Published in final edited form as:

Cancer Epidemiol. 2009 July ; 33(1): 24–30. doi:10.1016/j.canep.2009.04.009.

African American-preponderant single nucleotide polymorphisms (SNPs) and risk of breast cancer

Ikuko Kato^{a,b,*}, Michelle Cichon^a, Cecilia L. Yee^a, Susan Land^{a,c}, and Jeannette F. Korczak^{a,d}

^a Karmanos Cancer Institute, Wayne State University School of Medicine, Detroit, MI 48201

^b Department of Pathology, Wayne State University School of Medicine, Detroit, MI 48201

^c Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, MI 48201

^d Division of Extramural Activities, National Cancer Institute, Bethesda, MD 20892

Abstract

Background—African American women more often present with more aggressive types of breast cancer than Caucasian women, but little is known whether genetic polymorphisms specific to or disproportionate in African Americans are associated with their risk of breast cancer.

Methods—A population-based case-control study was conducted including 194 cases identified through the Metropolitan Detroit Cancer Surveillance System and 189 controls recruited through random digit dialing to examine polymorphisms in genes involved in estrogen metabolism and action.

Results—The African American-specific *CYP1A1* 5639C allele was associated with an increased risk of breast cancer (odds ratio(OR)=2.34, 95%confidence interval (CI): 1.23–4.44) and this association with the *CYP1A1* 5639 locus was dependent on another polymorphism in the *CYP3A4* gene (P=0.043 for the interaction). In addition, African American-predominant *CYP1B1* 432 Val allele was significantly more often found in the cases than in the controls overall and the *HSD17B1* 312 Gly allele was specifically associated with premenopausal breast cancer risk (OR=3.00, 95% CI: 1.29–6.99).

Conclusion—These observations need to be confirmed in larger studies due to the limited statistical power of the study based on a small number of cases.

Keywords

Breast cancer; African American; SNP; case-control study; estrogen

INTRODUCTION

Breast cancer in African American women is characterized by younger age at onset, more advanced stage at diagnosis and aggressive histological features, including higher nuclear

*Correspondence: Ikuko Kato, 110 East Warren Avenue, Detroit, MI 48201, telephone 313-578-4206, fax 313-578-4306, katoi@karmanos.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

grade and the absence of steroid hormone receptor expression, than in Caucasian women [1–2]. These characteristics are believed to contribute to poorer survival and, thus, higher mortality from breast cancer in African American women compared to Caucasians [3]. Genetic susceptibility may account for some of these differences, as epidemiological studies have failed to delineate differences in conventional risk factors for breast cancer between the two racial groups [4–6]

The present study focuses on selected single nucleotide polymorphisms (SNPs) in genes involved in estrogen metabolism or action, which are African-American-specific or else differ considerably in allele frequencies between African Americans and Caucasians, but were previously associated with risk of estrogen-related diseases, cancer or circulating hormone levels [7–15]. Specifically, these genes include (1) cytochrome P450 (*CYP*) 17, whose product catalyzes the synthesis of precursors of both estrogen and testosterone; (2) 17 β -hydroxysteroid dehydrogenase type 1 (*HSD17B1*), which converts estrone to the more potent estradiol; (3) *CYP1A1* and *CYP3A4*, which preferentially catabolize estrogen to 2-hydroxy products that lack peripheral biological activity; (4) *CYP1B1*, which metabolizes estrogen to 4-hydroxy products that maintain relatively strong biological activity; and (5) estrogen receptor (*ER*)-*beta*, which may antagonize ER-alpha mediated estrogenic action [16–17]. We report here the results of a population-based case-control study of African American breast cancer patients to investigate whether the selected SNPs in these genes are associated with breast cancer risk overall or with particular subsets of breast cancer cases.

MATERIALS AND METHODS

Study population

The study population for this case-control study was comprised of the following three groups of subjects: (1) existing cases recruited for a segregation analysis of African American breast cancer; (2) controls from a population-based case-control study for colorectal cancer conducted in approximately the same time period; and (3) additional population-based controls recruited specifically for the current study. All subjects gave written informed consent to participate in the respective studies. Eligible cases were African American women, between 18 and 79 years of age, who were diagnosed with invasive breast cancer between December 2001 and August 2003, and were identified through a rapid case-reporting system implemented in the Metropolitan Detroit Cancer Surveillance System (MDCSS), one of the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER)-cancer registries [3], which allowed patient access within 3–4 months from their diagnosis. Cases with a prior history of any cancer other than non-melanoma skin cancer were excluded. Of 719 potentially eligible cases, 666 were allowed by their physicians to contact. Among these 260 (39%) participated, while 60 were found to be deceased. Eleven cases who participated were subsequently found to be ineligible and were excluded from the current study. Because biological specimen collection from these breast cancer cases was initially optional, 138 of the eligible women provided a sample. We subsequently tried to re-contact 81 of the participants who did not provide a biospecimen but gave permission to be re-contacted. Of these, 56 provided a biospecimen, while 8 were found to be deceased. Thus, the total case population consisted of 194 women.

