

Genetic polymorphisms in the catechol estrogen metabolism pathway and breast cancer risk

Kerryn W. Reding^{1,2}, Noel S. Weiss^{1,2}, Chu Chen^{1,2}, Christopher I. Li^{1,2}, Christopher S. Carlson², Hui-Wen Wilkerson¹, Federico M. Farin¹, Kenneth E. Thummel¹, Janet R. Daling², and Kathleen E. Malone^{1,2}

1. University of Washington, Seattle, WA;

2. Fred Hutchinson Cancer Research Center, Seattle, WA

Running Title: Estrogen metabolizing genes and breast cancer risk

Keywords: breast cancer, catechol estrogen metabolism, genetic variation, combined hormone therapy

Corresponding Author:

Kerryn W. Reding, PhD, MPH

1100 Fairview Ave N.

Mailstop M4-B874

Seattle, WA 98109

kreding@u.washington.edu

Acknowledgements: This work was supported in part by the UW NIEHS sponsored Center for Ecogenetics and Environmental Health, Grant #: NIEHS P30ES07033. KWR was supported by the Cancer Epidemiology and Biostatistics Training Grant (2 T32 CA 09168) and the Department of Defense Breast Cancer Research Program Pre-doctoral Training Grant (06-1-0312). Data and blood specimen collection was supported by R01 CA72787.

Abstract

Background: This study investigated whether single nucleotide polymorphisms (SNPs) in genes within the catechol estrogen metabolism pathway altered the risk of breast cancer alone or in combination, as well as whether menopausal hormone therapy (HT) modified the effect of these SNPs on breast cancer risk.

Methods: In a population-based case-control study of breast cancer, 891 cases and 878 controls were genotyped for six functional SNPs in the *COMT*, *CYP1B1*, *GSTM1*, *GSTP1*, and *GSTT1* genes.

Results: Women homozygous with the T allele in *CYP1B1**2 (Ser119; rs1056827) were at 1.69 (95% confidence interval [CI]: 1.17-2.46) times the risk of women homozygous with the G allele; women homozygous with the G allele in *GSTP1* (Val105; rs1695) were at 0.73 (95% CI: 0.54-0.99) times the risk of breast cancer compared to women homozygous with the A allele. No other SNPs tested were associated with breast cancer to any appreciable degree. Potential gene-gene and gene-HT interactions were investigated.

Conclusion: With the exception of *GSTP1* and possibly *CYP1B1**2, our findings do not provide support for the role of genetic variation in the catechol estrogen metabolism pathway and breast cancer risk in post-menopausal women.

Introduction

A vast accumulation of data has demonstrated that estrogen-related exposures play a role in breast cancer etiology (1,2). Conceivably, this could be due in part to the mutagenic effects of estrogen's intermediate metabolites within the catechol estrogen (CE) pathway. There is evidence that genotoxicity may operate through the formation of reactive estrogen metabolites, namely catechol estrogen semiquinones and quinones which damage DNA via the formation of superoxide radicals and depurinating DNA adducts, although it is unclear what proportion of estradiol is converted into catechol estrogens (3-6).

It is plausible that estrogen's genotoxic potential varies across individuals and may be influenced by genetic variation within the CE metabolism pathway. Specifically, *CYP1B1* plays a role in the conversion of 17 β -estradiol (E₂) into 4-OHE₂ CE (3,5,7-9). Several SNPs within *CYP1B1* have been shown to have functional effects on the catalytic properties of the CYP1B1 enzyme (9,10), and variant alleles of *CYP1B1**2 (Ala119Ser) and *CYP1B1**3 (Leu432Val) have been observed to be associated with breast cancer risk in some (11,12), but not all, previous studies (13-15). Catechol-*O*-methyltransferase (*COMT*) is involved in methylating (and thereby inactivating) CEs (5,16). The *COMT* Met158 allele has been hypothesized to produce an enzyme with reduced functionality, although prior epidemiologic studies have not shown an association with breast cancer (9,14). If CEs are not inactivated, they can be easily oxidized into CE semiquinones and then into CE quinones (5). Members of the glutathione-S-transferase (*GST*) family are thought to play a role in the conjugation of CE quinones, with *GSTP1* being the predominant GST enzyme found in the breast (16,17). The *GSTP1* Val105 allele has

been shown to confer different catalytic enzyme activity (16), however, the epidemiologic findings regarding breast cancer risk have been mixed (14,18-21). Null mutations have been characterized for both *GSTM1* and *GSTT1*; for *GSTM1*, one pooled analysis reported a marginally significant, modestly increased risk of breast cancer associated with the null variant (14), while no increased risk was observed in a meta-analysis (22); for the *GSTT1* null variant, no pooled analyses have observed an association with breast cancer risk (14).

The amount of catechol estrogens accumulating in the breast tissue from daily exposure to estrogen supplied by estrogen-progestin therapy (EPT) or estrogen-alone therapy (ET) may conceivably vary among post-menopausal women and may be affected by the variation in genes within the catechol estrogen metabolism pathway. Several studies have investigated interactions between genes in this pathway, as well as gene-EPT interactions, but no specific gene-gene combination has been observed to be associated with breast cancer in more than one study (19-21,23-25). We investigated whether variation within five genes in the CE metabolism pathway was associated with an increased risk of breast cancer, and whether EPT and/or ET modified the effect of genotype on breast cancer risk.

