

Functional Genetic Polymorphisms in the Aromatase Gene *CYP19* Vary the Response of Breast Cancer Patients to Neoadjuvant Therapy with Aromatase Inhibitors

Liewei Wang¹, Katarzyna A. Ellsworth¹, Irene Moon¹, Linda L. Pelleymounter¹, Bruce W. Eckloff², Yvette N. Martin¹, Brooke L. Fridley³, Gregory D. Jenkins³, Anthony Batzler³, Vera J. Suman³, Saranya Ravi¹, J. Michael Dixon^{5,6}, William R. Miller⁵, Eric D. Wieben², Aman Buzdar⁷, Richard M. Weinshilboum¹, and James N. Ingle⁴

Abstract

Aromatase (*CYP19*) is a critical enzyme in estrogen biosynthesis and aromatase inhibitors (AI) are employed widely for endocrine therapy in postmenopausal women with breast cancer. We hypothesized that single nucleotide polymorphisms (SNPs) in the *CYP19* gene may alter the effectiveness of AI therapy in the neoadjuvant setting. Genomic DNA was obtained for sequencing from 52 women pre-AI and post-AI treatment in this setting. Additionally, genomic DNA obtained from 82 samples of breast cancer and 19 samples of normal breast tissue was subjected to resequencing. No differences in *CYP19* sequence were observed between tumor and germ-line DNA in the same patient. A total of 48 SNPs were identified including 4 novel SNPs when compared with previous resequencing data. For genotype-phenotype association studies, we determined the levels of aromatase activity, estrone, estradiol, and tumor size in patients pre-AI and post-AI treatment. We defined two tightly linked SNPs (rs6493497 and rs7176005 in the 5'-flanking region of *CYP19* exon 1.1) that were significantly associated with a greater change in aromatase activity after AI treatment. In a follow-up study of 200 women with early-stage breast cancer who were treated with adjuvant anastrozole, these same two SNPs were also associated with higher plasma estradiol levels in patients pre-AI and post-AI treatment. Electrophoretic mobility shift and reporter gene assays confirmed likely functional effects of these two SNPs on transcription of *CYP19*. Our findings indicate that two common genetic polymorphisms in the aromatase gene *CYP19* vary the response of breast cancer patients to aromatase inhibitors. *Cancer Res*; 70(1); 319–28. ©2010 AACR.

Introduction

In postmenopausal women, the major source of estrogen is the peripheral synthesis of estrone (E_1) and estradiol (E_2) by the cytochrome P450 enzyme aromatase (*CYP19A1*). The third-generation aromatase inhibitors (AI) anastrozole, exemestane, and letrozole are potent and specific inhibitors of aromatase (1, 2) and have become established for the treat-

ment of metastatic disease (3) and as adjuvant therapy in early breast cancer (4). These three AIs have been studied in the neoadjuvant setting by two of the authors (W.R.M. and J.M.D.), and of 59 patients, 46 (78%) had a clinically determined major response (defined as >50% reduction in tumor volume), 8 (14%) had 25% to 50% reduction in tumor volume, and 5 (8%) had <25% reduction (5).

Selected *CYP19* polymorphisms have been investigated for their possible association with the therapeutic efficacy of AIs and with sex hormone levels or risk for estrogen-dependent diseases (6–17). Although the results of those studies are controversial, they suggest an important role for genetic variation in aromatase function and, as a result, in estrogen production. Previously, we performed a *CYP19* resequencing study with 240 germ-line DNA samples from four ethnic groups to identify 88 common single-nucleotide polymorphisms (SNP) and 44 common haplotypes (18). In this study, we investigated whether a difference in *CYP19* sequence between breast tumor and germ-line DNA might exist and whether SNPs in tumor *CYP19* might have a significant effect on response to AI therapy. Specifically, we obtained DNA samples isolated from breast tumor and normal breast tissue from 52 Edinburgh women treated with AIs in a neoadjuvant setting for which *in vivo* assays had been done to measure levels of aromatase activity, E_1 ,

Authors' Affiliations: ¹Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics, ²Department of Biochemistry and Molecular Biology, ³Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, and ⁴Department of Oncology, Mayo Clinic, Rochester, Minnesota; ⁵Breast Unit, Department of Clinical Oncology, Western General Hospital, University of Edinburgh; ⁶Breakthrough Research Unit, Edinburgh, United Kingdom; and ⁷M.D. Anderson Cancer Center Oncology, Houston, Texas

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

Corresponding Author: Liewei Wang, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905. Phone: 507-284-5264; Fax: 507-284-4455; E-mail: wang.liewei@mayo.edu.

doi: 10.1158/0008-5472.CAN-09-3224

©2010 American Association for Cancer Research.

and E_2 in the tumors as well as tumor size before and after AI treatment (5). Gene resequencing was done with DNA from both tumor and normal breast tissue before and after AI treatment. A genotype-phenotype association study was then performed with the resequenced SNPs and all of the *in vivo* phenotypes. Functional genomic studies and a follow-up study were also done to confirm observations made during the initial genotype-phenotype association study.

Materials and Methods

Clinical samples and phenotype data. Both tumor and normal breast tissue samples were obtained from patients as previously described in detail (5). Specifically, women with estrogen receptor-positive (ER^+ ; defined as >20 fmol/mg cytosol protein) primary breast cancer larger than 3 cm (stage $T_{2,3,4b}$, $N_{0,1}$, M_0) and no prior treatment or concurrent treatment with hormone preparations were recruited to a neoadjuvant study to assess levels of E_1 , E_2 , and aromatase activity in the tumor as well as tumor size before and after AI treatment. All patients had received either letrozole (2.5 or 10 mg/d), anastrozole (1 or 10 mg/d), or exemestane (25 mg/d). Patients in the study from whom the specimens were obtained underwent tumor measurement before, at monthly intervals, and at the end of the 3 mo of AI therapy. In addition, patients were given an infusion of [3H]androstenedione and [^{14}C]E $_1$ before and at the end

of the 3-mo treatment period, with blood, tumor, and non-malignant breast tissue taken immediately after the infusion. That original study was designed to determine the effect of neoadjuvant AI therapy on the tumor, whereas the present study uses those data to determine genotype-phenotype associations for *CYP19* polymorphisms.

