model or did not change the point estimate associated with breast cancer risk by more than 10% were not included as confounder variables in this analysis. These included use of hormone replacement therapy and history of bilateral oophorectomy. Two sets of primary hypotheses were tested: First, we evaluated whether there was an association of breast cancer with each genotype adjusted for covariates described above. Second, we tested the first-order interaction of each variant with each other variant. Significance was assessed using likelihood ratio tests.

We did all analyses stratified by race to minimize the potential for confounding due to race. Because we have done multiple hypothesis tests, we considered the potential for false-positive associations. Accordingly, we applied the false-positive report probability (FPPR) approach (19), an approach that allows the investigator to interpret the results of hypothesis testing to insure against making false-positive inferences. All analyses were done in STATA (version 9.0, STATA Corporation, College Station, TX).

Results

The main effects of each genotype are presented in Table 2. No significant main effects were observed for any gene in African Americans. In European Americans, we observed significant main effects of genotypes containing any CYP1A2*2C (OR, 1.71; 95% CI, 1.09-2.67).

We identified three statistically significant interactions among the candidate genes studied here (Table 3). All of these interactions involved a gene that metabolizes estrogens to catecholestrogens and a sulfotransferase gene, which are involved in the sulfation and ultimate excretion of estrogens and catecholestrogens. First, we identified a significant interaction of CYP1A1*2C and SULT1A1*2 (Pinteraction = 0.0004) in European Americans. To further describe the risk conferred by these alleles, the adjusted OR for the effect of SULT1A1*2 and breast cancer among European American women with no CYP1A1*2C allele was 1.34 (95% CI, 1.01-1.76), whereas the adjusted OR for SULT1A1*2 and breast cancer among European American women with any CYP1A1*2C allele was 0.14 (95% CI, 0.04-0.56). Alternatively, the adjusted OR for the effect of CYP1A2*1F and breast cancer among European American women with no SULT1A1*2 allele was 4.33 (95% CI, 1.95-9.61), whereas the adjusted OR for the effect of CYP1A2*1F and breast cancer among European American women with any SULT1A1*2 allele was 0.57 (95% CI, 0.28-1.17). Second, we identified a significant interaction of CYP1A2*1F and SULT1E1 (Pinteraction = 0.0064) in European Americans. To further describe the risk conferred by these alleles, the adjusted OR for the effect of SULT1E1 and breast cancer among European American women with no CYP1A2*1F allele was 0.15 (95% CI, 0.03-0.64), whereas the adjusted OR for SULT1E1 and breast cancer among European American women with any CYP1A2*1F allele was 1.55 (95% CI, 1.09-2.19).

Table 2. Adjusted ORs and 95% CIs of the association of breast cancer case-control status and genotype main effects

<table>
<thead>
<tr>
<th>Genotype</th>
<th>European Americans (%)</th>
<th>African Americans (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls</td>
<td>Cases</td>
</tr>
<tr>
<td>No COMT Met</td>
<td>512 (78%)</td>
<td>377 (77%)</td>
</tr>
<tr>
<td>Any COMT Met</td>
<td>147 (22%)</td>
<td>115 (23%)</td>
</tr>
<tr>
<td>No CYP1A1*2C</td>
<td>682 (94%)</td>
<td>487 (91%)</td>
</tr>
<tr>
<td>Any CYP1A1*2C</td>
<td>40 (6%)</td>
<td>50 (9%)</td>
</tr>
<tr>
<td>No CYP1A2*1F</td>
<td>57 (9%)</td>
<td>45 (8%)</td>
</tr>
<tr>
<td>Any CYP1A2*1F</td>
<td>611 (91%)</td>
<td>451 (91%)</td>
</tr>
<tr>
<td>No CYP1B1*4</td>
<td>494 (73%)</td>
<td>343 (69%)</td>
</tr>
<tr>
<td>Any CYP1B1*4</td>
<td>186 (27%)</td>
<td>155 (31%)</td>
</tr>
<tr>
<td>No CYP3A4*1B</td>
<td>650 (92%)</td>
<td>464 (90%)</td>
</tr>
<tr>
<td>Any CYP3A4*1B</td>
<td>57 (9%)</td>
<td>86 (16%)</td>
</tr>
<tr>
<td>No SULT1A2</td>
<td>297 (53%)</td>
<td>199 (47%)</td>
</tr>
<tr>
<td>Any SULT1A2</td>
<td>259 (47%)</td>
<td>226 (53%)</td>
</tr>
<tr>
<td>No SULT1E1 A</td>
<td>613 (86%)</td>
<td>432 (83%)</td>
</tr>
<tr>
<td>Any SULT1E1 A</td>
<td>97 (14%)</td>
<td>89 (17%)</td>
</tr>
</tbody>
</table>

