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. Strom[1,2,3]

[1]Center for Clinical Epidemiology and Biostatistics and Department of Biostatistics and Epidemiology; [2]Abramson Cancer Center; [3]Department of Medicine; [4]Division of Oncology, Children's Hospital of Philadelphia; [5]Fox Chase Cancer Center, Philadelphia, Pennsylvania; and [6]Department 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

Received 9/20/06; revised 12/14/06; accepted 1/2/07.

Grant support: P01-CA77596.

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.

Note: Current address for R. Blanchard: Merck, Blue Bell, Pennsylvania; Current address for M. Berlin: Center for Women's Health and Department of Obstetrics and Gynecology, Oregon Health and Science University, Portland, Oregon; Current address for J.A. Berlin: Johnson and Johnson, Pharmaceutical Research and Development, Titusville, New Jersey.

Requests for reprints: Timothy Rebbeck, Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, 904 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021. Phone: 215-898-1793; Fax: 215-898-2265. E-mail: trebbeck@cceb.med.upenn.edu

doi:10.1158/1055-9965.EPI-06-0800

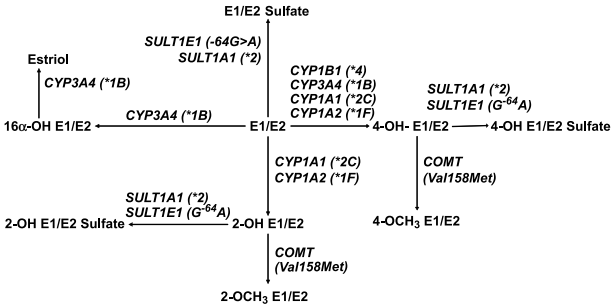

Figure 1. Steroid hormone metabolism pathways with genetic variants studied at these genes shown in parentheses.

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 reactivity. 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 interview (within 3 months). Eligible controls could not have a history of breast cancer. 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.

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-group 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

Gene	SNP nucleotide designation	Rs no.	SNP aliases	Primary variant function	Hypothesized estrogenicity of variant	Hypothesized mutagenicity of variant	Hypothesized breast cancer risk of variant
COMT	1947G>A	rs4680	$Val^{158}Met$, $Val^{108}Met$	Methionine allozyme has decreased activity (42)	—	Increased	Increased
CYP1A1	6750A>G	rs1048943	$Ile^{462}Val$, *2C, m2	Valine allozyme has increased inducibility to produce 2-, 4-, and 16α-catecholestrogens and the main source of extrahepatic 16α catecholestrogen (43)	Weak	Antioxidant, DNA damage protective	Protective
CYP1A2	734C>A, 163C>A	rs762551	*1F	*1F has increased inducibility and ultrarapid activity to produce 2-OH- and 4-OH-catecholestrogens (44)	Weak	Antioxidant, DNA damage protective	Protective
CYP1B1	1358A>G or 3290A>G	rs1800440	$Asn^{452}Ser$, *4	Serine allozyme associated with higher catalytic efficiency toward 4-OH-estradiol (45)	Potent	Genotoxic, Free radical generator	Increased
CYP3A4	729A>G	rs2740574	*1B	*1B allele may confer increased CYP3A4 expression (46)	Potent	Genotoxic, Free radical generator	Increased
SULT1A1	638G>A	rs9282861	$Arg^{213}His$, *2	*2 allozyme has lower thermostability, lower enzyme activity, and lower estrogen sulfation ability (36, 37, 47)	Increased	Increased	Increased
SULT1E1	−64G>A	rs3736599	Promoter variant	Not determined	Unknown	Unknown	Unknown

adequate power for testing first-order interactions (Fig. 1; Table 1). The variants selected were as follows: *COMT* Val^{158}Met (rs4680), *CYP1A1* Ile^{462}Val (*2C; rs1048943), *CYP1A2*1F* (rs762551), *CYP1B1* (Asn^{452}Ser, *4; rs1800440), *CYP3A4*1B* (rs2740574), *SULT1A1* Arg^{213}His (*2; rs9282861), 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 (FPRP) 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 *CYP1A1*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* ($P_{interaction}$ = 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 *CYP1A1*2C* 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 *CYP1A1*2C* 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* ($P_{interaction}$ = 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 of *SULT1E1* and breast cancer among European American women with any *CYP1A2*1F* allele was 1.55 (95% CI,

