Genetic Polymorphisms of the *CYP19A1* Gene and Breast Cancer Survival

Ji-Rong Long,[1] Nobuhiko Kataoka,[1] Xiao-Ou Shu,[1] Wanqing Wen,[1] Yu-Tang Gao,[2] Qiuyin Cai,[1] and Wei Zheng[1]

[1]Department of Medicine, Center for Epidemiologic Research, Vanderbilt University School of Medicine, Nashville, Tennessee and [2]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China

Abstract

The CYP19A1 protein (aromatase) plays a critical role in estrogen biosynthesis and thus may be related to the progression of breast cancer. We examined the association between *CYP19A1* genetic polymorphisms and breast cancer survival in a cohort of 1,136 patients who were recruited as part of a population-based case-control study in Shanghai, China from 1996 to 1998 and who has donated a DNA sample to the study. Patients were followed for cancer recurrence and mortality through July 2005. Nineteen haplotype tagging single-nucleotide polymorphisms (SNP) in the *CYP19A1* gene were evaluated. For each of the five SNPs located in haplotype block 2, patients homozygous for the minor alleles had a reduced 5-year disease-free survival rate compared with those carrying the major allele. The age-adjusted hazard ratios (HR) and 95% confidence intervals (95% CI) were 1.5 (1.1-2.1), 2.1 (1.2-3.6), 1.5 (1.1-2.0), 1.4 (1.0-2.0), and 1.4 (1.0-2.0) for *hCV1664178, rs12900137, rs730154, rs936306*, and *rs1902586,* respectively. Haplotype analyses showed that the haplotype *CCCTA* (all minor alleles of the five SNPs in block 2) was associated with decreased disease-free survival (HR, 1.9; 95% CI, 1.1-3.3). The nonsynonymous SNP, rs700519 (*Arg264Cys*), located in haplotype block 4, was also associated with breast cancer survival. The age-adjusted HR for the *Cys*/*Cys* (*T*/*T*) genotype was 2.2 (95% CI, 1.2-4.1) for overall survival and 2.1 (95% CI, 1.1-3.9) for disease-free survival, compared with those carrying the *Arg* (*C*) allele. These results suggest that polymorphisms in the *CYP19A1* gene may have effects on breast cancer prognosis. (Cancer Epidemiol Biomarkers Prev 2006;15(11):2115–22)

Introduction

Estrogens stimulate breast cell division and play a crucial role in the pathogenesis and progression of breast cancer. Women with high levels of serum estrogens have an increased risk of breast cancer (1). Estrogen concentrations are 10 to 50 times higher in malignant breast tissue than in plasma (2), and estrogen in breast tissue may act locally to promote the growth of breast carcinomas (3).

During and after menopause, a woman's endogenous estrogen production converts from a predominantly ovarian source to a peripheral source. Aromatase, encoded by the *CYP19A1* gene, is the main enzyme that catalyzes the final and rate-limiting step of estrogen biosynthesis, aromatization of androstenedione and testosterone to estrone and estradiol, respectively. A direct effect of aromatase on *in situ* estrogen synthesis in the breast has been reported (4). Elevated levels of aromatase expression have been observed in breast tumors relative to normal breast tissue (5). This evidence indicates a potential role for the *CYP19A1* gene in the development and progression of breast cancer. The importance of aromatase in the pathogenesis of breast cancer has also clearly been shown in a clinical setting, as inhibitors of this enzyme have been regularly used in the treatment of postmenopausal breast cancer (6). A recent study suggested that aromatase inhibitors might be more effective than modulators of the estrogen receptor in slowing tumor progression (7).

To date, a number of studies have been done to evaluate genetic polymorphisms in the *CYP19A1* gene in relation to breast cancer risk (8-13). However, only one study has examined the association of a particular genetic marker, the (*TTTA*) repeat in intron 4 in this gene, with breast cancer survival (14) and found a significant association between longer repeat length and improved survival. In the present study, we comprehensively evaluated *CYP19A1* genetic variants in relation to breast cancer survival in a large cohort of patients recruited as part of a population-based case-control study, the Shanghai Breast Cancer Study (15).

Materials and Methods

Subjects and Data Collection. In the Shanghai Breast Cancer Study, breast cancer patients were identified through a rapid case ascertainment system, supplemented by the Shanghai Cancer Registry, a population-based tumor registry. A total of 1,602 patients with a primary breast cancer diagnosis were identified between August 1996 and March 1998. Of them, 1,459 were recruited into the study with a response rate of 91% (16). Of the 1,459 breast cancer patients, 4 subjects were excluded from the survival study due to lack of adequate follow-up information. A peripheral blood sample (10 mL from each woman) was obtained from 1,193 patients, 82% of the 1,455 study participants in the survival study. The blood samples were processed within 6 hours of collection and stored at −70°C until the relevant bioassays were conducted. Medical charts were reviewed to abstract information on cancer diagnosis, tumor-node-metastasis (TNM) stage, estrogen receptor and progesterone receptor status, and cancer treatment. Pathologic slides for all cases were reviewed independently by two senior pathologists to confirm the cancer diagnosis.