134 African American controls through random digit dialing (RDD) were available from the colorectal cancer case-control study (overall control response rate of 59%). Sixty additional controls in younger age groups were recruited through RDD (response rate 71%), resulting in a total of 194 African American women controls. These controls were frequency matched to the cases in 5-year age intervals. Specifically, first we projected the numbers of cases for each 5-year age group by tabulating the cases who donated a biospecimen for the initial segregation analysis and adding 70% (estimated response rate) of the cases in each 5-year

age group who did not donate a biospecimen initially. Then, we selected available controls from the colorectal cancer case-control study and recruited additional controls to match this distribution. During this process, we closely monitored the accrual of additional cases and controls to ensure the balance of the cases and controls by age group and to make any correction for age-distribution of the controls if necessary.

Data and specimen collection

The cases were interviewed in person (66%) or by telephone (34%) for a detailed family history of cancer and information on other basic risk factors for breast cancer. The controls from the colorectal cancer study were interviewed by telephone regarding their usual diet and other basic risk factors for colorectal cancer. The additional controls recruited by RDD for the present study were interviewed by telephone for questions common to both studies. Fifty percent of the cases and 68% of the controls chose a blood draw, 47% of the cases and 32% of the controls provided mouthwash specimens, and 3% of the cases allowed access to their tissue blocks.

SNPs and Genotyping

Table 1 lists the SNPs selected for this study, other allele designations in earlier studies, some of which used different nomenclature to designate the alleles [7–15] and reported minor allele frequencies for African Americans and Caucasians in the NCBI SNP database [18]. A SNP in *CYP17A1* was included, despite relatively small racial variation in allele frequencies, due to a reported association with aggressive types of breast cancer [8]. All laboratory assays were performed at the Wayne State University Applied Genomics Technology Center. DNA was isolated with a Gentra Autopure under standard conditions. TaqMan genotyping assays were employed for SNP analysis using an Applied Biosystems 7900 (Foster City, CA). Representative Taqman plots for the seven SNPs are presented in Fig. 1. For quality control, 10% of the assays were repeated and amplicons for each of the assays were sequenced. For added sensitivity, DNA isolated from low yield samples was preamplified in an outer PCR reaction using oligonucleotide primers designed with Primer Express software (Applied Biosystems).

Statistical analysis

Odds ratios (ORs) and 95% confidence intervals (CIs) for breast cancer associated with each SNP were calculated using an unconditional logistic regression model [19], adjusting for selected covariates. Covariates showing an association with breast cancer risk at the $p < 0.10$ level were tested in a multivariable model. The ORs were obtained for heterozygotes and homozygotes for the minor allele of each SNP, using the homozygotes for the major allele as the referent group, whenever feasible. The gene-dosage effect was also tested using regression coefficients for the number of minor alleles (i.e., 0, 1 or 2) of each SNP. The interactions between pairs of SNPs were tested by including multiplicative interaction terms of these variables and by their p-values for the interaction coefficients, as well as by calculating strata-specific ORs. Analyses were repeated by menopausal status, as well as for some breast cancer characteristics. All statistical analyses were performed using SAS 9.1.

RESULTS

Five control subjects with a missing value for one or more of the selected covariates for the final adjusted model were excluded from the current analyses. Mean ages at diagnosis (for cases) and at the time identified by RDD (for controls) were 53 (range 25–79) and 54 (range 28–80) years, respectively. Among the three selected covariates, higher levels of pack-years of cigarettes smoking (4 levels) and of education (4 levels) were associated with case status, while an indicator variable for decaffeinated coffee usage was twice as prevalent in the

controls as in the cases. The number of cases and controls included in each analysis varied according to the number of subjects with undetermined genotypes, ranging from 177 to 194 for the cases, and from 187 to 189 for the controls. When Hardy-Weinberg equilibrium were tested for the cases and controls separately based on a chi square test, all SNPs were found to be in equilibrium using $p=0.01$ as the threshold, and all but the *CYP1B* polymorphism in the controls ($P=0.046$) using $p=0.05$ as the threshold. Minor allele frequencies were comparable to those reported in Table 1 for *CYP1A1* and *CYP17A1*, somewhat higher for *HSD17B1* (48.5%) and *ER2 (beta)* (13.8%) and lower for *CYP1B1* (24.9%) and *CYP3A4* (34.6%).

Table 2 shows uni-variable and multi-variable ORs associated with each SNP. There were few differences in the associations. The overall risk of breast cancer was significantly increased in women carrying at least one C allele of the *CYP1A1* T5639C polymorphism, yielding an OR of 2.34 (95% CI 1.23–4.44). On the other hand, women with the *CYP1B1* LeuLeu genotype had a significantly reduced risk of breast cancer (OR=0.33, 95% CI: 0.14–0.78) compared with those with at least one Val allele. In addition, a marginally significantly increased risk was observed for the *HSD17B1* Gly allele (OR =1.59, 95% CI 0.98–2.59 for the dominant model).