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

Genotype	European Americans (%)			African Americans (%)		
	Controls	Cases	Adjusted OR* (95% CI)	Controls	Cases	Adjusted OR* (95% CI)
No COMT Met	512 (78%)	377 (77%)	1	372 (91%)	149 (87%)	1
Any COMT Met	147 (22%)	115 (23%)	1.07 (0.81-1.44)	39 (9%)	23 (13%)	1.41 (0.80-2.48)
*No CYP1A1*2C*	682 (94%)	487 (91%)	1	404 (90%)	187 (94%)	1
*Any CYP1A1*2C*	40 (6%)	50 (9%)	1.71 (1.09-2.67)	43 (10%)	11 (6%)	0.62 (0.31-1.25)
*No CYP1A2*1F*	57 (9%)	45 (9%)	1	67 (16%)	32 (19%)	1
*Any CYP1A2*1F*	611 (91%)	451 (91%)	1.03 (0.67-1.57)	340 (84%)	137 (81%)	0.86 (0.53-1.38)
*No CYP1B1*4*	494 (73%)	343 (69%)	1	402 (93%)	164 (89%)	1
*Any CYP1B1*4*	186 (27%)	155 (31%)	1.13 (0.87-1.47)	32 (7%)	20 (11%)	1.48 (0.80-2.72)
*No CYP3A4*1B*	650 (92%)	464 (90%)	1	96 (22%)	43 (24%)	1
*Any CYP3A4*1B*	57 (8%)	54 (10%)	1.35 (0.90-2.02)	332 (78%)	137 (76%)	0.94 (0.62-1.45)
*No SULT1A1*2*	297 (53%)	199 (47%)	1	193 (56%)	85 (59%)	1
*Any SULT1A1*2*	259 (47%)	226 (53%)	1.24 (0.96-1.61)	153 (44%)	59 (41%)	0.87 (0.58-1.32)
No SULT1E1 A	613 (86%)	432 (83%)	1	293 (68%)	130 (71%)	1
Any SULT1E1 A	97 (14%)	89 (17%)	1.30 (0.94-1.79)	141 (32%)	53 (29%)	0.89 (0.60-1.31)

*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.

Table 3. Adjusted ORs, 95% CIs, and *P* values for pairwise genotype interactions and breast cancer case-control status

Gene	*COMT*	*CYP1A1*2C*	*CYP1A2*1F*	*CYP1B1*3*
COMT		14.27 (0.99-204.95) [0.051]	1.45 (0.27-7.78) [0.665]	2.05 (0.58-7.24) [0.263]
*CYP1A1*2C*	0.82 (0.28-2.36) [0.706]		1.65 (0.14-19.49) [0.692]	1.74 (0.39-4.83) [0.470]
*CYP1A2*1F*	1.37 (0.39-7.74) [0.628]	2.79 (0.45-17.5) [0.272]		1.49 (0.53-4.14) [0.449]
*CYP1B1*3*	0.95 (0.46-1.94) [0.880]	2.05 (0.63-6.66) [0.235]	0.68 (0.24-1.94) [0.477]	
*CYP1B1*4*	0.59 (0.31-1.12) [0.104]	0.96 (0.35-2.67) [0.940]	0.57 (0.21-1.54) [0.267]	*
CYP3A4	0.79 (0.25-2.45) [0.680]	3.73 (0.37-37.34) [0.262]	1.06 (0.24-4.77) [0.931]	0.78 (0.31-1.94) [0.590]
*SULT1A1*2*	1.91 (0.97-3.74) [0.061]	0.15 (0.05-0.42) [0.0004]	0.50 (0.18-1.34) [0.167]	0.87 (0.45-1.68) [0.673]
*SULT1A1*3*	0.35 (0.03-8.72) [0.387]	†	†	2.81 (0.29-27.49) [0.374]
SULT1E1	1.20 (0.48-2.96) [0.695]	1.81 (0.41-8.08) [0.440]	6.95 (1.72-28.00) [0.0064]	0.87 (0.40-1.87) [0.715]

NOTE: European American results are shown on the lower left triangle; African American results are shown on the upper right triangle. ORs were 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. Data are OR (95% CI) [*P* value].
*Interactions not computed for variants at the same locus.
† Not estimable.