Materials and Methods

Study Design and Data Collection

The study participants were recruited for a population-based case-control study, the ‘Puget Sound Area Breast Cancer Evaluation (PACE) Study’ of invasive breast cancer among post-menopausal women of 65-79 years of age in western Washington

State. The methods have been described in detail previously (26), and thus are summarized only briefly here. Eligible case participants were: 1) women 65-79 years of age when diagnosed with primary, invasive breast cancer between April 1, 1997 and May 31, 1999; 2) residents of the three county (King, Pierce, and Snohomish counties) Seattle-Tacoma Metropolitan area at diagnosis; 3) women with no previous history of *in situ* or invasive breast cancer; and 4) women with the ability to communicate in English. All cases were ascertained through the Cancer Surveillance System (CSS), the population-based tumor registry serving the Seattle-Puget Sound region of western Washington State and a participant in the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute. The PACE study interviewed 975 cases (80.6 %) of those identified through the CSS.

Eligible control participants for the PACE Study were identified through the Center for Medicare and Medicaid Services' (CMS) list of Social Security recipients from the general population of the same three county Metropolitan Seattle area from which the cases were drawn. The CMS Social Security list includes all individuals age 65 years and older residing in the United States who are eligible for Medicare benefits, including women whose income precludes eligibility to receive funds and women who elect not to receive cash benefits. The controls for this study were randomly selected from the CMS rolls and were frequency matched to the expected distribution of cases by five-year age group. Of the 1,365 controls selected, 1,007 women (73.8%) were interviewed.

Data collection for cases and controls was performed in an identical manner. For each participant, upon the attainment of informed consent, a structured, in-person

interview was conducted by a trained study interviewer on established and suspected breast cancer risk factors, including: demographic characteristics, reproductive history, menstrual history, hormonal contraception history, medical history, certain medications, weight and height history, lifestyle factors, family history of cancer, and hormonal therapy use. The questions on hormonal therapy were extensive and solicited information on lifetime use, including drug name, start and stop dates of use, separate and combined uses of estrogen and progestin, strength, and monthly pattern of pill use. All interview questions were limited to events occurring before each participant's diagnosis date (reference date for controls). Within the 5-year age groups on which controls were frequency matched to cases, reference dates for controls were assigned in a distribution matching the cases' expected diagnosis dates.

Blood samples were provided by 891 (91.4%) of the interviewed cases and 878 (87.1%) of the interviewed controls. This study was approved by the Fred Hutchinson Cancer Research Center's Institutional Review Board.

Laboratory Methods

DNA was extracted from buffy coats using manual phenol chloroform method. Genotyping was performed at the Functional Genomics Laboratory of the Center for Ecogenetics and Environmental Health at the University of Washington.

The 5'-nuclease TaqMan Detection System-based assays were developed to discriminate the following alleles: *CYP1B1**2 (Ala119Ser), *CYP1B1**3 (Leu432Val), *COMT* Val158Met and *GSTP1* Ile105Val (Supplemental Table A.1). Primers and dual-labeled allele specific probes were designed through Assays-by-DesignSM Service-SNP

Genotyping by Applied Biosystems Inc. (Foster City, CA). Each TaqMan minor groove binder (MGB) probe consisted of an oligonucleotide labeled both with a particular 5' reporter dye and a 3' nonfluorescent quencher. Amplification was performed by initial denaturation at 95°C for 10 minutes, followed by 40 cycles of amplification at 92°C for 15 seconds and annealing at 60°C for one minute. End-point analysis was performed on ABI 7900 Sequence Detection System (Applied Biosystems Inc., Foster City, CA) to determine the genotypes.

The presence or absence of the glutathione-S transferase mu (*GSTM1*) and theta (*GSTT1*) genes was determined using a multiplex PCR (27). Briefly, two sets of primers were used to amplify a 215-bp segment of the *GSTM1* gene and a 480-bp segment of the *GSTT1* gene. Primers were also included to amplify a 268-bp segment of the β -globin gene, which was used as a positive control for the PCR. The products were separated by electrophoresis with ethidium bromide-stained 2.5% agarose gel (ISC BioExpress, Kaysville, UT) and genotyped by visual inspection using UV illumination.

Positive controls consisting of DNA aliquots representing wild-type/wild-type, wild-type/mutant, mutant/mutant genotypes (characterized by DNA-sequencing) and a negative control (no DNA) were included in each assay performed. Additionally, randomly selected samples (10%) were identified for replicate testing and were integrated randomly into the genotyping plates. Laboratory staff members were blinded to case-control status and to replicate status.

Among the 10% of the samples for which genotypes were replicated for quality control assessment, the level of agreement between the samples was 100% for the *COMT* Val158Met alleles; 100% for *CYP1B1**2 Ala119Ser; 94.7% for *CYP1B1**3 Leu432Val;

93.2% for *GSTM1**0; 96.2% for *GSTP1* Ile105Val; and 96.4% for *GSTT1**0. When discordance occurred, the genotype was coded as a missing variable. Additionally, the amount of missing data (including those removed for QC purposes) for each genotype was 0.0% (*COMT*), 2.0% (*CYP1B1**2), 0.6% (*CYP1B1**3), 0.5% (*GSTM1*), 0.2% (*GSTP1*), and 0.2% (*GSTT1*).