DNA samples were isolated from 52 patients, including 19 from normal and 82 from tumor tissue. Thirteen were normal-tumor pairs obtained from the same individuals. Tumor DNA included pre- and post-AI treatment samples. After quality control for the gene resequencing results and merging with clinical phenotypes, 77 DNA samples from 45 individuals were available for use in the genotype-phenotype studies. The characteristics of these 45 patients are listed in Supplementary Table S1. This clinical protocol was reviewed and approved by the Lothian Research Ethics Committee (Edinburgh, United Kingdom), and all participants provided written consent. The genotyping study reported subsequently was reviewed and approved by the Mayo Institutional Review Board.

Follow-up study sample sets. Two hundred DNA samples were isolated from women treated with 1 mg anastrozole per day in an ongoing Mayo Clinic-M.D. Anderson collaborative clinical pharmacogenomic study. Plasma samples were collected from these women to measure levels of E_1 and E_2 as well as anastrozole concentrations. E_1 and E_2 levels before and after treatment with anastrozole were measured by gas chromatography-negative ion chemical ionization-tandem

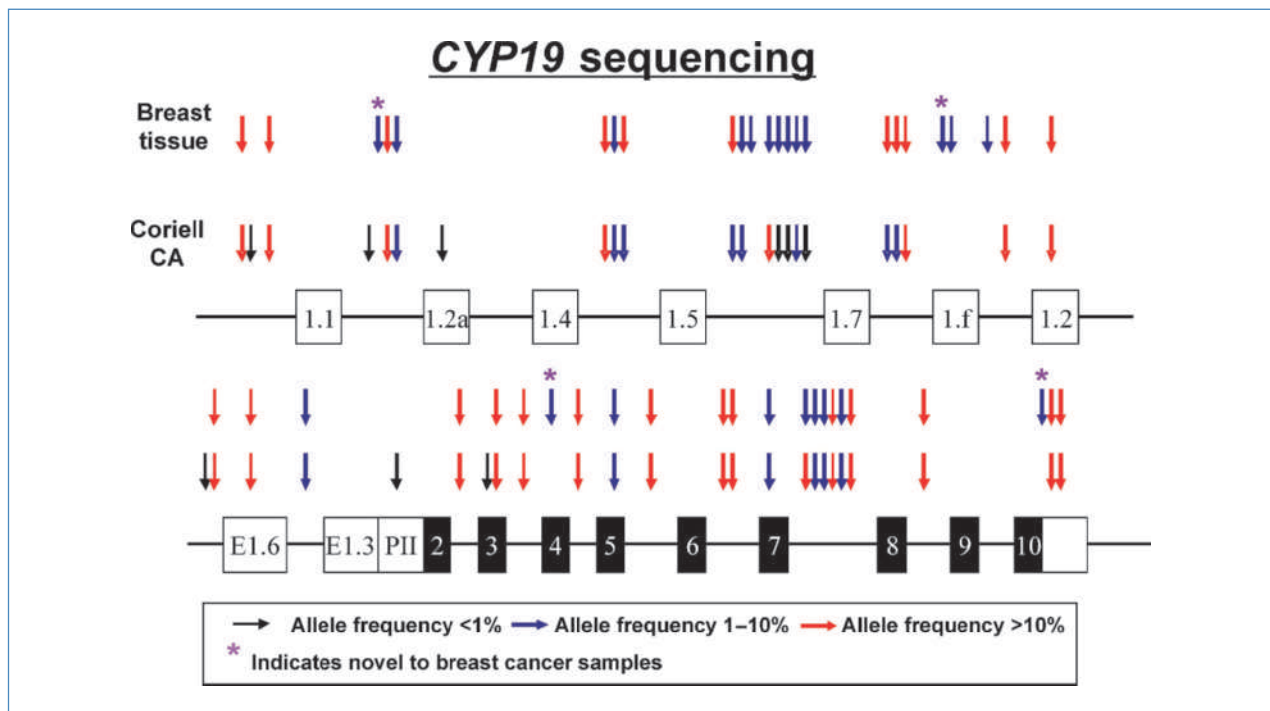
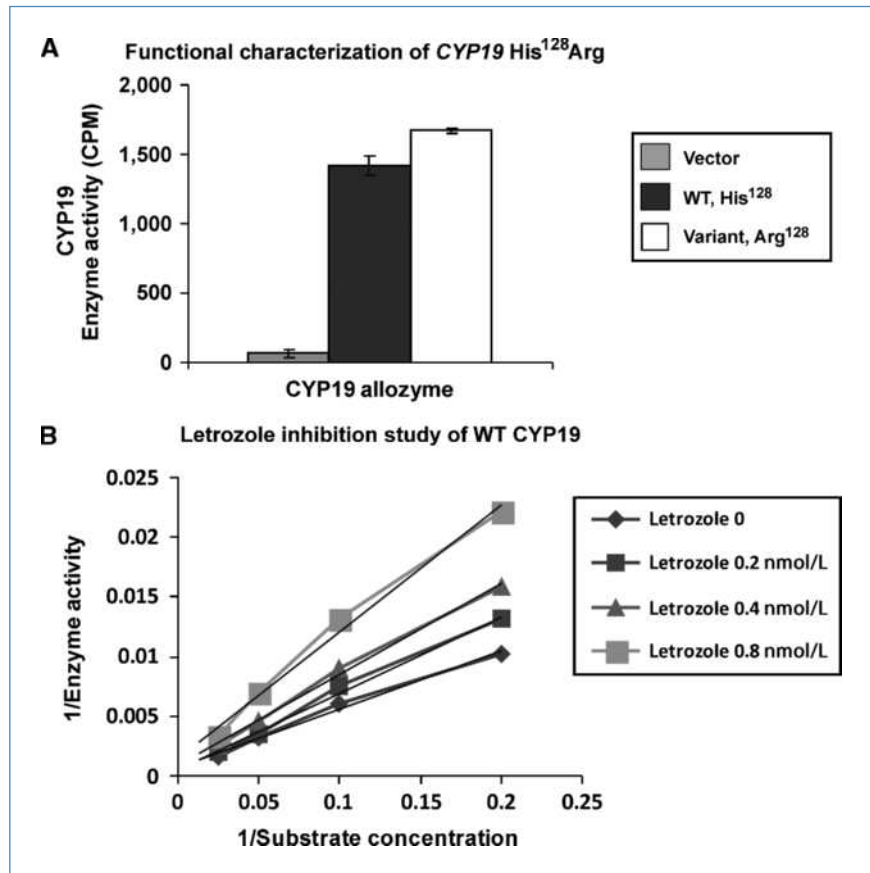