*Estimated from conditional logistic regression matched on age, and adjusted for the following variables: age at menarche, number of full-term pregnancies, age at menopause, duration of oral contraceptive use, body mass index, and existence of first-degree relative with breast cancer.

Cancer Epidemiol Biomarkers Prev 2007;16(3). March 2007
Downloaded from cebp.aacjournals.org on December 15, 2021. © 2007 American Association for Cancer Research.
Table 3. Adjusted ORs, 95% CIs, and P values for pairwise genotype interactions and breast cancer case-control status

<table>
<thead>
<tr>
<th>Gene</th>
<th>COMT</th>
<th>CYP1A1*2C</th>
<th>CYP1A2*IF</th>
<th>CYP1B1*3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SULT1E1</td>
<td>1.20 (0.48-2.96) [0.695]</td>
<td>1.81 (0.41-8.08) [0.440]</td>
<td>6.95 (1.72-28.00) [0.0064]</td>
<td>0.87 (0.40-1.87) [0.715]</td>
</tr>
</tbody>
</table>
| CYP1B1*4      | assumed that there is moderate evidence for an interaction = 0.041) in African Americans. The adjusted OR for the effect of SULT1A1*2 and breast cancer among African American women with any SULT1E1 allele was 0.79 (95% CI, 0.49-1.27), whereas the adjusted OR for the effect of CYP1A2*IF and breast cancer among European American women with any SULT1E1 allele was 6.77 (95% CI, 1.57-29.20). Third, we identified a significant interaction of CYP1B1*4 and SULT1A1*2 (interaction = 0.041) in African Americans. The adjusted OR for the effect of SULT1A1*2 and breast cancer among African American women with any SULT1E1 allele was 0.74 (95% CI, 0.47-1.17), whereas the adjusted ORs for SULT1A1*2 and breast cancer among African American women with any SULT1A1*2 allele was 3.25 (95% CI, 0.42-25.14). Similarly, the adjusted OR for the effect of CYP1B1*4 and breast cancer among African American women with no SULT1A1*2 allele was 0.75 (95% CI, 0.27-2.06), whereas the adjusted OR for the effect of CYP1B1*4 and breast cancer among African American women with any SULT1A1*2 allele was 3.60 (95% CI, 1.02-12.70).

To evaluate the possibility that these statistically significant associations represent false-positive findings, we calculated the FPRP. (19) Based on our knowledge of prior associations of these genes with breast cancer, the functional role of these genes in estrogen metabolism, and the putative functional significance of the variants in these genes, we assumed prior probabilities of observing the interaction effect of 0.3%, 1%, 5%, 10%, and 20% to represent a range of values that reflect low to high probabilities of a true association. For the main effect of CYP1A1*2C, the FPRP was <0.1 for a 10% prior probability. This suggests that this main effect is not likely to represent a false-positive report. For the first interaction in European Americans between CYP1A1*2C and SULT1A1*2, the FPRP was 7% for a prior probability of 0.005, and 10% for a prior probability of 0.01, making this result extremely unlikely to be a false-positive report. For both the second and the third interactions, FPRP was high (>50%) for prior probabilities of 0.005. Therefore, if the a priori chance of these associations is very low, these results may represent false-positive associations. However, given the strong biological plausibility of the genes and functional significance of the variants involved, it is reasonable to consider somewhat higher prior probabilities of association. For the second interaction in European Americans between CYP1A1*2F and SULT1E1, the FPRP was <0.1 for a 0.1 prior probability and 10% for a 0.2 prior probability. Again, this result is unlikely to be a false-positive report if it can be assumed that there is moderate a priori evidence for an association. For the interaction in African Americans between CYP1B1*4 and SULT1A1*2, the FPRP was <0.2 for a 0.2 prior probability. Although these values still represent FPRP values that suggest this finding is unlikely to be a false-positive report, this is the least likely of the three significant associations to be a true positive report.