Statistical Methods

The EPT variable was categorized as never use, short-term use (< 60 months of use), and long-term use (\geq 60 months of use). The ET variable was similarly categorized as never use, short-term use, and long-term use. In addition, analysis examining ET use excluded any women who had ever used EPT due to the association observed between EPT use and breast cancer risk in this dataset.

Concordance between replicate samples included for quality control purposes was determined for each SNP. Testing for deviation from Hardy-Weinberg Equilibrium (HWE) was performed within population controls using a χ^2 goodness of fit test. Mantel-Haenszel Chi-Square test was used for tests of associations in bivariate analyses.

Odds Ratios (OR) and corresponding 95% CIs were calculated using unconditional logistic regression. ORs were adjusted for age and year of diagnosis because these were frequency matched variables. The Wald's test statistic was used to test trends in models containing variables with multiple levels. We also used the Wald's test statistic to determine if the trend in odds ratios associated with various genotypes differed by HT use.

In the multi-gene analysis, we considered only SNPs statistically significantly associated with breast cancer in the single gene analyses. Similarly, only when a single- or multi-gene model was statistically significantly associated with breast cancer risk did we investigate the potential for a gene-hormone therapy interaction. The exception to this was the investigation of interactions which have been previously reported in the literature.

Interaction terms were investigated using the likelihood ratio test (LRT). We estimated *CYP1B1* haplotype frequencies and estimated ORs and 95% CIs associated with breast cancer. Haplotype estimation was performed using Phase v.2.1. The *CYP1B1* haplotype with the highest frequency (G at A119S and G at L432V) served as the referent category.

We accounted for multiple testing in our analyses by using the false-positive reporting probability (FPRP) and preset the FPRP-level criterion at 20% (based on the number of studies previously investigating these genes and our adequate sample size), and evaluated the FPRP using prior probabilities ranging from 0.1 to 0.01 (28).

Results

Demographic and hormone related characteristics of breast cancer cases and controls are presented in Table 1. Similar to findings previously reported in the overall PACE study (26), cases had a higher BMI and were more likely to be long-term users of EPT than controls.

In the single gene analyses, women homozygous for the *CYP1B1**2 Ser allele (T/T) had a 1.69-fold increased risk of breast cancer (95% CI: 1.17-2.46) and the

heterozygotes (G/T) were at no increased risk (OR = 0.99 [95% CI: 0.81-1.20]) compared to homozygous wildtypes (Table 2). Women homozygous for the *GSTP1* Val allele (G/G) were at 0.73 times the risk of breast cancer (95% CI: 0.54-0.99) and the heterozygotes (A/G) had no altered risk (OR = 1.04 [95% CI: 0.85-1.27]) compared to the homozygous wildtype. The *COMT* Met158 allele (rs4680), *CYP1B1* Val432 allele (rs1056836), *GSTMI* null allele, and *GSTT1* null allele were not observed to be associated with risk of breast cancer.

We investigated the possibility of reporting a false positive association between breast cancer and the SNPs in our study by calculating a false positive report probability (FPRP) for prior probabilities ranging from 0.1, 0.05, to 0.01. We observed FPRPs equal to 0.11, 0.21, and 0.58, at these prior probability levels, respectively, for *CYP1B1**2; and FPRPs equal to 0.06, 0.11, and 0.39, respectively, for *GSTP1*, indicating that the false-positive reporting probability was within the 20% criterion set *a priori* for both of these SNPs at a prior probability of 0.1 but not at 0.01.

There was no indication that the associations seen for the homozygous variant allele of *CYP1B1**2 differed according to the use of EPT or ET (Table 3). For *GSTP1*, the reduced risk of breast cancer among women homozygous for the Val allele was limited to never users of ET (p-value for interaction= 0.004). While we had data on duration and recency of use, our sample size was too limited to investigate gene-disease associations stratified by EPT duration and recency.

The two genes found to have associations with breast cancer, *CYP1B1**2 and *GSTP1*, were assessed for a possible gene-gene interaction. However, we observed no evidence of this (p-value = 0.36; Table 4). We also observed no elevation in risk

associated with the *CYP1B1**2-*3 haplotypes (compared to GG, OR = 1.0 [95% CI: 0.8-1.1] for GC; OR = 1.0 [95% CI: 0.8-1.3] for TG; OR = 1.1 [95% CI: 0.9-1.3] for TC; data not shown).

Additionally, we investigated possible interactions between genes previously reported in the literature (19-21,23-25) but did not detect any among the following combined gene variables: *GSTT1-GSTM1*, *COMT-GSTP1*, *COMT-GSTT1*, *COMT-GSTM1*, *COMT-GSTM1-GSTT1*, and *GSTP1-GSTM1-GSTT1* (Supplemental Table 1).