Figure 1. Human *CYP19* gene structure and polymorphisms. Schematic representation of the human *CYP19* gene. Exons are represented as rectangles. Black rectangles, open reading frame; white rectangles, UTRs. Arrows indicate the locations of polymorphisms, with colors representing MAFs. Red arrows, polymorphisms with frequencies of $>10\%$; blue and black arrows, polymorphisms with frequencies of 1% to 10% and $<1\%$, respectively. *, SNPs that were “novel” to this study.

Figure 2. Functional characterization of the CYP19 Arg¹²⁸ variant allozyme. **A**, enzyme activity assay. CYP19 enzyme activity was measured with microsomal preparations from cells transfected with empty vector and WT and Arg¹²⁸ constructs using [¹ β -³H]androst-4-ene-3,17-dione as substrate. Columns, mean of three independent experiments; bars, SE. **B**, representative letrozole inhibition study of WT CYP19. Double-inverse plots were created using six concentrations of [¹ β -³H]androst-4-ene-3,17-dione in the presence of increasing concentrations of letrozole (0–0.8 nmol/L).



mass spectrometry (MS/MS) and liquid chromatography-electrospray ionization-MS/MS bioanalytical methods, conducted at Taylor Technology (Princeton, NJ). Anastrozole and anastrozole metabolite levels were measured by liquid chromatography-MS/MS assay performed at Indiana University. This protocol was reviewed and approved by the Mayo Clinic and M.D. Anderson Institutional Review Boards, and all participants provided written consent.

CYP19 gene resequencing. Resequencing was done using the primers and PCR conditions described in Supplementary Table S2. Resequenced regions included all coding exons, noncoding upstream exons, intron-exon splice junctions, 1,000 bp of 5'-flanking regions for each of the noncoding exons, and the 3'-untranslated region (UTR). The primers and PCR conditions differed slightly from our previous resequencing study using 240 ethnically defined DNA samples (18) in that we sequenced longer regions than in the previous study. Because of this difference, we also resequenced additional regions included in the present study using the same 60 Caucasian DNA samples used in the previous study to allow us to directly compare the resequencing results between germ-line DNA from healthy individuals and tumor DNA from breast cancer patients. Amplicons were sequenced on both strands with an ABI 3730 DNA sequencer. Independent amplifications were done for samples in which a SNP was observed only once

or for samples with ambiguous chromatograms. The chromatograms were analyzed using Mutation Surveyor (SoftGenetics).

Cell culture and transfection. A mammalian expression construct was created for Arg¹²⁸ by performing site-directed mutagenesis using a wild-type (WT) construct described previously (18) as template for circular PCR. Primers used to perform mutagenesis are listed in Supplementary Table S3. COS-1 cells obtained from the American Type Culture Collection (ATCC) were transfected with the WT and variant CYP19 constructs, as well as empty vector as control, using the TranFast reagent (Promega). The pSV40- β -galactosidase vector encoding β -galactosidase (Promega) was used as a control for transfection efficiency.

Aromatase enzyme assay and inhibition study. Aromatase activity was determined as described previously (18–20). Specifically, microsomes were isolated from COS-1 cells transfected with WT, Arg¹²⁸, and empty vector together with β -galactosidase. Microsomal preparations were then used to perform the aromatase enzyme assay using [¹ β -³H]androst-4-ene-3,17-dione (NEN Life Science Products) as substrate (18). Aromatase activity was assayed by measuring the release of ³H₂O from radiolabeled substrate. Reactions were carried out for 45 min at 37°C with a NADPH regeneration system in a final volume of 100 μ L, as described previously. Blanks were samples that lacked

Table 1. Summary of the clinical and *in vivo* phenotypes of the breast cancer patients

Phenotype	Pretreatment	Posttreatment	Post-pre
Tumor volume			
<i>n</i>	44	44	44
Mean (SD)	8.5 (11.00)	2.2 (2.38)	-6.4 (9.08)
Aromatase activity (fmol/mg protein/h)			
<i>n</i>	45	45	45
Mean (SD)	33.9 (41.80)	6.1 (13.16)	-27.8 (40.71)
Median	23.0	0.0	-21.0
E ₁ level (dpm/g)			
<i>n</i>	30	30	30
Mean (SD)	1.9 (2.79)	0.4 (0.40)	-1.5 (2.73)
Median	0.7	0.3	-0.5
E ₂ level (dpm/g)			
<i>n</i>	30	30	30
Mean (SD)	1.5 (1.27)	0.3 (0.34)	-1.2 (1.25)
Median	1.2	0.3	-0.8

NOTE: These phenotypes were determined as part of a previous study (5).

enzyme. For substrate kinetic studies, substrate concentrations ranged from 0 to 80 nmol/L. For inhibition studies, assays were done in the presence of three concentrations of letrozole (0.2, 0.4, and 0.8 nmol/L). K_m and K_i values were then determined.

Electrophoretic mobility shift assay. Nuclear extracts from SKBR03 cells (ATCC) were isolated and protein was quantified by the Bradford method (21). Biotin-labeled sense and antisense oligonucleotides were designed to contain WT and variant sequences for the rs6493497 (G/A) and rs7176005 (C/T) SNPs. Probe sequences are listed in Supplementary Table S3. Electrophoretic mobility shift assays (EMSA) were done with the LightShift Chemiluminescent EMSA kit (Pierce). A 400-fold excess of unlabeled probe was added for the competition assays.