1.09-2.23). Similarly, the adjusted OR for the effect of *CYP1A2*1F* and breast cancer among European American women with no *SULT1E1* allele was 0.79 (95% CI, 0.49-1.27), whereas the adjusted OR for the effect of *CYP1A2*1F* 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* ($P_{interaction} = 0.041$) in African Americans. The adjusted OR for the effect of *SULT1A1*2* and breast cancer among African American women with no *CYP1B1*4* 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 *CYP1B1*4* 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 0.5%, 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 *CYP1A2*1F* and *SULT1E1*, the FPRP was <20% 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 <20% 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 metabolites in breast tissue that could influence breast carcinogenesis. CYP1A1 is involved in the generation of 2-, 4-, and 16α-hydroxylation of estrone and estradiol, with 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 low 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 (*2C) 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

Table 3. Adjusted ORs, 95% CIs, and *P* values for pairwise genotype interactions and breast cancer case-control status (Cont'd)

*CYP1B1*4*	*CYP3A4*	*SULT1A1*2*	*SULT1A1*3*	*SULT1E1*
2.63 (0.20-35.42) [0.465]	0.61 (0.16-2.26) [0.457]	1.24 (0.26-5.83) [0.784]	0.55 (0.11-2.74) [0.463]	1.44 (0.36-5.80) [0.607]
2.45 (0.30-20.17) [0.405]	0.54 (0.10-2.85) [0.472]	0.78 (0.11-5.28) [0.800]	1.72 (0.27-10.75) [0.564]	0.60 (0.12-3.01) [0.536]
1.18 (0.23-6.04) [0.836]	0.24 (0.06-1.01) [0.052]	0.40 (0.12-1.26) [0.117]	0.56 (0.18-1.75) [0.320]	0.67 (0.23-1.96) [0.462]
*	0.61 (0.24-1.55) [0.301]	0.73 (0.31-1.75) [0.483]	1.74 (0.74-4.12) [0.206]	0.46 (0.20-1.05) [0.066]
	1.14 (0.30-4.24) [0.850]	4.89 (1.06-22.48) [0.041]	2.23 (0.50-9.95) [0.295]	0.56 (0.12-2.67) [0.471]
1.00 (0.38-2.66) [0.993]		0.71 (0.23-2.23) [0.561]	0.80 (0.27-2.44) [0.701]	1.38 (0.47-4.01) [0.559]
0.71 (0.39-1.28) [0.251]	1.99 (0.77-5.11) [0.153]		*	0.65 (0.26-1.64) [0.362]
3.11 (0.47-20.82) [0.242]		*		0.88 (0.35-2.21) [0.785]
1.33 (0.63-2.83) [0.456]	1.37 (0.47-4.00) [0.567]	0.88 (0.42-1.83) [0.724]	1.34 (0.19-9.33) [0.768]	

†

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 (34, 35). 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 P4540 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* genotypes 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 protective 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* 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/progestin 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 P4540 multigene family involved in the generation of catecholestrogen, 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 catecholestrogen sulfation 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. Grisso 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 Fernandes for performing extensive quality control checks and helping with the development of the questionnaire database; Alanna Rebbeck 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