The ORs were further calculated stratified by menopausal status (Table 3). The associations with the *CYP1A1* T5639C and *CYP1B1* polymorphisms tended to be more pronounced in postmenopausal women, but the differences in the ORs were not statistically significant. On the other hand, the association with the *HSD17B1* Gly allele became more gene dosage dependent ($p=0.009$) in premenopausal women, with ORs for heterozygotes and homozygotes of 2.67 and 4.11, respectively, and a statistically significant difference in the gene-dose effect between pre- and post-menopausal women ($P=0.012$). When the association with *HSD17B1* Gly allele (dominant model) was analyzed by breast cancer subtype, we found a more pronounced increase in risk associated with aggressive phenotypes, i.e., ER/PR-negative (OR = 2.59, 95%CI 1.13–5.92), high nuclear grade (OR=2.23; 95% CI 1.18–4.22) and advanced stage(node-positive, distant stage or local invasion) (OR= 2.08 95% CI 1.05–4.09) (data not shown).

We also examined gene-gene interactions using combinations of indicator variables for dichotomized genotypes of paired polymorphisms. Based on the results of the overall analyses in Table 2, a recessive model was chosen for the *CYP1B1* and *CYP3A4* polymorphisms and a dominant model for the other polymorphisms, and analyses were limited to 12 pairs of polymorphisms where the smallest group of the 4 genotypic combinations included at least 10 subjects. Among those, an antagonistic (negative) interaction between the *CYP1A1* 5639 and *CYP3A4* polymorphisms was found to be statistically significant ($p=0.043$). The main effect of *CYP3A4* was not statistically significant ($P= 0.11$) in this model. In addition, we found a marginally significant antagonistic (negative) interaction between the *CYP1A1* 5639 C allele and *HSD17B1* Gly allele ($P=0.054$). To aid the interpretation of these interactions, we further present in Table 4 the ORs associated with the *CYP1A1* 5639 C allele (TC + CC genotype) according to dichotomized genotypes of the other polymorphisms. All ORs associated with *CYP1A1* 5639 C allele were more pronounced in the absence (left column) of other minor alleles/genotypes, and most ORs (except the one for the *CYP17A1*) were statistically significant in these combinations.

DISCUSSION

Estrogens are known to undergo extensive oxidative metabolism at various positions of the molecules catalyzed by enzymes present in the liver and other peripheral organs, which

primarily belong to CYP families. Three major pathways exist, namely 2-, 4- and 16 α -hydroxylation. While 2- metabolites have weak or no biological activity, 4- and 16 α -metabolites retain strong estrogenic activity [21]. The enzyme CYP1B1 predominantly metabolizes estrogens to 4-hydroxy products [22] which are capable of inducing DNA adduct formation [23] and are carcinogenic in animals [24]. Importantly, this Val432Leu polymorphism has been shown to have functional effects. Among several known SNPs in this gene, the Val432Leu polymorphism exerts the greatest impact on the catalytic properties, with three-fold higher 4-hydroxylation by the Val allele than by the Leu allele [25]. In addition, a small clinical study demonstrated that the ratios of 4- to 2-hydroxy metabolites were higher in subjects with Val alleles [26]. The associations between this polymorphism and breast cancer, however, have been very mixed with large heterogeneity. Paracchini et al summarized in their meta and pooled analyses that the Val allele increases the risk of breast cancer in Caucasians in a dominant fashion (summary OR=1.5, 95% CI: 1.1–2.1), but does not affect the risk of breast cancer in African Americans [27]. Albeit the frequency of Val allele in our African American population (26%) was not different from their summary frequency for African American (29%), the present study found an association similar to that reported for Caucasians. Because only four studies were included in this meta-analysis and because the number of study subjects in each study was rather small, the result of the present study is not necessarily considered to be contradictory.

In contrast to CYP1B1, the major metabolic pathway catalized by CYP1A1 is 2-hydroxylation [22], and thus genetic polymorphisms leading to higher enzymatic activity have been hypothesized to reduce the risk of breast cancer. There are several known SNPs in this gene. The two polymorphisms included in this study reside in the untranslated region, and thus it has been postulated that they primarily act through linkage disequilibrium with other as of yet unidentified polymorphisms. However, functional data for these two polymorphisms have been very inconsistent in terms of messenger RNA expression levels, enzymatic activity or circulating or urinary levels of various estrogen metabolites [7]. The most commonly studied polymorphism thus far is T6235C, which is also referenced as 2A, *m1*, or T3801C. In a meta-analysis to summarize the results from 13 studies, the ORs for the TC and CC genotypes compared with TT were estimated to be 0.91 (95% CI: 0.70–1.19) and 0.97 (95% CI: 0.52–1.80), respectively [7]. Two more recent studies did not find an association of breast cancer with this polymorphism in any ethnic groups [28–29], consistent with our observation. Due to its specificity to the African American population, the data for the T5639C polymorphism, which is also referenced as T3205C or *m3*, have been limited. Three studies of African Americans in New York [30], Tennessee [31], and North Carolina [29] failed to show an association with breast cancer. The discrepancy between our study and others concerning this polymorphism may arise from interactions with other genetic polymorphisms or with environmental factors. Interactions with other genetic polymorphisms in the same metabolic pathway are biologically plausible [32] In this context, the statistically significant interaction of *CYP1A1* (T5639C) with *CYP3A4* (A-392G) observed in this study is interesting, although no prior studies have reported an interaction between this particular combination of loci [7,32]. Moreover, this interaction may partly account for inconsistent observations regarding the association between the *CYP1A1* T5639C and breast cancer in earlier studies.