Reporter gene assay. Reporter gene constructs were created by PCR amplification of a 1,000-bp region surrounding the two SNPs using Coriell DNA as template. The WT PCR product was cloned into pGL3 Basic vector (Promega), followed by mutagenesis to obtain the variant constructs. Primers used to perform mutagenesis are listed in Supplementary Table S3. DNA sequences were verified by sequencing both strands. Vector without an insert was used as a control. SKBR03 cells were transfected with WT and variant constructs together with a pRL-TK DNA construct encoding *Renilla* luciferase as a control for transfection efficiency. Cells were then lysed and luciferase activity was measured using a dual-luciferase activity assay with a TD-20/20 luminometer (Turner Designs). Results were reported as the ratio of firefly luciferase to *Renilla* luciferase light units, and all values were expressed as a percentage of the activity of the WT construct. All assays were done in triplicate.

TaqMan assay. In a follow-up study, SNPs rs6493497 (G/A) and rs7176005 (C/T), upstream of *CYP19* exon 1.1, were gen-

typed using 200 germ-line DNA samples from breast cancer patients being treated with adjuvant anastrozole. TaqMan assays were done according to the manufacturer's instructions. Positive (Coriell DNA with known genotypes) and negative controls (lacking DNA) were included in the assays.

Statistical methods. SNPs obtained during the gene resequencing study were subjected to quality control to remove those with minor allele frequencies (MAF) <5%, SNPs that deviated significantly from Hardy-Weinberg equilibrium ($P < 0.001$), or SNPs with call rates <95%. Due to the skewness of the phenotypes, the van der Waerden rank normal score transformation was done for all phenotypes. Phenotypes were compared among anastrozole, exemestane, and letrozole using a Monte Carlo approximated exact Kruskal-Wallis test. SNP associations with transformed phenotypes were assessed using ordinary regression, treating genotype effect as the count of minor alleles [i.e., 1 degree of freedom (*df*) or trend test] in all treatment groups as well in a subset analysis by treatment group. For the genotypic analyses, *P* values were corrected for multiple testing by a permutation step-up procedure (22). Linkage disequilibrium among *CYP19* polymorphisms was determined by calculating D' and R^2 values for all possible pairwise combinations of polymorphisms, and the data were plotted using Haploview software (23). Haplotypes were estimated for the markers passing quality control and compared among phenotypes using score tests from the haplo.stats software⁸ (24). For the follow-up study, the relationship between genotypes and pre-drug and post-drug E₁ and E₂ concentrations was modeled using a negative bimodal generalized linear model due to the extreme skewness of the phenotypes (25, 26). Models of posttreatment

⁸ http://mayoresearch.mayo.edu/mayo/research/schaid_lab/software.cfm

E_1 and E_2 were also adjusted for anastrozole levels. Differences in E_1 and E_2 , before or after treatment, were left on their original scale. Ordinary regression was used to assess the association of phenotypes and genotypes treated as the count of minor alleles (i.e., 1 *df*).

Differences in enzyme activity, protein levels, and luciferase activities were tested using Student's *t* test with the software Prism. K_i and K_m values were calculated with the Prism program.

Results

CYP19 gene resequencing. *CYP19* resequencing identified 48 SNPs (Fig. 1; Supplementary Table S4). No differences were observed between normal germ-line and tumor DNA. Because one reaction covering part of the 5'-flanking region of exon 1.7 contained two indels, one located 1,051 bp and the other 1,000 bp upstream of exon 1.7, we could not deter-

mine the allele frequencies of the indels or of a SNP at -787 upstream of exon 1.7. We also could not determine the genotype of an indel and variable number of tandem repeats in intron 4 that were reported previously. While using the 60 CA DNA samples to resequence the additional regions that were not covered in our previous study (18), we identified 11 additional SNPs in Caucasians in addition to the original 37 SNPs identified during our previous study (Fig. 1; Supplementary Table S4). Four SNPs, including A383G that resulted in a 128His>Arg amino acid alteration, were observed only in breast tissue, including both normal and tumor tissues (Fig. 1; Supplementary Table S4).

Functional characterization of His¹²⁸Arg. Because His¹²⁸Arg was only observed in breast tumor and normal breast germ-line DNA at 1% allele frequency and had not been observed in the 60 Caucasian subjects included in our previous resequencing study, we performed functional genomic studies with this SNP. Microsomal preparations from COS-1 cells transiently transfected with WT, Arg¹²⁸, and

Table 2. Genotype-phenotype association

(A) Results of genotype-phenotype correlation study for rs7176005 and rs6493497

Phenotype	Marker	Pearson <i>r</i>	Unadjusted <i>P</i>	Adjusted <i>P</i>
Change in aromatase activity	5'-FR of exon 1.2 (-588) rs7176005	-0.40	0.006	0.039
	5'-FR of exon 1.1 (-144) rs6493497	-0.40	0.006	0.039
Change in E_1	5'-FR of exon 1.2 (-588) rs7176005	-0.38	0.038	0.180
	5'-FR of exon 1.1 (-144) rs6493497	-0.38	0.038	0.180
Pretreatment aromatase activity	5'-FR of exon 1.2 (-588) rs7176005	0.33	0.028	0.151
	5'-FR of exon 1.1 (-144) rs6493497	0.33	0.028	0.151

(B) Effect of the two SNPs on plasma estrogen levels before and after anastrozole treatment in 200 postmenopausal patients recruited for the Mayo-MD Anderson pharmacogenomics study

Phenotype	Marker	Multiplicative effect of 1 rare allele increase in phenotype (95% CI)	<i>P</i> for marker effect	<i>P</i> for anastrozole effect
Pre- E_1	rs6493497	1.20 (1.00-1.45)	0.055	—
	rs7176005	1.14 (0.95-1.36)	0.15	—
Pre- E_2	rs6493497	1.31 (1.03-1.67)	0.028	—
	rs7176005	1.24 (0.99-1.56)	0.057	—
Post- E_1 *	rs6493497	3.38 (0.79-14.36)	0.099	0.63
	rs7176005	3.02 (0.78-11.68)	0.110	0.63
Post- E_2 *	rs6493497	12.58 (3.39-46.68)	0.0002	0.93
	rs7176005	10.88 (3.12-37.94)	0.0002	0.80

Abbreviations: FR, flanking region; 95% CI, 95% confidence interval.