Discussion

We report that functionally active variants in estrogen metabolism genes are associated with breast cancer risk both as main effects and in combination with one another. These results involve putative functionally significant genetic variants in which the direction of an hypothesized effect could be predicted a priori, and are therefore biologically plausible associations. We have identified three significant interactions, each of which involve one phase I catecholestrogen-generating gene and one phase II sulfation gene. The biological function of the catecholestrogen formation genes and the sulfation genes overlap, such that these genes are multifunctional and redundant in their activity (Fig. 1). These effects may depend on other factors not accounted for in the present analyses, including other exposures or lifestyle, which may also differ by race. Therefore, it is possible that interactions of genes involved in catecholestrogen formation with genes involved in estrogen or catecholestrogen sulfation are critical in breast cancer etiology; however, we may not have yet identified all of these genes involved in this interaction.

A statistically significant main effect of CYP1A1*2C on breast cancer risk was observed in European Americans. CYP1A1 is expressed at low levels in the liver, but is expressed at higher levels in breast tissues (20). Thus, we can hypothesize that the primary effect of CYP1A1 on breast cancer risk in postmenopausal women is in the generation of local metabolism in breast tissue that could influence breast carcinogenesis. CYP1A1 is involved in the metabolism of 2-, 4-, and 16-OH estradiol, by a preferential generation of 2-hydroxylated compounds (Fig. 1; ref. 20). However, CYP1A1 is the major extrahepatic source of 16α-hydroxylated estrogens in the context of low estrogen concentrations, as might be expected to occur in postmenopausal women. Furthermore, it has been hypothesized that a 2-OH/16α-OH estrogen ratio may be associated with increased breast cancer risk (21), although this has not been confirmed in epidemiologic studies (22). The CYP1A1*2C variant studied here is a nonsynonymous missense variant that results in an isoleucine to valine change at amino acid 462. The valine (26α) allele has increased inducibility to produce catecholestrogens (23, 24). Therefore, we would hypothesize that the increased production of catecholestrogens, particularly the 16α form in breast tissue, by CYP1A1*2C would be associated with increased breast cancer risk.

Associations of CYP1A1 and breast cancer have been widely reported, but these associations have been inconsistent with some papers reporting positive associations (25-29), and others showing no main effect (13, 30, 31) or effects only in the context of exposure such as smoking (32). However, our association is in the opposite direction of that previously reported by Doherty et al. (10), who reported that this polymorphism is inconsistent with their findings...
inversely associated with breast cancer. Many (but not all) of the null results may have been underpowered to detect increased ORs <2.0, despite reports of ORs of 1.5 to 2.0. Therefore, although we report a significant main effect of CYP1A1*2C in a population-based sample of European American women, it is not clear that this association is sufficiently robust across the literature to conclude that it is a breast cancer risk factor independent of other genes or exposures. Because of this discrepancy, the association between CYP1A1 and breast cancer risk cannot be confirmed.

In addition to main effects, we identified three significant first-order interactions among the genes studied here. Each of these interactions involved one member of the cytochrome P450 multigene family known to generate potentially genotoxic catecholestrogens (i.e., CYP1A1, CYP1A2, or CYP1B1) and a member of the sulfotransferase family (i.e., SULT1A1 or SULT1E1) involved in the sulfation of estrogens and catecholestrogens to more hydrophilic less biologically active forms. Sulfation of estrogens may protect breast cells from the mitogenic and DNA-damaging activities of estradiol and its catecholestrogen metabolites (33). Sulfation of catecholestrogens can also compete with their methylation (Fig. 1), the product of which is antimitogenic and antiproliferative (33). The biological function of the SULT1A1*2 allele is well described to have lower enzyme thermostability, lower enzyme activity, and lower estrogen sulfation ability than the nonvariant form (36, 37). The finding that sulfation by SULT1A1 or SULT1E1 interacts with members of the cytochrome P450 multigene family to predispose to breast cancer risk suggests that estrogen and catecholestrogen sulfation is an important mechanism of breast carcinogenesis.