Sequence homology analysis does not indicate functional significance of the *HSD17B1* Ser312Gly polymorphism, and no associations have been found between this polymorphism and circulating levels of estrone and estradiol [12] Nonetheless, a nested case-control study from a multi-ethnic cohort demonstrated that this polymorphism was associated with the risk of advanced breast cancer in conjunction with a *CYP17A1* polymorphism [33]. The combined effect of the Ser312Gly polymorphism with the *CYP17A1* polymorphism was also noted for breast cancer in Chinese women in Singapore [34], although there was no such

indication in the present study. More recently, Feigelson et al. [35] reported that this polymorphism and a haplotype that includes this SNP were associated with the risk of ER-negative breast cancer. This suggests that the observed association with the Ser312Gly polymorphism was possibly through linkage disequilibrium.

There are several limitations to the present study. First, metabolic pathways catalized by CYP family enzymes are very complex with a large extent of redundancy. While one reaction is mediated by several different enzymes, one enzyme acts on several different pathways, which may result in the synthesis of molecules with opposite biological properties [22]. Although a dominant pathway for each enzyme can be determined by *in vitro* experiments [22], there is uncertainty as to whether such results can be directly extrapolated to the *in vivo* situation, where various enzymes function simultaneously and physiological conditions may vary over time. Moreover, substrates for these enzymes comprise a wide range of chemicals besides estrogen, including some carcinogens to humans [36], and, thus, the net effect of these genetic polymorphisms on breast cancer risk is difficult to predict. In addition, these enzymes are known to be induced by a variety of environmental exposures, such as medications and pesticides, as well as by substrates (steroid hormones) themselves [11,16]. Thus, the contribution of genomic control of overall enzymatic activities may be relatively limited.

A second limitation is that there is a high likelihood that some of the associations observed in this study were chance findings due to multiple comparisons, although the number of genes tested in this study is approximately equal to the number of environmental risk factors typically studied in epidemiological investigations. Thus, caution should be exercised in interpreting the results, taking consistency with other available data into consideration. Although the number of African Americans included in this study was not small compared with those in earlier studies, we realize that our study was not adequately powered to test interactions between variants with low allele frequencies. Our original power calculation for 200 cases and 200 controls indicated that the minimum detectable odds ratios (MDOR) ranged from 1.50 to 1.81 for main effects with minor allele frequencies of 0.1–0.5 and from 2.2 to 2.8 for gene-gene interactions between minor allele frequencies of 0.1–0.5 and a dichotomized genotype frequency of 0.5.

We also realize the potential effect of survival bias as a part of the study (additional sample collection) was based on prevalent cases. In fact, when we analyzed all eligible cases in the SEER database, the cases included in this study found to be 6 years younger than those not included and that cases with distant stage were less (3%) in those included than those who were not (11%). However, the facts that the majority of the cases with a biospecimen were recruited through the rapid case-reporting system in the registry, which limited the loss of the patients before the initial contact, and that only 8 patients were excluded from additional biospecimen collection because of death ensure that the potential effect of this bias was relatively small.

Due to the low participation rate, self-selection bias is certainly a concern. The bias may account for the unexpected association with decaffeinated coffee use as well as cigarette smoking. Thus, we adjusted these factors in calculating multivariable ORs. However, we believe the effect of self-selection bias on genotype distributions of SNPs of interest is relatively limited, compared with that on environmental risk factors. Moreover, although we analyzed SNPs with substantial ethnic variation in allele frequencies in a recently admixed population, studies have demonstrated that the effect of population stratification is very small when a difference in baseline disease risk between ethnicities is small or moderate [37,38]. This is the case for breast cancer and, thus, such biases are unlikely to explain the observed associations.

Despite the limitations discussed above, there was some suggestion that genetic variation in the *CYP1A1*, *CYP1B1* and *HSD17B1* loci may be associated with breast cancer risk in African Americans. Importantly, the association with the *CYP1B1* locus was consistent with functional data and that with the *HSD17B1* locus corroborates previous reports of the association with ER-negative breast cancer in mixed race populations, while our finding of an association with an African American-specific *CYP1A1* polymorphism was relatively novel. Thus, these observations warrant further investigation in larger studies taking into account the full range of variation in these genes as well as other estrogen metabolic genes.

Acknowledgments

This research was supported by grants from Susan G. Komen for the Cure, POP0100608 and POP0504545, and from NIH R01-CA93817.

The authors thank the study participants for their generosity in donating time and biospecimens, the Metropolitan Detroit Cancer Surveillance System for rapid case-ascertainment and Ms. Erin Crosby and Ms. LaTasha Hudgens for recruitment of breast cancer cases and Ms. Deborah Wilson for random digit dialing.