*Adjusted for anastrozole plasma level.

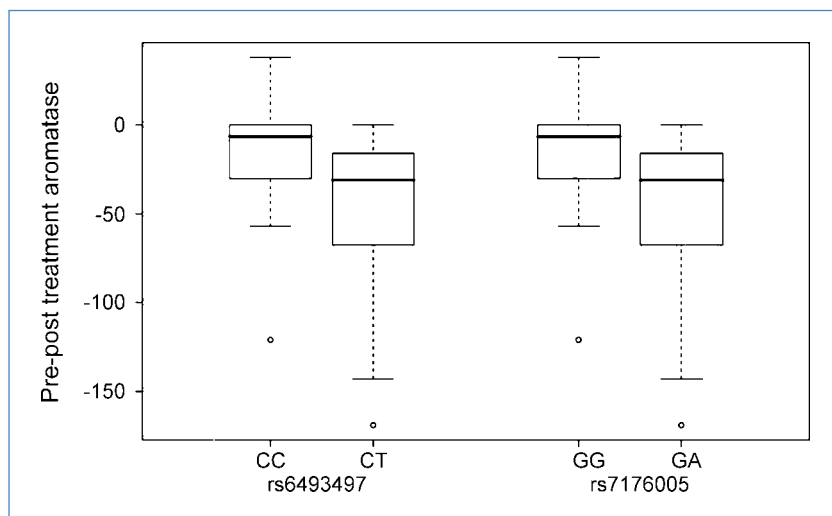


Figure 3. Genotype-phenotype association study. Quartile box plot was plotted between the genotypes of the two SNPs, rs6493497 (-588) and rs7176005 (-144), and change in aromatase activity measured in tumor tissue before and after AI treatment.

empty vector were used to perform enzyme assays with [1β - ^3H]androst-4-ene-3,17-dione as substrate. No significant difference in aromatase enzyme activity was observed between the WT and Arg¹²⁸ allozymes (Fig. 2A). Substrate kinetic studies were also done with six different concentrations of [1β - ^3H]androst-4-ene-3,17-dione (2.5, 5, 10, 20, 40, and 80 nmol/L). Representative substrate kinetic studies for WT and Arg¹²⁸ showed that apparent K_m values for the WT and Arg¹²⁸ allozymes were 81 ± 4 and 164 ± 32 nmol/L, respectively. These values did not differ significantly ($P = 0.08$). We also performed inhibition studies with increasing concentrations of letrozole (0.2, 0.4, and 0.8 nmol/L). Double-inverse plots for the WT inhibition study with letrozole are shown in Fig. 2B. The K_i value for WT was 1.6 ± 0.11 nmol/L and that for Arg¹²⁸ was 1.6 ± 0.27 nmol/L ($P = 0.9$). Western blot analysis using antibody directed against aromatase showed no difference in expression (data not shown).

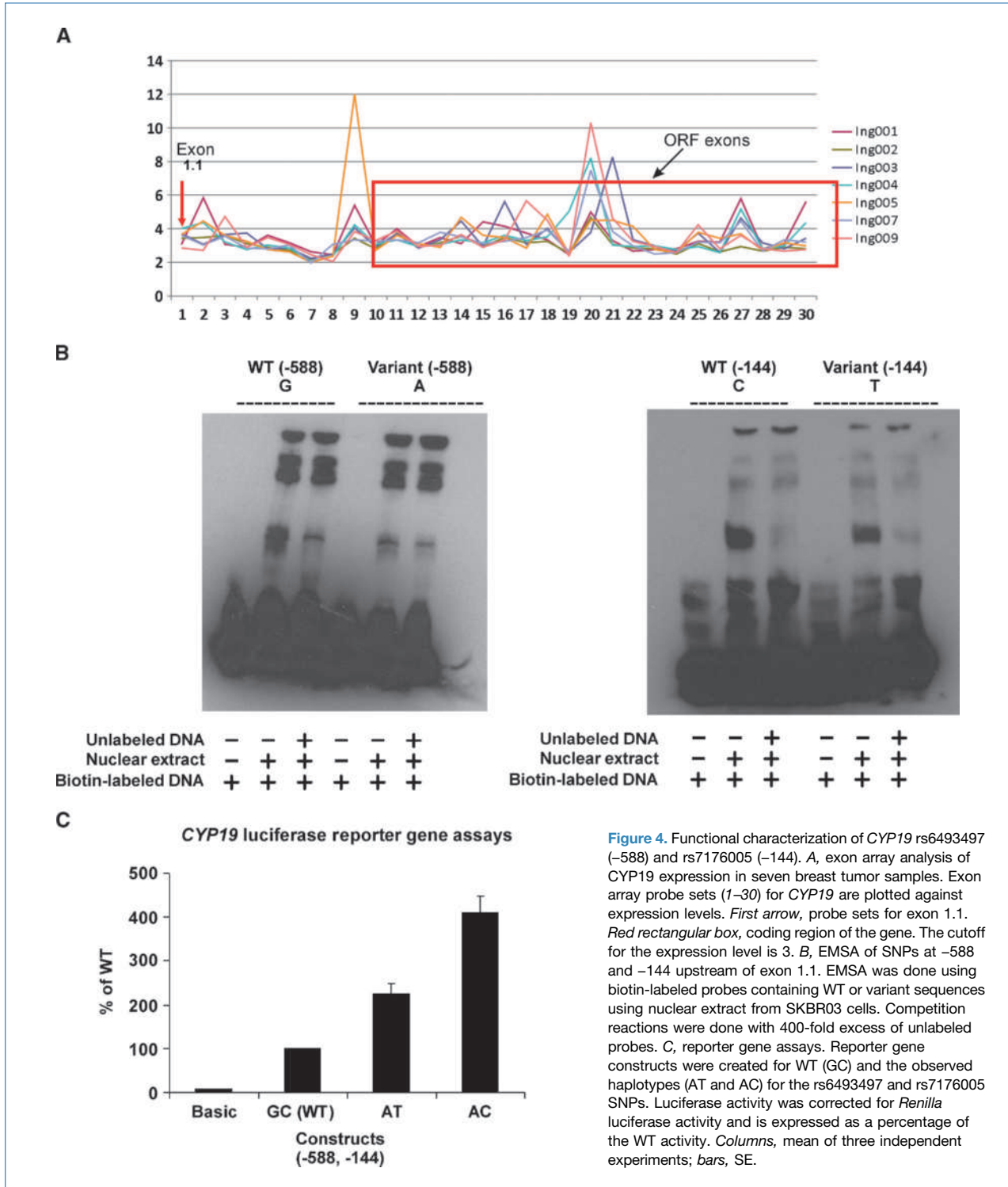
Genotype-phenotype association study. To identify *CYP19* SNPs in breast tumor tissue that might contribute to variation in response to AI treatment in women with ER⁺ breast cancer, we performed genotype-phenotype studies with our resequencing data using tumor DNA together with *in vivo* phenotypes that had been obtained in a neoadjuvant setting, as described previously (5). DNA samples from 45 patients were used in this analysis. Supplementary Table S1 lists the demographic and clinical characteristics of these patients. These patients were treated with three different AIs, with information on aromatase activity before and after AI treatment recorded for 45 patients and on tumor size before and after AI treatment for 44 patients. However, only the 30 patients treated with letrozole and anastrozole had information on levels of E_1 and E_2 before and after treatment because exemestane is a steroid analogue and interferes with the assays used to measure E_1 and E_2 (27). As expected, most patients treated with AIs had decreases in E_1 , E_2 , and tumor size. However, there was variation in changes in these phenotypes before and after AI treatment (Supplementary Fig. S1). Before performing the genotype-phenotype association