The first significant interaction we identified involved CYP1A1*2C and SULT1A1*2 alleles in European Americans. CYP1A1*2C alleles were associated with significantly increased risk in women who inherited SULT1A1*1 variants. As described above, CYP1A2*2C is associated with increased 2-OH-catecholestrogen formation, which is the major species of catecholestrogen and may play a major role in postmenopausal exposure to catecholestrogens because of its local activity in adipose tissue (20). High levels of 2-OH catecholestrogen have been hypothesized to be protective of breast cancer, particularly in conjunction with low 16α-OH-catecholestrogens. That is, a low 2-OH/16α-OH catecholestrogen ratio may be associated with increased breast cancer risk (21). In addition, SULT1A1*2 is associated with impaired estrogen sulfation. The combination of impaired estrogen sulfation to limit catecholestrogen formation, along with increased generation of the protective 2-OH catecholestrogen formation is consistent with the protective interaction of CYP1A1*2C and SULT1A1*2 that we have observed. In particular, postmenopausal European American women who have inherited any SULT1A1*2 and any CYP1A1*2C allele are significantly protected from developing breast cancer (OR, 0.14; 95% CI, 0.04-0.56), compared with women with only a CYP1A1*2C allele.

The second interaction observed here involves CYP1A2*1F and SULT1E1 —64G>A promoter variant in European Americans. CYP1A2*1F was associated with increased risk in women who carried a variant SULT1E1 A allele. CYP1A2*1F is involved in the generation of both 2-OH and 4-OH catecholestrogens. As described above, 2-OH catecholestrogens are protectors of breast cancer risk, whereas 4-OH catecholestrogens have strong estrogenic activity, are associated with free radical generation, and are genotoxic (33). Because CYP1A2*1F is involved in the generation of both 2-OH and 4-OH catecholestrogens (38), our data are consistent with the hypothesis that breast cancer is elevated among women who carry both CYP1A2*1F alleles and an impaired catecholestrogen sulfation ability conferred by SULT1E1. However, the function of the SULT1E1 —64G>A allele is not well known; thus, this hypothesis will only be supported if additional research confirms that this allele is associated with impaired catecholestrogen sulfation. Assuming this to be the case, our data suggest that the effect of CYP1A2*1F to generate 4-OH catecholestrogens predominates in postmenopausal European American women to confer breast cancer risk, in conjunction with impaired catecholestrogen sulfation by SULT1E1. In our data, women were observed to be at substantially increased breast cancer risk if they had inherited both CYP1A2*1F and SULT1E1 alleles (OR, 6.77; 95% CI, 1.57-29.20) compared with women who inherited no CYP1A2*1F allele but had still inherited a SULT1E1 allele.

The third interaction observed here involves CYP1B1*4 and SULT1A1*2 alleles in African Americans. CYP1B1*4 was associated with increased risk in women who carried SULT1A1*2 alleles. CYP1B1*4 is associated with the increased generation of 4-OH catecholestrogens, which have strong estrogenic activity and are genotoxic. Therefore, breast cancer risk would be predicted to be elevated among women who carry both CYP1B1*4 and an impaired catecholestrogen sulfation ability conferred by SULT1A1*2. In our data, women were observed to be at substantially increased risk if they had inherited both alleles (OR, 3.60; 95% CI, 1.02-12.70), compared with women with only the CYP1B1*4 allele.