References

1. Newman LA. Breast cancer in African-American women. *Oncologist*. 2005; 10:1–14. [PubMed: 15632248]
2. Bowen RL, Stebbing J, Jones LJ. A review of the ethnic differences in breast cancer. *Pharmacogenomics*. 2006; 7:935–42. [PubMed: 16981851]
3. Ries, LAG.; Melbert, D.; Krapcho, M.; Mariotto, A.; Miller, BA.; Feuer, EJ., et al. SEER Cancer Statistics Review 1975–2004. Bethesda, MD: National Cancer Institute; 2007. http://seer.cancer.gov/csr/1975_2004/, based on November 2006 SEER data submission, posted to the SEER web site
4. Ursin G, Bernstein L, Wang Y, Lord SJ, Deapen D, Liff JM, et al. Reproductive factors and risk of breast carcinoma in a study of white and African-American women. *Cancer*. 2004; 101:353–62. [PubMed: 15241834]
5. Hall IJ, Moorman PG, Millikan RC, Newman B. Comparative analysis of breast cancer risk factors among African-American women and White women. *Am J Epidemiol*. 2005; 161:40–51. [PubMed: 15615914]
6. Bernstein L, Patel AV, Ursin G, Sullivan-Halley J, Press MF, Deapen D, et al. Lifetime recreational exercise activity and breast cancer risk among black women and white women. *J Natl Cancer Inst*. 2005; 97:1671–9. [PubMed: 16288120]
7. Masson LF, Sharp L, Cotton SC, Little J. Cytochrome P-450 1A1 gene polymorphisms and risk of breast cancer: a HuGE review. *Am J Epidemiol*. 2005; 161:901–15. [PubMed: 15870154]
8. Feigelson HS, Coetzee GA, Kolonel LN, Ross RK, Henderson BE. A polymorphism in the CYP17 gene increases the risk of breast cancer. *Cancer Res*. 1997; 57:1063–5. [PubMed: 9067272]
9. Bartsch H, Nair U, Risch A, Rojas M, Wikman H, Alexandrov K. Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol Biomarkers Prev*. 2000; 9:3–28. [PubMed: 10667460]
10. Dally H, Edler L, Jager B, Schmezer P, Spiegelhalder B, Dienemann H, et al. The CYP3A4*1B allele increases risk for small cell lung cancer: effect of gender and smoking dose. *Pharmacogenetics*. 2003; 13:607–18. [PubMed: 14515059]
11. Keshava C, McCanlies EC, Weston A. CYP3A4 polymorphisms--potential risk factors for breast and prostate cancer: a HuGE review. *Am J Epidemiol*. 2004; 160:825–41. [PubMed: 15496535]
12. Olson SH, Bandera EV, Orlov I. Variants in estrogen biosynthesis genes, sex steroid hormone levels, and endometrial cancer: a HuGE review. *Am J Epidemiol*. 2007; 165:235–45. [PubMed: 17110639]
13. Bailey LR, Roodi N, Dupont WD, Parl FF. Association of cytochrome P450 1B1 (CYP1B1) polymorphism with steroid receptor status in breast cancer. *Cancer Res*. 1998; 58:5038–41. [PubMed: 9823305]

14. Sowers MR, Wilson AL, Kardia SR, Chu J, McConnell DS. CYP1A1 and CYP1B1 polymorphisms and their association with estradiol and estrogen metabolites in women who are premenopausal and perimenopausal. *Am J Med.* 2006; 119:S44–S51. [PubMed: 16949388]
15. Peter I, Shearman AM, Vasani RS, Zucker DR, Schmid CH, Demissie S, et al. Association of estrogen receptor beta gene polymorphisms with left ventricular mass and wall thickness in women. *Am J Hypertens.* 2005; 18:1388–95. [PubMed: 16280269]
16. Tsuchiya Y, Nakajima M, Yokoi T. Cytochrome P450-mediated metabolism of estrogens and its regulation in human. *Cancer Lett.* 2005; 227:115–24. [PubMed: 16112414]
17. Ascenzi P, Bocedi A, Marino M. Structure-function relationship of estrogen receptor alpha and beta: impact on human health. *Mol Aspects Med.* 2006; 27:299–402. [PubMed: 16914190]
18. NCBI Single Nucleotide Polymorphism, dbSNP. [accessed August 17, 2007]. <http://www.ncbi.nlm.nih.gov/SNP/>
19. Breslow, NE.; Day, NE. IARC Scientific Publications No 32. Lyon: IARC; 1980. Statistical Methods in Cancer Research, Vol I: The Analysis of Case-Control Studies.
20. Nock NL, Tang D, Rundle A, Neslund-Dudas C, Savera AT, Bock CH, et al. Associations between smoking, polymorphisms in polycyclic aromatic hydrocarbon (PAH) metabolism and conjugation genes and PAH-DNA adducts in prostate tumors differ by race. *Cancer Epidemiol Biomarkers Prev.* 2007; 16:1236–45. [PubMed: 17548691]
21. Martucci CP, Fishman J. P450 enzymes of estrogen metabolism. *Pharmacol Ther.* 1993; 57:237–57. [PubMed: 8361994]
22. Lee AJ, Cai MX, Thomas PE, Conney AH, Zhu BT. Characterization of the oxidative metabolites of 17beta-estradiol and estrone formed by 15 selectively expressed human cytochrome P450 isoforms. *Endocrinology.* 2003; 144:3382–98. [PubMed: 12865317]
23. Belous AR, Hachey DL, Dawling S, Roodi N, Parl FF. Cytochrome P450 1B1-mediated estrogen metabolism results in estrogen-deoxyribonucleoside adduct formation. *Cancer Res.* 2007; 67:812–7. [PubMed: 17234793]
24. Newbold RR, Liehr JG. Induction of uterine adenocarcinoma in CD-1 mice by catechol estrogens. *Cancer Res.* 2000; 60:235–7. [PubMed: 10667565]
25. Li DN, Seidel A, Pritchard MP, Wolf CR, Friedberg T. Polymorphisms in P450 CYP1B1 affect the conversion of estradiol to the potentially carcinogenic metabolite 4-hydroxyestradiol. *Pharmacogenetics.* 2000; 10:343–53. [PubMed: 10862525]
26. Paracchini V, Pedotti P, Raimondi S, Garte S, Bradlow HL, Sepkovic DW, et al. A common CYP1B1 polymorphism is associated with 2-OHE1/16-OHE1 urinary estrone ratio. *Clin Chem Lab Med.* 2005; 43:702–6. [PubMed: 16207128]
27. Paracchini V, Raimondi S, Gram IT, Kang D, Kocabas NA, Kristensen VN, et al. Meta-and pooled analyses of the cytochrome P-450 1B1 Val432Leu polymorphism and breast cancer: a HuGE-GSEC review. *Am J Epidemiol.* 2007; 165:115–25. [PubMed: 17053044]
28. 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. [PubMed: 16103451]
29. Li Y, Millikan RC, Bell DA, Cui L, Tse C-K, Mewman B, et al. Cigarette smoking, cytochrome P4501A1 polymorphisms, and breast cancer among African-American and white women. *Breast Cancer Res.* 2004; 6:R460–R473. [PubMed: 15217514]
30. Taioli E, Garbers S, Bradlow HL, Carmella SG, Akerkar S, Hecht SS. Effects of indole-3-carbinol on the metabolism of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in smokers. *Cancer Epidemiol Biomarkers Prev.* 1997; 6:517–22. [PubMed: 9232339]
31. Bailey LR, Roodi N, Verrier CS, Yee CJ, Dupont WD, Parl FF. Breast cancer and CYP1A1, GSTM1, and GSTT1 polymorphisms: evidence of a lack of association in Caucasians and African Americans. *Cancer Res.* 1998; 58:65–70. [PubMed: 9426059]
32. Rebbeck TR, Troxel AB, Walker AH, Panossian S, Gallagher S, Shatalova EG, et al. Pairwise combinations of estrogen metabolism genotypes in postmenopausal breast cancer etiology. *Cancer Epidemiol Biomarkers Prev.* 2007; 16:444–50. [PubMed: 17372239]