study, we determined the effect of the three treatments on the phenotypes of interest. No significant differences for phenotypes were observed among the drugs. Because the effects of the different AI drugs on the phenotypes were minimal, statistical analysis was completed on the combined data for all three AIs.

Genotype-phenotype association analysis was then done using the SNPs identified by resequencing *CYP19* in breast tumor DNA samples from 45 patients. After quality control, 27 SNPs were used in the analysis. Table 1 lists the phenotypes studied, including changes in E_1 , E_2 , aromatase activity, and tumor size as well as each of these phenotypes measured before and after AI treatment. Two SNPs, one located at 144 bp (rs6493497) and the other located at 588 bp (rs7176005) 5' upstream of exon 1.1, the major "placental" exon 1, were significantly associated with change in aromatase activity before and after AI treatment (nominal $P = 0.0063$, $P = 0.039$ after adjustment for multiple testing; Table 2A; Fig. 3). Specifically, these SNPs result in greater change in aromatase activity before and after AI treatment, which indicated that samples containing these two SNPs had greater inhibition of aromatase activity. These same SNPs were also associated with change in E_1 level ($P = 0.038$) and with higher aromatase activity at baseline ($P = 0.028$), although these differences were not significant after correction for multiple comparisons (Table 2A). rs6493497 and rs7176005 were in linkage disequilibrium ($R^2 = 1$), with observed MAFs of 17% in these samples (Supplementary Fig. S2). Another two linked SNPs ($R^2 = 1$), one located (-690) and the other (-725) bp upstream of exon 1.f, were significantly associated with post-drug aromatase activity, with $P = 0.022$ adjusted for multiple testing. Marginally, the SNP at (-628) upstream of the 5'-flanking region for exon 1.5 was also associated with change in aromatase ($P = 0.06$ after adjustment for multiple testing). We also performed haplotype analysis using the resequenced SNPs, but none of the haplotypes showed a significant association with the phenotypes of interest after correction for multiple testing (data not shown).

Because rs6493497 and rs7176005 were associated with multiple phenotypes that were related to aromatase activity and aromatase inhibition, we also performed a follow-up study by genotyping these two SNPs using an additional 200 DNA samples isolated from postmenopausal women

with breast cancer who were treated with 1.0 mg anastrozole per day in the adjuvant setting. Plasma samples had been collected before and after the initiation of anastrozole treatment for the measurement of E₁ and E₂ levels as well as plasma anastrozole concentrations, although tissue aromatase



enzyme activity was not available for these patients. The MAFs for rs6493497 and rs7176005 were 12% and 13%, respectively, in these 200 samples. The genotype-phenotype association study for these 200 subjects was done using levels of E_1 , E_2 , and change in E_1 and E_2 before and after anastrozole as phenotypes. Both SNPs were significantly associated with pre-drug ($P = 0.028$ for rs6493497 and $P = 0.057$ for rs7176005) and post-drug ($P = 0.0002$ for both SNPs) E_2 concentrations after adjustment for plasma anastrozole concentration (Table 2B). The average plasma concentration of anastrozole in these 200 subjects was 32.2 ng/mL, with a range from 0.0 to 98.8 ng/mL. Specifically, with each added minor allele, rs6493497 and rs7176005 were associated with 13- and 11-fold increases in post-drug E_2 level, respectively. We will discuss the possible implications of the genotype-phenotype association results in the Discussion.

Functional characterization of rs6493497 and rs7176005.

To further understand the functional effect of the two significant SNPs located in the 5'-flanking region of *CYP19* exon 1.1, we performed a series of functional genomic studies. These SNPs are located in the major placental promoter, exon 1.1, and thus we first attempted to determine whether this exon was expressed in breast tumor tissue using exon array data obtained with seven RNA samples isolated from breast tumor tissue, a subset of the tumor samples included in this study. Figure 4A shows these exon array data and indicates that although the expression of this exon in breast tumor tissue was low, as shown by the first arrow, it was clearly expressed. We then performed EMSA to test for possible DNA-protein binding within the regions containing these two SNPs using nuclear extract from the SKBR03 cells. SKBR03 cells have WT sequence for both of these SNPs (Fig. 4B). We chose this cell line because a previous study done by Kinoshita and Chen (28) showed that aromatase is highly expressed in these cells and that estrogen can upregulate aromatase expression through *CYP19* promoter 1.1. EMSA showed that the WT sequence at position -588 displayed greater DNA-protein binding than did the variant sequence. This effect was less striking for the SNP at position -144, although binding for the WT sequence was slightly stronger than that for the variant (Fig. 4B). Reporter gene assay indicated that the AT and AC combinations both displayed increased transcription activity, with AC having the most significant effect (Fig. 4C), that is, the effect on transcription was most striking for the -144 SNP. The TC combination had not been observed in any of the patients. Therefore, we did not study that haplotype.