1. Bernstein L, Pike MC, Ross RK, Judd HL, Brown JB, Henderson BE. Estrogen and sex hormone-binding globulin levels in nulliparous and parous women. J Natl Cancer Inst 1985;74:741–5.
2. Madigan MP, Troisi R, Potischman N, Dorgan JF, Brinton LA, Hoover RN. Serum hormone levels in relation to reproductive and lifestyle factors in postmenopausal women (United States). Cancer Causes Control 1998;9:199–207.
3. Potischman N, Swanson CA, Siiteri P, Hoover RN. Reversal of relation between body mass and endogenous estrogen concentrations with menopausal status. J Natl Cancer Inst 1996;88:756–8.
4. Colditz GA. Estrogen, estrogen plus progestin therapy, and risk of breast cancer. Clin Cancer Res 2005;11:909–17s.
5. Rebbeck TR, Troxel AB, Wang Y, et al. Estrogen sulfation genes, hormone replacement therapy, and endometrial cancer risk. J Natl Cancer Inst 2006; 98:1311–20.
6. Haiman C, Hankinson SE, Colditz GA, Hunter DJ, De Vivo I. A polymorphism in CYP17 and breast cancer risk. Cancer Res 2001;61: 3955–60.
7. Paynter RA, Hankinson SE, Colditz GA, Kraft P, Hunter DJ, De Vivo I. CYP19 (aromatase) haplotypes and endometrial cancer risk. Int J Cancer 2005;116:267–74.
8. McGrath M, Hankinson SE, Arbeitman L, Colditz GA, Hunter DJ, De Vivo I. Cytochrome P450 1B1 and catechol-*O*-methyltransferase polymorphisms and endometrial cancer susceptibility. Carcinogenesis 2004;25:559–65.
9. Rylander-Rudqvist T, Wedren S, Granath F, et al. Cytochrome P450 1B1 gene polymorphisms and postmenopausal breast cancer risk. Carcinogenesis 2003;24:1533–9.
10. Doherty JA, Weiss NS, Freeman RJ, et al. Genetic factors in catechol estrogen metabolism in relation to the risk of endometrial cancer. Cancer Epidemiol Biomarkers Prev 2005;14:357–66.
11. Sasaki M, Tanaka Y, Kaneuchi M, Sakuragi N, Dahiya R. CYP1B1 gene polymorphisms have higher risk for breast cancer, and positive correlations with estrogen receptor α and estrogen receptor β expressions. Cancer Res 2003;63:3913–8.
12. De Vivo I, Hankinson SE, Colditz GA, Hunter DJ. A functional polymorphism in the progesterone receptor gene is associated with an increase in breast cancer risk. Cancer Res 2003;63:5236–8.
13. Le Marchand L, Donlon T, Kolonel LN, Henderson BE, Wilkens LR. Estrogen metabolism-related genes and breast cancer risk: the multiethnic cohort study. Cancer Epidemiol Biomarkers Prev 2005;14:1998–2003.
14. Bunin GR, Baumgarten M, Norman SA, Strom BL, Berlin JA. Practical aspects of sharing controls between case-control studies. Pharmacoepidemiol Drug Saf 2005;14:523–30.
15. Strom B, Schinnar R, Weber AL, et al. Case-control study of postmenopausal hormone replacement therapy and endometrial cancer. Am J Epidemiol 2006;164:775–86.
16. Shatalova EG, Walther SE, Favorova OO, Rebbeck TR, Blanchard RL. Genetic polymorphisms in human SULT1A1 and UGT1A1 genes associate