33. Feigelson HS, McKean-Cowdin R, Coetzee GA, Stram DO, Kolonel LN, Henderson BE. Building a multigenic model of breast cancer susceptibility: CYP17 and HSD17B1 are two important candidates. *Cancer Res.* 2001; 61:785–9. [PubMed: 11212283]
34. Wu AH, Seow A, Arakawa K, Van Den Berg D, Lee HP, Yu MC. HSD17B1 and CYP17 polymorphisms and breast cancer risk among Chinese women in Singapore. *Int J Cancer.* 2003; 104:450–7. [PubMed: 12584742]
35. Feigelson HS, Cox DG, Cann HM, Wacholder S, Kaaks R, Henderson BE, et al. Haplotype analysis of the HSD17B1 gene and risk of breast cancer: a comprehensive approach to multicenter analyses of prospective cohort studies. *Cancer Res.* 2006; 66:2468–75. [PubMed: 16489054]
36. Guengerich FP, Shimada T. Oxidation of toxic and carcinogenic chemicals by human cytochrome P-450 enzymes. *Chem Res Toxicol.* 1991; 4:391–407. [PubMed: 1912325]
37. Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiol Biomarkers Prev.* 2002; 11:513–20. [PubMed: 12050091]
38. Wang Y, Localio R, Rebbeck TR. Evaluating bias due to population stratification in epidemiologic studies of gene-gene or gene-environment interactions. *Cancer Epidemiol Biomarkers Prev.* 2006; 15:124–32. [PubMed: 16434597]

Fig. 1.

Representative Taqman plots for the seven SNPs studied. There are 3 clusters, indicating wild type of homozygotes, heterozygotes, and minor allele homozygotes. A: *CYP1A1* rs4986884, B: *CYP1A1* rs4646903, C: *CYP1B1* rs1056836, D: *CYP3A4* rs2740574, E: *CYP17A1* rs743572, F: *HSD17B1* rs605059, and G: *ER Beta* rs1256059. See Table 1 for details of each SNP.

Table 1

Single nucleotide polymorphisms (SNPs) selected for this study

Gene	SNP ID ^a	Location	Alteration	Other allele designations	Minor allele frequency ^b	
					African	Caucasian
<i>CYP1A1</i>	<i>rs4986884</i>	3'UTR	T5639C	3/m3[7,9]	6.2%	0.0%
<i>CYP1A1</i>	<i>rs4646903</i>	3'UTR	T6235C	2A/m1[7,9]	19.0%	28.3%
<i>CYP1B1</i>	<i>rs1056836</i>	Exon 3	val432leu	3/m1[13,14]	35.2%	57.4%
<i>CYP3A4</i>	<i>rs2740574</i>	Promoter	A-392G	1B[10,11]	50.0%	3.3%
<i>CYP17A1</i>	<i>rs743572</i>	Promoter	T-34C	A2[8,12]	33.0%	40.0%
<i>HSD17B1</i>	<i>rs605059</i>	Exon 6	Ser312Gly	--	33.3%	57.1%
<i>ER beta</i>	<i>rs1256059</i>	Intron 7	C/T	--	7.5%	39.7%

^a *rs 4986884* and *rs4646903* have been designated as *rs17861087* and *rs17861083*, respectively, in the Applied Biosystem database.