Discussion

Breast cancer remains the most common cancer among women, and the majority of postmenopausal women have tumors that are ER⁺ and, as a result, are candidates for endocrine therapy (29). Over the past decade, AIs have become established in both the metastatic and the adjuvant settings and are a main focus of research in the prevention setting in postmenopausal women at high risk of developing breast cancer (3, 4, 30). Aromatase (CYP19) catalyzes

the conversion of androgens to E_1 or E_2 (31). Therefore, inhibition of aromatase has become a mainstream treatment of postmenopausal women with ER⁺ breast cancer. Third-generation AIs include the steroid derivative exemestane and the nonsteroidal agents letrozole and anastrozole. However, response to treatment with these drugs varies widely, ranging from lack of efficacy to severe side effects (32, 33). One possible explanation is that genetic variation in the genes encoding the drug target, or drug-metabolizing enzymes, might contribute to this variation in response. Previous resequencing studies of *CYP19* identified 88 SNPs using 240 ethnically defined germ-line DNA samples (18). In the current study, we expanded our previous resequencing effort to include DNA isolated from breast tumor and normal tissue to test the hypothesis that genetic variation in *CYP19* might play an important role in variations in aromatase enzyme activity, estrogen production, and clinical response. Our study took advantage of a unique sample set collected at the University of Edinburgh (5). These samples were collected in a neoadjuvant setting, with information on *in vivo* E_1 and E_2 levels, aromatase activity in the tumor, and tumor size before and after AI treatment.

Resequencing of *CYP19* in these samples identified a total of 48 SNPs, including 4 novel SNPs, as compared with previous resequencing efforts (Fig. 1; Supplementary Table S4). Functional characterization of a novel nonsynonymous coding SNP, Arg¹²⁸, failed to show an effect on aromatase enzyme activity, apparent K_m value, or K_i values with letrozole (Fig. 2). However, genotype-phenotype association studies done using 27 SNPs identified during resequencing with *in vivo* and clinical data for 45 patients as phenotypes identified two tightly linked SNPs, rs6493497 and rs7176005, which were significantly associated with a greater change in aromatase activity before and after AI treatment (Table 2A; Fig. 3), with adjusted P values of 0.038 for both SNPs. These same SNPs were also moderately associated with change in E_1 level and with baseline aromatase activity in these tumor samples. These two SNPs were associated with higher basal aromatase activity, compatible with the results of our luciferase reporter gene assay, which showed high activity for the variant SNPs, as well as greater inhibitory effect of AIs, which might be due to the elevated baseline level of aromatase activity (Figs. 3 and 4B and C). However, these two SNPs were not associated with levels of either E_1 or E_2 before or after AI treatment in tumor tissue, which could be due to the small sample size because E_1 and E_2 levels were only available for the 30 samples from patients treated with letrozole or anastrozole.

We next performed a follow-up study with an additional 200 DNA samples obtained from women with breast cancer who had been treated with 1.0 mg anastrozole. We had measured plasma E_1 and E_2 concentrations as well as anastrozole levels in these patients before and after AI treatment. These phenotypes were used to perform a genotype-phenotype correlation study to determine whether the two SNPs identified in the Edinburgh patients might be associated with plasma estrogen concentrations. Although the

samples in the follow-up study were germ-line DNA, we hypothesized that if the SNPs had functional effects on aromatase activity and inhibition by AIs, they might have similar effects on plasma estrogen levels. We performed the follow-up study with the two SNPs independently to avoid multiple comparisons because they were tightly linked. The results showed that both SNPs were significantly associated with plasma E_2 level before and after anastrozole treatment (Table 2B), adjusted for plasma drug concentrations, a potential confounding variable. We observed this association only with E_2 and not with E_1 , which might be due to different precursors for E_1 and E_2 , with E_1 being formed from androstenedione and E_2 from testosterone. Both SNPs were associated with higher E_2 levels, particularly after treatment with anastrozole (Table 2B), which could be related to our observations with the tumor DNA in which the two SNPs were moderately associated with higher aromatase activity at baseline. Although inhibition was greater with samples harboring these SNPs in tumor tissue (Fig. 3), this might not result in lower estrogen levels after AI treatment because of the higher basal aromatase activity for samples with these two SNPs.

We next determined how these two SNPs might affect aromatase function. We performed reporter gene and EMSA assays, which showed that the two SNPs showed different DNA-protein binding patterns for WT and variant sequences, particularly for the -588 SNP, and that they displayed different transcriptional activity when compared with the WT sequence (Fig. 4B and C). These functional studies provided additional evidence that these SNPs might result in increased aromatase transcription and, thus, higher aromatase activity, resulting in higher estrogen production, although inhibition of aromatase activity was greater in the presence of the SNPs, which might be due to the higher aromatase expression (more drug targets for AIs). It would be interesting to determine whether tumor sam-

ples containing these SNPs have different CYP19 expression than WT. However, because the quantity of tumor tissue was very limited, we were unable to test this hypothesis. This possibility should be tested in future studies. We did not observe an association with tumor size or change in tumor size before and after AIs, nor did we observe an association with E_1 or E_2 before and after AI treatment in tumors, which might be due to the small sample size studied. However, the fact that rs6493497 and rs7176005 in the 5'-flanking region of exon 1.1 of *CYP19* were associated with aromatase activity and also displayed an association with plasma E_2 levels before and after AI treatment suggests an important functional role for these two SNPs in variation in response to AI therapy. It also raises the possibility that these same SNPs might play a role in the risk of breast cancer because estrogen is a major risk factor for breast cancer. All of these possibilities need to be tested further in larger independent patient cohorts.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

NIH grants U01 GM61388 (The Pharmacogenetics Research Network), P50 CA116201 (Mayo Clinic Breast Cancer Specialized Program of Research Excellence), and R01 CA138461 and PhRMA Foundation "Center of Excellence in Clinical Pharmacology" Award.

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.

Received 8/28/09; revised 10/20/09; accepted 10/21/09; published online 1/4/10.