with breast tumor characteristics: a case-series study. Breast Cancer Res 2005;7:R909–21.
17. Feigelson HS, McKean-Cowdin R, Pike MC, et al. Cytochrome P450c17α gene (CYP17) polymorphism predicts use of hormone replacement therapy. Cancer Res 1999;59:3908–10.
18. Kadlubar FF, Berkowitz GS, Delongchamp RR, et al. The CYP3A4*1B variant is related to the onset of puberty, a known risk factor for the development of breast cancer. Cancer Epidemiol Biomarkers Prev 2003;12:327–31.
19. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.
20. Cribb AE, Knight MJ, Dryer D, et al. Role of polymorphic human cytochrome P450 enzymes in estrone oxidation. Cancer Epidemiol Biomarkers Prev 2006;15:551–8.
21. Bradlow HL, Hershcopf R, Fishman JF. Oestradiol 16 α-hydroxylase: a risk marker for breast cancer. Cancer Surv 1986;5:573–83.
22. Ursin G, London S, Stanczyk FZ, et al. Urinary 2-hydroxyestrone/16α-hydroxyestrone ratio and risk of breast cancer in postmenopausal women. J Natl Cancer Inst 1999;91:1067–72.
23. Petersen DD, McKinney CE, Ikeya K, et al. Human CYP1A1 gene: cosegregation of the enzyme inducibility phenotype and an RFLP. Am J Hum Genet 1991;48:720–5.
24. Landi MT, Bertazzi PA, Shields PG, et al. Association between CYP1A1 genotype, mRNA expression and enzymatic activity in humans. Pharmacogenetics 1994;4:242–6.
25. Ambrosone CB, Freudenheim JL, Graham S, et al. Cytochrome P4501A1 and glutathione S-transferase (M1) genetic polymorphisms and postmenopausal breast cancer risk. Cancer Res 1995;55:3483–5.
26. Shen Y, Li DK, Wu J, Zhang Z, Gao E. Joint effects of the CYP1A1 MspI, ERα PvuII, ERα XbaI polymorphisms on the risk of breast cancer: results from a population-based case-control study in Shanghai, China. Cancer Epidemiol Biomarkers Prev 2006;15:342–7.
27. Huang CS, Chern HD, Chang KJ, Cheng CW, Hsu SM, Shen CY. Breast cancer risk associated with genotype polymorphism of the estrogen-metabolizing genes CYP17, CYP1A1, and COMT: a multigenic study on cancer susceptibility. Cancer Res 1999;59:4870–5.
28. Huang CS, Shen CY, Chang KJ, Hsu SM, Chern HD. Cytochrome P4501A1 polymorphism as a susceptibility factor for breast cancer in postmenopausal Chinese women in Taiwan. Br J Cancer 1999;80:1838–43.
29. Miyoshi Y, Takahashi Y, Egawa C, Noguchi S. Breast cancer risk associated with CYP1A1 genetic polymorphisms in Japanese women. Breast J 2002;8:209–15.
30. Bailey LR, Roodi N, Verrier CS, Yee CJ, Dupont WD, Parl FF. Breast cancer and CYPIA1, GSTM1, and GSTT1 polymorphisms: evidence of a lack of association in Caucasians and African Americans. Cancer Res 1998;58:65–70.
31. Modugno F, Zmuda JM, Potter D, et al. Estrogen metabolizing polymorphisms and breast cancer risk among older white women. Breast Cancer Res Treat 2005;93:261–70.
32. Ishibe N, Hankinson SE, Colditz GA, et al. Cigarette smoking, cytochrome P450 1A1 polymorphisms, and breast cancer risk in the Nurses' Health Study. Cancer Res 1998;58:667–71.
33. Yager JD, Liehr JG. Molecular mechanisms of estrogen carcinogenesis. Annu Rev Pharmacol Toxicol 1996;36:203–32.
34. Seegers JC, Aveling ML, Van Aswegen CH, Cross M, Koch F, Joubert WS. The cytotoxic effects of estradiol-17β, catecholestradiols and methoxyestradiols on dividing MCF-7 and HeLa cells. J Steroid Biochem 1989;32:797–809.
35. Fotsis T, Zhang Y, Pepper MS, et al. The endogenous oestrogen metabolite 2-methoxyoestradiol inhibits angiogenesis and suppresses tumour growth. Nature 1994;368:237–9.
36. Raftogianis RB, Wood TC, Otterness DM, Van Loon JA, Weinshilboum RM. Phenol sulfotransferase pharmacogenetics in humans: association of common SULT1A1 alleles with TS PST phenotype. Biochem Biophys Res Commun 1997;239:298–304.
37. Raftogianis RB, Wood TC, Weinshilboum RM. Human phenol sulfotransferases SULT1A2 and SULT1A1: genetic polymorphisms, allozyme properties, and human liver genotype-phenotype correlations. Biochem Pharmacol 1999;58:605–16.
38. Shou M, Korzekwa KR, Brooks EN, Krausz KW, Gonzalez FJ, Gelboin HV. Role of human hepatic cytochrome P450 1A2 and 3A4 in the metabolic activation of estrone. Carcinogenesis 1997;18:207–14.
39. Wacholder S, Rothman N, Caporaso N. Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. J Natl Cancer Inst 2000;92:1151–8.
40. Wang Y, Localio R, Rebbeck TR. Evaluating bias due to population stratification in case-control association studies of admixed populations. Genet Epidemiol 2004;27:14–20.
41. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. JAMA 2002;288:321–33.
42. Lachman HM, Papolos DF, Saito T, Yu YM, Szumlanski CL, Weinshilboum RM. Human catechol-*O*-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. Pharmacogenetics 1996;6:243–50.
43. Cosma G, Crofts F, Taioli E, Toniolo P, Garte S. Relationship between genotype and function of the human CYP1A1 gene. J Toxicol Environ Health 1993;40:309–16.
44. Sachse C, Brockmoller J, Bauer S, Roots I. Functional significance of a C→A polymorphism in intron 1 of the cytochrome P450 CYP1A2 gene tested with caffeine. Br J Clin Pharmacol 1999;47:445–9.
45. Shimada T, Watanabe J, Kawajiri K, et al. Catalytic properties of polymorphic human cytochrome P450 1B1 variants. Carcinogenesis 1999;20:1607–13.
46. Amirimani B, Walker AH, Weber BL, Rebbeck TR. RESPONSE: re: modification of clinical presentation of prostate tumors by a novel genetic variant in CYP3A4. J Natl Cancer Inst 1999;91:1588–90.
47. Adjei AA, Weinshilboum RM. Catecholestrogen sulfation: possible role in carcinogenesis. Biochem Biophys Res Commun 2002;292:402–8.

Pairwise Combinations of Estrogen Metabolism Genotypes in Postmenopausal Breast Cancer Etiology

Timothy R. Rebbeck, Andrea B. Troxel, Amy H. Walker, et al.

Cancer Epidemiol Biomarkers Prev 2007;16:444-450.