^b Based on the dataset which includes the largest number of African Americans or Africans (if no data are available for African Americans) along with Caucasians in the NCBI SNP database [18]

Table 2

Univariable and multivariable odds ratios (OR) and 95% confidence intervals (CI) associated with the selected SNPs among all subjects

Genes	Genotypes ^b	No. of cases	No. of controls	Univariable Odds Ratio(95%CI)	Multivariable ^a Odds Ratio(95%CI)
CYP1A1(5639)	TT	157	170	1.00	1.00
	TC	32	17	-	-
	CC	2	0	-	-
	TC+CC	34	17	2.17(1.16 – 4.03)	2.34(1.23 – 4.44)
CYP1A1(6235)	TT	117	109	1.00	1.00
	TC	68	67	0.95(0.62 – 1.45)	0.94(0.61 – 1.45)
	CC	9	13	0.65(0.27 – 1.57)	0.69(0.28 – 1.72)
	Dominant	77	80	0.90(0.60 – 1.35)	0.90(0.59 – 1.36)
	Recessive	9	13	0.66(0.28 – 1.58)	0.71(0.29 – 1.74)
	Dose-effect			P=0.430	P=0.487
	ValVal	106	103	1.00	1.00
CYP1B	ValLeu	68	66	1.00(0.65 – 1.55)	0.97(0.62 – 1.52)
	LeuLeu	8	20	0.39(0.16 – 0.92)	0.33(0.14 – 0.79)
	Dominant	76	86	0.86(0.57 – 1.30)	0.81(0.53 – 1.24)
	Recessive	8	20	0.39(0.17 – 0.91)	0.33(0.14 – 0.79)
	Dose-effect			P=0.132	P=0.067
	GG	82	82	1.00	1.00
	AG	81	84	0.96(0.63 – 1.49)	0.90(0.57 – 1.39)
CYP3A4	AA	31	23	1.35(0.73 – 2.51)	1.27(0.67 – 2.40)
	Dominant	112	107	1.05(0.70 – 1.57)	0.97(0.64 – 1.48)
	Recessive	31	23	1.37(0.77 – 2.45)	1.34(0.74 – 2.44)
	Dose-effect			P=0.490	P=0.689
	TT	70	82	1.00	1.00
	TC	82	78	1.23(0.79 – 1.92)	1.11(0.70 – 1.76)
	CC	32	29	1.29(0.71 – 2.34)	1.12(0.61 – 2.07)
CYP17A1	Dominant	114	107	1.25(0.83 – 1.89)	1.12(0.73 – 1.71)
	Recessive	32	29	1.16(0.67 – 2.01)	1.06(0.60 – 1.87)
	Dose-effect			P=0.320	P=0.651

Genes	Genotypes ^b	No. of cases	No. of controls	Univariable Odds Ratio(95%CI)	Multivariable ^a Odds Ratio(95%CI)
<i>HSD17B1</i>	SerSer	40	57	1.00	1.00
	SerGly	95	86	1.57(0.96 – 2.59)	1.73(1.03 – 2.90)
	GlyGly	42	46	1.30(0.73 – 2.33)	1.34(0.73 – 2.44)
	Dominant	137	132	1.48(0.93 – 2.37)	1.59(0.98 – 2.59)
	Recessive	42	46	0.97(0.60 – 1.56)	0.93(0.57 – 1.53)
<i>ERbeta</i>	Dose-effect			P=0.350	P=0.318
	CC	145	139	1.00	1.00
	TC	44	46	0.92(0.57 – 1.47)	0.84(0.52 – 1.38)
	TT	4	4	0.96(0.24 – 3.91)	0.71(0.17 – 2.98)
	Dominant	48	50	0.92(0.58 – 1.46)	0.83(0.52 – 1.34)
	Recessive	4	4	0.98(0.24 – 3.97)	0.74(0.18 – 3.09)
	Dose-effect			P=0.747	P=0.427

^a Adjusted for pack-years of cigarette smoking and education and use of decaffeinated coffee

^b Dominant: comparison between one or two minor alleles and no alleles, Recessive: comparison between two minor alleles and others

Table 3

Multivariable odds ratios (OR)^a and 95% confidence intervals (CI) associated with the selected SNPs by menopausal status