References

- Geisler J, Haynes B, Anker G, Dowsett M, Lonning PE. Influence of letrozole and anastrozole on total body aromatization and plasma estrogen levels in postmenopausal breast cancer patients evaluated in a randomized, cross-over study. *J Clin Oncol* 2002;20:751-7.
- Geisler J, King N, Anker G, et al. *In vivo* inhibition of aromatization by exemestane, a novel irreversible aromatase inhibitor, in postmenopausal breast cancer patients. *Clin Cancer Res* 1998;4:2089-93.
- Ingle JN, Suman VJ. Aromatase inhibitors for therapy of advanced breast cancer. *J Steroid Biochem Mol Biol* 2005;95:113-9.
- Ingle JN, Dowsett M, Cuzick J, Davies C. Aromatase inhibitors versus tamoxifen as adjuvant therapy for postmenopausal women with estrogen receptor positive breast cancer: meta-analyses of randomized trials of monotherapy and switching strategies. *Cancer Res* 2009;69:66S.
- Miller WR, Dixon JM. Local endocrine effects of aromatase inhibitors within the breast. *J Steroid Biochem Mol Biol* 2001;79:93-102.
- Watanabe J, Harada N, Suemasu K, Higashi Y, Gotoh O, Kawajiri K. Arginine-cysteine polymorphism at codon 264 of the human CYP19 gene does not affect aromatase activity. *Pharmacogenetics* 1997;7:419-24.
- Kristensen VN, Harada N, Yoshimura N, et al. Genetic variants of CYP19 (aromatase) and breast cancer risk. *Oncogene* 2000;19:1329-33.
- Probst-Hensch NM, Ingles SA, Diep AT, et al. Aromatase and breast cancer susceptibility. *Endocr Relat Cancer* 1999;6:165-73.
- Siegelmann-Danieli N, Buetow KH. Constitutional genetic variation at the human aromatase gene (*Cyp19*) and breast cancer risk. *Br J Cancer* 1999;79:456-63.
- Haiman CA, Hankinson SE, Spiegelman D, et al. A tetranucleotide repeat polymorphism in CYP19 and breast cancer risk. *Int J Cancer* 2000;87:204-10.
- Bershtein LM, Imianitov EN, Suspitsyn EN, et al. A polymorphism study of the CYP19 gene in endometrial cancer patients. *Vopr Onkol* 2000;46:302-5.
- Baxter SW, Choong DY, Eccles DM, Campbell IG. Polymorphic variation in CYP19 and the risk of breast cancer. *Carcinogenesis* 2001;22:347-9.
- Masi L, Becherini L, Gennari L, et al. Polymorphism of the aromatase gene in postmenopausal Italian women: distribution and correlation with bone mass and fracture risk. *J Clin Endocrinol Metab* 2001;86:2263-9.

14. Miyoshi Y, Ando A, Hasegawa S, et al. Association of genetic polymorphisms in CYP19 and CYP1A1 with the oestrogen receptor-positive breast cancer risk. *Eur J Cancer* 2003;39:2531–7.
15. Thyagarajan B, Brott M, Mink P, et al. CYP1B1 and CYP19 gene polymorphisms and breast cancer incidence: no association in the ARIC study. *Cancer Lett* 2004;207:183–9.
16. 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.
17. Lloveras B, Monzo M, Colomer R, et al. Letrozole efficacy is related to human aromatase CYP19 single nucleotide polymorphisms (SNPs) in metastatic breast cancer patients. *Am Soc Clin Oncol Annu Meet Proc (Post-Meet Ed)* 2004;22:507.
18. Ma CX, Adjei AA, Salavaggione OE, et al. Human aromatase: gene resequencing and functional genomics. *Cancer Res* 2005;65:11071–82.
19. Brueggemeier RW, Richards JA, Joomprabutra S, Bhat AS, Whetstone JL. Molecular pharmacology of aromatase and its regulation by endogenous and exogenous agents. *J Steroid Biochem Mol Biol* 2001;79:75–84.
20. Hahn EF, Fishman J. Immunological probe of estrogen biosynthesis. Evidence for the 2 β -hydroxylative pathway in aromatization of androgens. *J Biol Chem* 1984;259:1689–94.
21. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal Biochem* 1976;72:248–54.
22. Peter H, Westfall RDT, Rom D, Wolfinger RD, Hochberg Y. Multiple comparisons and multiple tests: using the SAS system. Cary (NC): SAS Institute; 2000.
23. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–5.
24. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002;70:425–34.
25. Agresti A. *Categorical data analysis*. Hoboken (NJ): John Wiley & Sons, Inc.; 2002.
26. Wedderburn RWM. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* 1974;61:439–47.
27. Johannessen DC, Engan T, Di Salle E, et al. Endocrine and clinical effects of exemestane (PNU 155971), a novel steroidal aromatase inhibitor, in postmenopausal breast cancer patients: a phase I study. *Clin Cancer Res* 1997;3:1101–8.
28. Kinoshita Y, Chen S. Induction of aromatase (CYP19) expression in breast cancer cells through a nongenomic action of estrogen receptor α . *Cancer Res* 2003;63:3546–55.
29. Anderson WF, Chatterjee N, Ershler WB, Brawley OW. Estrogen receptor breast cancer phenotypes in the Surveillance, Epidemiology, and End Results database. *Breast Cancer Res Treat* 2002;76:27–36.
30. Ingle JN. Endocrine therapy trials of aromatase inhibitors for breast cancer in the adjuvant and prevention settings. *Clin Cancer Res* 2005;11:900–5.
31. Simpson ER, Mahendroo MS, Means GD, et al. Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis. *Endocr Rev* 1994;15:342–55.
32. Forbes JF, Cuzick J, Buzdar A, Howell A, Tobias JS, Baum M. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 100-month analysis of the ATAC trial. *Lancet Oncol* 2008;9:45–53.
33. Coates AS, Keshaviah A, Thurlimann B, et al. Five years of letrozole compared with tamoxifen as initial adjuvant therapy for postmenopausal women with endocrine-responsive early breast cancer: update of study BIG 1-98. *J Clin Oncol* 2007;25:486–92.