We also investigated a possible EPT interaction with combined genes as reported by Mitrunen, et al, and did not observe evidence that the risk of breast cancer associated with combined genes varied according to EPT use within the *COMT-GSTP1* combined variable (p-value = 0.19; Supplemental Table 2). With the *COMT-GSTT1* and *COMT-GSTM1* combined variables, while we observed a greater risk in short-term users of EPT than was predicted by their separate associations (p-value = 0.004 and 0.001, respectively), the pattern of risk when stratified by EPT was not what we would have expected *a priori* (comparing two to zero high risk genotypes in *COMT-GSTT1*, OR = 0.8 [95% CI: 0.5-1.4] for never users, OR = 24.4 [95% CI: 3.0-200.2] for EPT use < 30 months, OR = 0.8 [95% CI: 0.3-2.4] for EPT use of 30+ months; in *COMT-GSTM1*, OR = 0.7 [95% CI: 0.4-1.1] for never users, OR = 8.5 [95% CI: 1.4-52.4] for EPT use < 30 months, and OR = 1.3 [95% CI: 0.7-2.3] for EPT use of 30+ months), admittedly with a constrained sample size for this analysis.

Our results were unchanged when limited to White women for all of the analyses presented above.

Discussion

Our study observed that variants in the *CYP1B1* and *GSTP1* genes were associated with breast cancer risk. In interpreting these results, we must first consider the potential limitations of our study. Lab errors resulting in non-differential misclassification of genotypes could bias ORs toward the null. However, QC analyses indicated high correlations between replicate samples. Also, allele distributions were consistent with HWE, and there was a minimal amount of missing genotype data.

Another potential limitation was the omission of a portion of the interviewed cases (8.6%) and controls (12.9%) who were otherwise eligible for the parent study but did not donate a blood specimen. Comparison of the women who donated blood with all participants interviewed is reassuring in that there were no discernible differences in risk factor distributions and risk estimates, nor did the cases who donated blood vary from the entire case series in terms of stage and other disease features. Another possible source of selection bias, however, could arise from the 19.4% of cases and 26.2% of controls who were identified by CSS but did not participate in the parent study. For analyses that consider interview data, the ability of women to recall past exposures is a potential concern. Specific tools designed to assist recall were employed in this study, including a lifetime calendar and colored photographs of medications. As a check on the quality of recall in PACE participants, we previously compared self-reported data on several categories of medications with pharmacy records and found overall agreement to be quite good (29).

A large number of individual studies have reported on associations between genes in this pathway and breast cancer risk (11-13,15,18,20-25,30-36). However, no genome-

wide association studies have reported findings for regions encompassing these genes (37,38). In addition to several large individual studies, a meta-analysis of *CYP1B1**2 found no association between breast cancer risk and the Ala/Ser genotype (OR = 1.1 [95% CI: 0.9-1.2]), or the Ser/Ser genotype (OR = 1.0 [95% CI: 0.8-1.2]) (9,11,13,15,31,33). Findings from haplotype analyses have not been any more compelling. Among three large studies investigating the risk of breast cancer associated with haplotypes in the *CYP1B1* gene, one observed a 1.5-fold increased risk comparing women homozygous for G,T,C at positions 48, 119, and 432, respectively, to those with the most common haplotype (95% CI: 1.0-2.1; p-value = 0.03), and two others reported no alteration in breast cancer risk in any *CYP1B1* haplotype (15,31,33). The generally negative findings from prior studies argue that the 1.7-fold increased risk of breast cancer that we observed among women with the *CYP1B1**2 Ser allele should be interpreted with caution.

For *GSTP1*, a pooled analysis of 301 cases and 397 controls, reported that the G/G genotype (Val homozygotes) was associated with an increased risk of breast cancer (OR = 1.86 [95% CI: 1.05-3.3]) (14), but two larger studies have observed either a decreased risk (OR = 0.6 [95% CI: 0.3-1.0]) or no association among the Val homozygotes (OR = 0.9 [95% CI: 0.5-1.4]) (19,20). However, among post-menopausal women these studies have observed borderline decreased risks associated with Val homozygotes compared to wildtype homozygotes (OR = 0.5 [95% CI: 0.2-1.1], and OR = 0.7 [95% CI: 0.4-1.1]) (19,20). In the two studies that stratified by use of HT, one reported a greater reduction in breast cancer risk observed among ever users of ET (OR = 0.2 [95% CI: 0.04-0.8]) than never users (OR = 0.6 [95% CI: 0.2-2.0]).(20) But another

study observed no differential association with Val homozygosity according to use of EPT (OR = 0.8 [95% CI: 0.5-1.2] for never users; OR = 0.8 [95% CI: 0.4-1.5] for ever users) (20).

Our study, the largest to date examining the association between *GSTP1* and breast cancer risk, observed women with the *GSTP1* Val/Val genotype generally to be at a reduced risk of breast cancer, and is in broad agreement with the findings of other epidemiologic studies reported for post-menopausal women. We evaluated the probability that this finding represented a false-positive association and calculated the likelihood of that to be 11% at a prior probability of 0.05 (at lower prior probabilities, the false-reporting probability is higher). Thus, if the prior probability of an association between the *GSTP1* Val/Val allele is at least 5% (given the *a priori* biological rationale for investigating SNPs within this gene, it is reasonable to assume the prior probability is 5% or greater) the probability of this finding being a false-positive association is relatively small.