Despite the biological plausibility of the associations reported here, there are limitations to the present study. First, we have limited our inferences to only one SNP in each of our candidate genes. There are multiple other potentially functional SNPs in many of the candidate we studied, and it is possible that we have not identified important associations because of our limited SNP choice. However, we chose to limit the number of SNPs considered here to limit the potential for false-positive associations by conducting an excessive number of hypothesis tests. We specified 56 a priori hypotheses (14 associated with main effects and 42 associated with interactions). Although this is not an excessive number of hypothesis tests, it is still possible that we have detected false-positive associations. We have computed the FPRP of Wacholder et al. (39) to aid in interpreting our findings. In situations in which we have observed statistically significant associations, the FPRP supported the inferences that these results represent “noteworthy” associations unlikely to be explained by false-positive results. We generated prior
probabilities of association based on the strong evidence from the literature about prior association studies, the functional relevance of the genes considered here, and the putative functional significance of the variants in these genes, as described above. Because the “true” prior probability for an association is difficult to determine, we follow the recommendations of Wacholder et al. (19) and considered a range of prior probabilities that represent low to high values. Despite the support from these calculations that the associations are noteworthy, we cannot rule out the possibility that these associations are due to chance and therefore need to be confirmed in additional large epidemiologic studies. Another concern is that our study sample comprised both African Americans and European Americans, creating the potential for confounding by ethnicity (population stratification). We have stratified our analyses on race, thus eliminating the majority of the variability that could lead to this confounding (40). In addition, we observed significant main effects in European Americans but not African Americans. Because European Americans are less prone to false-positive associations than a more highly admixed population of African Americans (39, 40), it is unlikely that population stratification has led to serious false inferences of association. Although we have limited our analyses to evaluate first-order interactions between genotypes, the sample sizes for the joint effects of these factors became very small in some situations. Therefore, we may have missed some important effects in groups for which the joint genotype-genotype groups were small. Therefore, nonsignificant interactions involving small cell sizes should be interpreted with caution, and future large-scale studies of interaction among these genes will be required to confirm that these are in fact null associations. This study was not designed to be able to address the three-way interaction involving two genes and race. Therefore, we cannot formally infer whether race-specific differences existed in the gene–gene interactions observed here. Therefore, any differences between African American and European American groups seen here could be explained by chance alone, and we cannot infer any biologically meaningful effects from our data. As with many population-based case-control studies, response rates were relatively low. We estimated that 78% of eligible and accessible cases and controls participated in this study. Although many studies of this type report rates of this magnitude, we acknowledge that these rates may have induced responder biases that could have affected the results of our study. However, because individuals are not aware of their genotypes, it is also less likely that the genotype results may be influenced by the expected forms of bias compared with biases in responses to questionnaire administration. Finally, we have only considered the effects of estrogen metabolism. It is likely that other steroid hormones, particularly progesterone or progestins, also affect breast cancer risk, as suggested by recent findings of excess breast cancer risk in women who used combined hormone replacement therapy that includes progestins (41). Future studies will be required to better understand the combined role of estrogen and progesterone/progesterin metabolism on breast cancer risk.

In summary, we have reported novel first-order interactions between breast cancer risk and genotypes in important estrogen metabolism genes that involve a member of the cytochrome P450 multigene family involved in the generation of catecholestrogens, and a gene involved in the sulfation of estrogens and catecholestrogens. These results suggest that combinations of catecholestrogen and sulfation pathways may together be involved in breast carcinogenesis. Despite the a priori biological plausibility of the genes and variants studied here, the ability to predict the direction of effects for interactions among these genes is problematic. For example, sulfation reactions are involved in the metabolism of both estrogens as well as catecholestrogens. A genetic variant associated with impaired sulfation may be associated with decreased estrogenicity due to its impaired estrogen metabolism (i.e., lower breast cancer risk). The same variant may result in lower catecholestrogensulfation leading to increased mutagenicity (i.e., increased breast cancer risk). These effects may only be relevant in the context of specific catecholestrogen generation scenarios. Therefore, the joint effects of multiple genes in pathways may be difficult to predict a priori. Therefore, the present results are hypothesis generating in that they point to new biological interactions in the etiology of breast cancer, and are consistent with the hypothesis that metabolism of estrogens by multiple genes may be associated with breast cancer etiology.

Acknowledgments

We thank Dr. J. A. Grasso for her central role in the development and execution of this research; the database manager Dr. Anita L. Weber, the Project Manager for the Hospital Network Core, Elene Turzo, and the Project Manager for the Field Core, Desiree Burgh, for their incredible efforts in coordinating the logistical aspects of obtaining institutional review board approvals in participating hospitals and for ascertaining and recruiting the large number of subjects in this study; Karen Venuto for managing the tracking database and the vast correspondence involved in this study; Shawn Fernandez for performing extensive quality control checks and helping with the development of the questionnaire database; Alanna Rebeck for assistance with data processing; and the hospitals in the Greater Delaware Valley for their cooperation and the physicians who sponsored our study in these institutions.

References

16. Shatalova EG, Walther SE, Favorovaa VO, Rebeck TR, Blanchard RL. Genetic polymorphisms in human SULT1A1 and UGT1A1 genes associate
22. Ursin G, London S, Stanczyk FZ, et al. Urinary 2-hydroxyestrone/16α-