Genes	Genotypes ^b	Pre-menopausal (N=153)			Post-menopausal (N=230)		
		No. of cases	No. of controls	Odds Ratio(95%CI)	No. of cases	No. of controls	Odds Ratio(95%CI)
<i>CYP11A1</i> (5639)	TT	65	64	1.00	92	106	1.00
	TC	11	9	-	21	8	-
	CC	2	0	-	0	0	-
	TC+ CC	13	9	1.53(0.60 – 3.91)	21	8	3.50(1.41 – 8.74)
<i>CYP11A1</i> (6235)	TT	49	39	1.00	68	70	1.00
	TC	25	26	0.68(0.33 – 1.39)	43	41	1.09(0.62 – 1.93)
	CC	5	9	0.40(0.12 – 1.35)	4	4	0.99(0.23 – 4.35)
	Dominant	30	35	0.61(0.31 – 1.18)	47	45	1.08(0.62 – 1.89)
<i>CYP11B</i>	Recessive	5	9	0.46(0.14 – 1.51)	4	4	0.96(0.22 – 4.12)
	Dose-effect	P=0.097			P=0.822		
	ValVal	44	44	1.00	62	59	1.00
	ValLeu	25	23	1.22(0.59 – 2.52)	43	43	0.88(0.49 – 1.59)
<i>CYP3A4</i>	LeuLeu	4	7	0.55(0.15 – 2.06)	4	13	0.20(0.06 – 0.68)
	Dominant	29	30	1.05(0.53 – 2.06)	47	56	0.70(0.40 – 1.23)
	Recessive	4	7	0.52(0.14 – 1.88)	4	13	0.22(0.07 – 0.70)
	Dose-effect	P=0.745			P=0.034		
<i>CYP17A1</i>	GG	31	35	1.00	51	47	1.00
	AG	33	30	1.13(0.56 – 2.29)	48	54	0.69(0.39 – 1.25)
	AA	15	9	1.62(0.61 – 4.34)	16	14	0.93(0.39 – 2.20)
	Dominant	48	39	1.24(0.64 – 2.40)	64	68	0.74(0.43 – 1.29)
<i>CYP17A1</i>	Recessive	15	9	1.52(0.61 – 3.80)	16	14	1.12(0.50 – 2.51)
	Dose-effect	P=0.362			P=0.534		
	TT	36	31	1.00	34	51	1.00
	TC	27	30	0.84(0.41 – 1.73)	55	48	1.40(0.76 – 2.59)
<i>CYP17A1</i>	CC	12	13	0.75(0.29 – 1.92)	20	16	1.48(0.65 – 3.39)
	Dominant	39	43	0.81(0.42 – 1.56)	75	64	1.42(0.80 – 2.54)
	Recessive	12	13	0.81(0.33 – 1.96)	20	16	1.22(0.58 – 2.59)

Genes	Genotypes ^b	Pre-menopausal (N=153)			Post-menopausal (N=230)		
		No. of cases	No. of controls	Odds Ratio(95%CI)	No. of cases	No. of controls	Odds Ratio(95%CI)
HSD17B1	Dose-effect	P=0.506			P=0.269		
	SerSer	11	24	1.00	29	33	1.00
	SerGly	39	38	2.67(1.11 – 6.43)	56	48	1.42(0.73 – 2.77)
	GlyGly	18	12	4.11(1.41 – 12.03)	24	34	0.76(0.35 – 1.63)
	Dominant	57	50	3.00(1.29 – 6.99)	80	82	1.14(0.61 – 2.11)
	Recessive	18	12	2.04(0.87 – 4.76)	24	34	0.61(0.32 – 1.15)
ERbeta	Dose-effect	P=0.009			P=0.507		
	CC	55	56	1.00	90	83	1.00
	TC	22	17	1.25(0.58 – 2.68)	22	29	0.64(0.33 – 1.24)
	TT	1	1	0.89(0.05 – 15.01)	3	3	0.60(0.11 – 3.16)
	Dominant	23	18	1.23(0.58 – 2.60)	25	32	0.64(0.34 – 1.19)
	Recessive	1	1	0.84(0.05 – 14.04)	3	3	0.67(0.13 – 3.50)
	Dose-effect	P=0.644			P=0.173		

^a Adjusted for pack-years of cigarette smoking and education and use of decaffeinated coffee

^b Dominant: comparison between one or two minor alleles and no alleles, Recessive: comparison between two minor alleles and others

Odds ratios (OR)^a and 95% confidence intervals (CI) associated with *CYP1A1* 5639 C-allele in combination of other SNPs

Table 4

Other SNPs	Genotype	CYP1A1	No. of cases	No. of controls	Odds Ratio (95%CI)	Genotype	No. of cases	No. of controls	Odds Ratio (95%CI)	P-values for interaction
<i>CYP3A4</i>	GG+AG	TT	131	152	1.00	AA	25	18	1.00	0.043
		TC+CC	29	12	3.20 (1.53 – 6.70)		5	5	0.71 (0.16 – 3.11)	
<i>CYP17A1</i>	TT	TT	60	75	1.00	TC+CC	91	95	1.00	0.904
		TC+CC	10	6	2.83 (0.84 – 9.56)		23	11	2.31 (1.05 – 5.10)	
<i>HSD17B1</i>	SerSer	TT	31	54	1.00	SerGly +	113	116	1.00	0.056
		TC+CC	9	1	15.90 (1.82 – 139)	GlyGly	23	16	1.73 (0.85 – 3.54)	
<i>ERbeta</i>	CC	TT	115	124	1.00	TC+TT	41	46	1.00	0.727
		TC+CC	28	13	2.48 (1.20 – 5.11)		6	4	2.13 (0.47 – 9.72)	

^a Adjusted for pack-years of cigarette smoking and education and use of decaffeinated coffee