Additionally, while it is plausible that the reduction in breast cancer risk associated with the *GSTP1* Val/Val genotype is most pronounced in never users of ET, these findings would need to be replicated before any inferences can be made. Adding complexity to the issue are the conflicting findings demonstrating varying levels of catalytic activity between the variant protein and the wildtype protein. Some demonstrated a higher Vmax for the Val105 isoform, while others showed a lower catalytic efficiency for the Val105 isoform compared to wildtype (39-42).

Our null findings with *COMT*, *GSTM1*, and *GSTT1* and breast cancer are in broad agreement with previous reports from pooled and meta-analyses (9,14,22). Also, if

account is taken of the effect modification by age in the HuGE review of *CYP1B1**3 (for Val/Leu compared to Leu/Leu genotypes in Caucasians women over 60 years of age: OR = 1.1 [95% CI: 0.8-1.4]; 51-54 year olds: OR = 2.5 [95% CI: 1.1-5.7]; 55-59 year olds: OR = 1.9 [95% CI: 1.1-3.6]) (43), then our findings in this study of women aged 65 and older are consistent with prior works as well.

Previous studies have also investigated possible gene-gene combinations within this set of genes as there is strong biologic rationale for considering the joint effects of genes within this shared pathway, with several studies (18,20,23-25,44) but not all (21), reporting associations between high risk genotypes from combined genes in this pathway and breast cancer risk. However, with the exception of two studies reporting marginally significant increased risks of breast cancer associated with combined *GSTP1*, *GSTM1*, and *GSTT1*, no specific gene-gene combination has been observed in more than one study (18,20,23-25,44). We investigated but did not detect any gene-gene interactions within the catechol estrogen pathway.

A previous study reported on a potential interaction between combined gene variables and EPT in breast cancer (24). Specifically, among EPT users of greater than 30 months, Mitrunen, et al observed the risk of breast cancer to be heightened in relation to *COMT-GSTM1* (OR = 9.1 [95% CI: 1.8-45.0]), *COMT-GSTT1* (OR = 8.4 [95% CI: 1.4-49.0]), and *COMT-GSTP1* (OR = 7.0 [95% CI: 1.2-40.6] comparing 2 to 0 high risk genotypes). An increased risk was not similarly observed among EPT users of less than 30 months (*COMT-GSTM1*: OR = 0.7 [95% CI: 0.2-3.5] comparing 2 to 0 high risk genotypes); *COMT-GSTT1*: OR = 1.7 [95% CI: 0.3-10.5]), and *COMT-GSTP1* (OR = 2.3 [95% CI: 0.4-14.1]). While we observed a greater risk associated with two high risk

genotypes for short-term users of EPT for both the *COMT-GSTT1* combined variable and the *COMT-GSTM1* combined variable, these patterns of risk were not what we had hypothesized *a priori* (nor were the patterns in agreement with those reported by Mitrunen, et al), and thus, we cannot draw any conclusions based on these findings.

Our study observed a 27% reduction in breast cancer risk associated with the *GSTP1* Val/Val genotype. However, we are the first study to report a statistically significant decreased risk among post-menopausal women; clearly, replication of our results is needed before any firm conclusion can be drawn. Based on this study and the existing epidemiologic literature, there is little evidence that any SNPs in the catechol estrogen metabolism pathway, with the exception of *GSTP1*, have a main effect on breast cancer risk and there is little evidence supporting interactions between any of these genes and HT use.

References

1. Feigelson HS. Breast cancer: epidemiology and molecular endocrinology. In: Henderson BE, Ponder B, Ross RK. Hormones, Genes, and Cancer. New York: Oxford University Press, 2003:120-38.
2. Thomas HV, Reeves GK, Key TJ. Endogenous estrogen and postmenopausal breast cancer: a quantitative review. *Cancer Causes Control* 1997;8:922-8.
3. Cavalieri EL, Stack DE, Devanesan PD, et al. Molecular origin of cancer: catechol estrogen-3,4-quinones as endogenous tumor initiators. *Proc.Natl Acad.Sci U.S.A* 1997;94:10937-42.
4. Yager JD. Endogenous estrogens as carcinogens through metabolic activation. *J Natl Cancer Inst Monogr* 2000;67-73.
5. Cavalieri E, Frenkel K, Liehr JG, Rogan E, Roy D. Estrogens as endogenous genotoxic agents--DNA adducts and mutations. *J Natl Cancer Inst Monogr* 2000;75-93.
6. Yager JD, Liehr JG. Molecular mechanisms of estrogen carcinogenesis. *Annu.Rev Pharmacol.Toxicol.* 1996;36:203-32.
7. Goth-Goldstein, R., Erdmann, C. A., and Russell, M. Cytochrome P4501B1 expression in normal breast tissue. *Polycyclic Aromatic Compounds* 23(259), 275. 2003.

8. Murray GI, Melvin WT, Greenlee WF, Burke MD. Regulation, function, and tissue-specific expression of cytochrome P450 CYP1B1. *Annu.Rev Pharmacol.Toxicol.* 2001;41:297-316.
9. Wen W, Cai Q, Shu XO, et al. Cytochrome P450 1B1 and catechol-O-methyltransferase genetic polymorphisms and breast cancer risk in Chinese women: results from the shanghai breast cancer study and a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2005;14:329-35.
10. Mitrunen K, Hirvonen A. Molecular epidemiology of sporadic breast cancer. The role of polymorphic genes involved in oestrogen biosynthesis and metabolism. *Mutat.Res* 2003;544:9-41.
11. Watanabe J, Shimada T, Gillam EM, et al. Association of CYP1B1 genetic polymorphism with incidence to breast and lung cancer. *Pharmacogenetics* 2000;10:25-33.
12. Zheng W, Xie DW, Jin F, et al. Genetic polymorphism of cytochrome P450-1B1 and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2000;9:147-50.
13. Rylander-Rudqvist T, Wedren S, Granath F, et al. Cytochrome P450 1B1 gene polymorphisms and postmenopausal breast cancer risk. *Carcinogenesis* 2003;24:1533-9.
14. de Jong MM, Nolte IM, te Meerman GJ, et al. Genes other than BRCA1 and BRCA2 involved in breast cancer susceptibility. *J Med Genet* 2002;39:225-42.

15. Dunning AM, Dowsett M, Healey CS, et al. Polymorphisms associated with circulating sex hormone levels in postmenopausal women. *J Natl Cancer Inst* 2004;96:936-45.
16. Raftogianis R, Creveling C, Weinshilboum R, Weisz J. Estrogen metabolism by conjugation. *J Natl. Cancer Inst. Monogr* 2000;113-24.
17. Kelley MK, Engqvist-Goldstein A, Montali JA, Wheatley JB, Schmidt DE, Kauvar LM. Variability of glutathione S-transferase isoenzyme patterns in matched normal and cancer human breast tissue. *Biochem J* 1994;304:843-8.
18. Helzlsouer KJ, Selmin O, Huang HY, et al. Association between glutathione S-transferase M1, P1, and T1 genetic polymorphisms and development of breast cancer. *J Natl Cancer Inst* 1998;90:512-8.
19. Millikan R, Pittman G, Tse CK, Savitz DA, Newman B, Bell D. Glutathione S-transferases M1, T1, and P1 and breast cancer. *Cancer Epidemiol Biomarkers Prev* 2000;9:567-73.
20. Mitrunen K, Jourenkova N, Kataja V, et al. Glutathione S-transferase M1, M3, P1, and T1 genetic polymorphisms and susceptibility to breast cancer. *Cancer Epidemiol Biomarkers Prev* 2001;10:229-36.
21. Curran JE, Weinstein SR, Griffiths LR. Polymorphisms of glutathione S-transferase genes (GSTM1, GSTP1 and GSTT1) and breast cancer susceptibility. *Cancer Lett.* 2000;153:113-20.

22. Sull JW, Ohrr H, Kang DR, Nam CM. Glutathione S-transferase M1 status and breast cancer risk: a meta-analysis. *Yonsei Med J* 2004;45:683-9.
23. Park SK, Yim DS, Yoon KS, et al. Combined effect of GSTM1, GSTT1, and COMT genotypes in individual breast cancer risk. *Breast Cancer Res Treat.* 2004;88:55-62.
24. Mitrunen K, Kataja V, Eskelinen M, et al. Combined COMT and GST genotypes and hormone replacement therapy associated breast cancer risk. *Pharmacogenetics* 2002;12:67-72.
25. Cheng TC, Chen ST, Huang CS, et al. Breast cancer risk associated with genotype polymorphism of the catechol estrogen-metabolizing genes: a multigenic study on cancer susceptibility. *Int J Cancer* 2005;113:345-53.
26. Li CI, Malone KE, Porter PL, et al. Relationship between long durations and different regimens of hormone therapy and risk of breast cancer. *JAMA* 2003;289:3254-63.
27. Sweeney C, Farrow DC, Schwartz SM, Eaton DL, Checkoway H, Vaughan TL. Glutathione S-transferase M1, T1, and P1 polymorphisms as risk factors for renal cell carcinoma: a case-control study. *Cancer Epidemiol Biomarkers Prev* 2000;9:449-54.
28. Wacholder S, Chanock S, Garcia-Closas M, El GL, 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.

29. Boudreau DM, Daling JR, Malone KE, Gardner JS, Blough DK, Heckbert SR. A validation study of patient interview data and pharmacy records for antihypertensive, statin, and antidepressant medication use among older women. *Am J Epidemiol* 2004;159:308-17.
30. De Vivo I, Hankinson SE, Li L, Colditz GA, Hunter DJ. Association of CYP1B1 polymorphisms and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2002;11:489-92.
31. Gaudet MM, Chanock S, Lissowska J, et al. Genetic variation of Cytochrome P450 1B1 (CYP1B1) and risk of breast cancer among Polish women. *Pharmacogenet. Genomics* 2006;16:547-53.
32. Kocabas NA, Sardas S, Cholerton S, Daly AK, Karakaya AE. Cytochrome P450 CYP1B1 and catechol O-methyltransferase (COMT) genetic polymorphisms and breast cancer susceptibility in a Turkish population. *Arch.Toxicol.* 2002;76:643-9.
33. Matyjasik J, Cybulski C, Masojc B, et al. CYP1B1 and predisposition to breast cancer in Poland. *Breast Cancer Res Treat.* 2007;106:383-8.
34. Millikan RC, Pittman GS, Tse CK, et al. Catechol-O-methyltransferase and breast cancer risk. *Carcinogenesis* 1998;19:1943-7.
35. Mitrunen K, Jourenkova N, Kataja V, et al. Polymorphic catechol-O-methyltransferase gene and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2001;10:635-40.

36. Thompson PA, Shields PG, Freudenheim JL, et al. Genetic polymorphisms in catechol-O-methyltransferase, menopausal status, and breast cancer risk. *Cancer Res* 1998;58:2107-10.
37. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat.Genet.* 2007.
38. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447:1087-93.
39. Ali-Osman F, Akande O, Antoun G, Mao JX, Buolamwini J. Molecular cloning, characterization, and expression in *Escherichia coli* of full-length cDNAs of three human glutathione S-transferase Pi gene variants. Evidence for differential catalytic activity of the encoded proteins. *J Biol.Chem.* 1997;272:10004-12.
40. Sundberg K, Johansson AS, Stenberg G, et al. Differences in the catalytic efficiencies of allelic variants of glutathione transferase P1-1 towards carcinogenic diol epoxides of polycyclic aromatic hydrocarbons. *Carcinogenesis* 1998;19:433-6.
41. Zimniak P, Nanduri B, Pikula S, et al. Naturally occurring human glutathione S-transferase GSTP1-1 isoforms with isoleucine and valine in position 104 differ in enzymic properties. *Eur.J Biochem.* 1994;224:893-9.
42. Eaton DL, Bammler TK. Concise review of the glutathione S-transferases and their significance to toxicology. *Toxicol.Sci* 1999;49:156-64.

43. Paracchini V, Raimondi S, Gram IT, et al. Meta- and Pooled Analyses of the Cytochrome P-450 1B1 Val432Leu Polymorphism and Breast Cancer: A HuGE-GSEC Review. *Am J Epidemiol* 2006.
44. Spurdle AB, Chang JH, Byrnes GB, et al. A systematic approach to analysing gene-gene interactions: polymorphisms at the microsomal epoxide hydrolase EPHX and glutathione S-transferase GSTM1, GSTT1, and GSTP1 loci and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2007;16:769-74.

Table 1. Characteristics of breast cancer cases and controls

Characteristic		Case n (%)^{1,2}	Control n (%)^{1,2}
		n=891	n=878
Age at reference date (yrs)	65-69	278 (31.2)	278 (31.7)
	70-74	613 (68.8)	600 (68.3)
Age at menopause (yrs) ³	< 39	74 (9.6)	98 (11.2)
	40-44	115 (14.9)	130 (14.8)
	45-49	233 (30.1)	229 (26.1)
	50-54	257 (33.2)	239 (27.2)
	≥ 55	95 (12.3)	103 (11.7)
	Missing data	117	79
Age at First FTP (yrs)	< 30	720 (81.1)	726 (82.8)
	≥ 30	168 (18.9)	151 (17.2)
	Missing data	3	1
Cause of menopause	Natural menopause	508 (57.9)	519 (59.9)
	Induced menopause	122 (13.9)	124 (14.3)
	Simple Hysterectomy	207 (23.6)	198 (22.8)
	Other	41 (4.7)	26 (3.0)
	Missing data	13	11
BMI quartiles	16.00-22.96	169 (19.5)	211 (24.8)
	22.97- 26.01	221 (25.6)	213 (25.1)
	26.02-30.11	244 (28.2)	211 (24.8)
	30.12- 48.70	231 (26.7)	214 (25.2)
	Missing data	26	29
Mean BMI		27.5 (5.7)	26.9 (5.4)
EPT	Never use	648 (73.0)	697 (79.8)
	< 60 months of use.	80 (9.0)	80 (9.2)
	≥ 60 months of use	160 (18.0)	96 (11.0)
	Missing data	3	5
ET	Never use	407 (46.1)	412 (47.2)
	< 60 months of use.	157 (17.8)	176 (20.2)
	≥ 60 months of use	319 (36.1)	284 (32.6)
	Missing data	8	6
Race	Caucasian	854 (95.9)	814 (92.7)
	African American	11 (1.2)	27 (3.1)

Estrogen Metabolizing Genes and Breast Cancer Risk

Asian	17 (1.9)	21 (2.4)
Other/unknown	9 (1.0)	16 (1.8)

FTP: Full term pregnancy; BMI: Body Mass Index

1. with the exception of mean BMI: mean (standard deviation)
2. Some percentages do not sum to 100% due to rounding
3. Using data imputed from hormone therapy use and bilateral oophorectomy status

Table 2. The risk of breast cancer associated with genetic variation as investigated in single gene models

Gene	Genotype	Case n (%)	Control n (%)	OR ¹	95% CI
<i>COMT</i>	G/G (Val/Val)	224 (25.1)	211 (24.0)	1.0 (ref)	
	G/A (Val/Met)	427 (48.0)	431 (49.1)	0.94	0.74 1.18
	A/A (Met/Met)	240 (26.9)	236 (26.9)	0.97	0.75 1.26
<i>CYP1B1*2</i>	G/G (Ala/Ala)	452 (51.5)	454 (53.0)	1.0 (ref)	
	G/T (Ala/Ser)	341 (38.8)	353 (41.2)	0.99	0.81 1.20
	T/T (Ser/Ser)	84 (9.6)	50 (5.8)	1.69	1.17 2.46
<i>CYP1B1*3</i>	C/C (Leu/Leu)	289 (32.6)	271 (31.0)	1.0 (ref)	
	C/G (Leu/Val)	409 (46.2)	427 (48.9)	0.90	0.73 1.11
	G/G (Val/Val)	188 (21.2)	176 (20.1)	1.00	0.77 1.30
<i>GSTP1</i>	A/A (Ile/Ile)	382 (42.9)	366 (41.8)	1.0 (ref)	
	A/G (Ile/Val)	417 (46.8)	390 (44.6)	1.04	0.85 1.27
	G/G (Val/Val)	92 (10.3)	119 (13.6)	0.73	0.54 0.99
<i>GSTM1</i>	Present	421 (47.4)	415 (47.4)	1.0 (ref)	
	Null	467 (52.5)	460 (52.6)	1.00	0.83 1.21
<i>GSTT1</i>	Present	744 (83.5)	738 (84.2)	1.0 (ref)	
	Null	147 (16.5)	139 (15.8)	1.04	0.81 1.34

1. Adjusted for age and year of diagnosis (reference date for controls)

Table 3. The risk of breast cancer associated with *CYP1B1**2 and *GSTP1* Ile105Val stratified by EPT and ET use

Genotype	Case ¹	Control ¹	OR ²	95% CI		p-value ³
<i>CYP1B1</i>						
<i>Never EPT Users</i>						
G/G	317 (49.5)	346 (51.0)	1.0 (ref)			
G/T	259 (40.5)	291 (42.9)	0.98	0.78	1.23	
T/T	64 (6.9)	42 (6.2)	1.63	1.07	2.48	0.14
<i>Ever EPT Users</i>						
G/G	133 (56.8)	105 (60.7)	1.0 (ref)			
G/T	81 (34.6)	60 (34.7)	1.07	0.70	1.62	
T/T	20 (8.6)	8 (4.6)	1.90	0.80	4.51	0.001
				p _{int} = 0.96		
<i>Never ET Users</i>						
G/G	134 (49.3)	155 (50.2)	1.0 (ref)			
G/T	113 (41.5)	137 (44.3)	0.98	0.70	1.38	
T/T	25 (9.2)	17 (5.5)	1.69	0.87	3.27	0.91
<i>Ever ET Users⁴</i>						
G/G	179 (49.3)	191 (51.6)	1.0 (ref)			
G/T	145 (39.9)	154 (41.6)	1.01	0.74	1.38	
T/T	39 (10.7)	25 (6.8)	1.57	0.91	2.72	0.23
				p _{int} = 0.64		
missing						
	14 (1.6)	21 (2.4)				
<i>GSTP1</i>						
<i>Never EPT Users</i>						
A/A	284 (43.8)	284 (40.8)	1.0 (ref)			
A/G	293 (45.2)	315 (45.2)	0.94	0.75	1.18	
G/G	71 (11.0)	95 (13.6)	0.73	0.51	1.03	0.11
<i>Ever EPT Users</i>						
A/A	97 (40.4)	80 (45.4)	1.0 (ref)			
A/G	122 (50.8)	72 (40.9)	1.40	0.92	2.11	
G/G	21 (8.8)	24 (13.6)	0.73	0.38	1.41	0.97
				p _{int} = 0.41		
<i>Never ET Users</i>						
A/A	125 (45.3)	103 (32.8)	1.0 (ref)			
A/G	125 (45.3)	166 (52.9)	0.63	0.44	0.90	
G/G	26 (9.4)	45 (14.3)	0.45	0.26	0.78	0.001
<i>Ever ET Users⁴</i>						
A/A	55 (59.8)	37 (40.2)	1.0 (ref)			
A/G	77 (75.5)	25 (24.5)	1.30	0.96	1.78	
G/G	12 (46.2)	14 (53.8)	0.97	0.61	1.54	0.53
				p _{int} = 0.004		
missing						
	0 (0.0)	3 (0.3)				

1. Some percentages do not sum to 100% due to rounding

Estrogen Metabolizing Genes and Breast Cancer Risk

2. Adjusted for age and year of diagnosis (reference date for controls).
3. p-value for test of trend
4. Limited to exclusive users of ET

Table 4. The risk of breast cancer associated with genetic variation modeled as *CYP1B1**2- *GSTP1* gene-gene interactions

Case n (%) ¹	Control n (%) ¹	# High Risk Genotypes ²		OR ³	95% CI	
		<i>GSTP1</i>	<i>CYP1B1</i> *2			
79 (9.0)	108 (12.6)	0	0	1.0 (ref)		
10 (1.1)	8 (0.9)	0	1	1.81	0.64	5.07
714 (81.4)	696 (81.5)	1	0	1.53	0.99	2.35
74 (8.4)	42 (4.9)	1	1	2.78	1.46	5.32
				p _{int} = 0.36		

1. Some percentages do not sum to 100% due to rounding
2. High risk genotypes as determined in the single gene analyses: A/A and A/G for *GSTP1* and T/T for *CYP1B1*.
3. Adjusted for age and year of diagnosis (reference date for controls)