Patients were followed until July 2005 for cancer recurrence and mortality with a combination of two active follow-up surveys and record linkage to death certificates kept by the Vital Statistics Unit of the Shanghai Center for Disease Control

Grant support: National Cancer Institute grants USPHS RO1CA64277 and RO1CA90899.

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.

Requests for reprints: Wei Zheng, Vanderbilt Center for Epidemiologic Research, Vanderbilt University School of Medicine, 1161 21st Avenue South, S-1121 Medical Center North, Nashville, TN 37232-2587. Phone: 615-936-0682; Fax: 615-322-1754. E-mail: wei.zheng@vanderbilt.edu

doi:10.1158/1055-9965.EPI-06-0464

Table 1. Overall survival by demographics and known breast cancer prognosis factors, the Shanghai Breast Cancer Study

Covariables/levels	All cases (n = 1,455)				Subjects with genotype (n = 1,136)			
	Cases	Deaths	5-y survival (%)*	P	Cases	Deaths	5-y survival (%)†	P
Age at diagnosis (y)								
<42	345	77	84.4	0.02	282	64	84.0	0.01
42-46	357	56	86.8		283	41	88.0	
47-52	365	84	81.4		281	67	81.5	
53-64	388	96	80.9		290	72	81.0	
Education								
<Middle school	177	44	80.2	0.42	140	37	79.3	0.27
Middle school	622	136	83.3		503	109	83.5	
>Middle school	656	133	84.3		493	98	85.2	
Menopause								
Premenopause	950	182	85.5	0.01	758	145	85.5	0.01
Postmenopause	499	130	79.9		373	98	79.6	
TNM†								
0-I	358	32	93.0	<0.001	285	25	93.3	<0.001
IIa	508	81	88.0		401	65	88.3	
IIb	320	83	79.7		252	66	80.2	
III-IV	165	83	59.4		125	61	60.0	
Surgery†								
Yes	1,446	308	83.5		1,130	241	83.8	
No	1	1	0.0		0	0		
Chemotherapy†								
Yes	1,367	290	83.7	0.08	1,063	224	84.1	0.19
No	70	16	81.4		59	15	79.7	
Radiotherapy†								
Yes	566	176	75.3	<0.001	436	139	75.5	<0.001
No	690	102	89.4		538	79	89.4	
Tamoxifen use†								
Yes	921	147	89.1	<0.001	730	127	88.5	<0.001
No	263	54	84.0		212	40	86.3	
Total	1,455	313	83.4		1,136	244	83.6	

*Log-rank test for P value; survival rate derived from Kaplan-Meier analysis.
†Data are missing for a small group of subjects.

and Prevention (17). The median follow-up time for the cohort was 7.1 years. Through interview of patients, or next of kin for deceased patients, we obtained information on disease progression, recurrence, quality of life, and cause of death (if deceased). Of the 1,455 eligible patients, 1,378 were followed-up via in-person contact or by phone (active follow-up) at least once during the two follow-ups. Among them, 266 deaths were identified, 237 from breast cancer, 26 from other diseases, and 3 from unclear causes. Survival status for the 77 participants who were not actively followed was established by linkage to the death registry, and 47 deaths were identified with all but one due to breast cancer. The

Table 2. Summary of the *CYP19A1* genetic markers evaluated in the study

SNP	Alleles*	Location	Position†	Haplotype block‡	Allele frequency (%)§	Genotype frequency (%)‖			P¶
						AA	AB	BB	
rs2446405	**A**/T	Exon 1	49,434,085	1	48.2	24.0	48.4	27.7	0.30
rs2445765	**C**/G	Exon 1	49,422,190	1	26.6	7.0	39.1	53.9	0.94
rs2470144	**T**/C	Exon 1	49,409,017	1	38.9	13.9	49.9	36.2	0.10
rs1004984	**A**/G	Exon 1	49,400,821	1	32.0	9.2	45.6	45.2	0.12
rs1902584	**T**/A	Exon 1	49,398,946	1	14.2	2.0	24.5	73.6	0.88
hCV1664178	**C**/A	Exon 1	49,388,433	2	32.5	10.1	44.8	45.1	0.47
rs12900137	**C**/G	Exon 1	49,386,645	2	16.2	2.7	27.0	70.3	0.81
rs730154	**C**/T	Exon 1	49,378,496	2	33.0	10.3	45.3	44.4	0.42
rs936306	**T**/C	Exon 1	49,366,890	2	33.0	10.2	45.7	44.2	0.28
rs1902586	**A**/G	Exon 1	49,358,145	2	32.3	9.9	44.7	45.4	0.44
rs749292	**A**/G	Exon 1	49,346,023	3	46.6	21.8	49.6	28.6	0.88
rs6493494	**A**/G	Exon 1	49,337,127	3	45.9	18.9	54.1	27.1	<0.01
rs1008805	**G**/A	Exon 1	49,336,891	3	29.5	7.6	43.8	48.6	0.07
rs12907866	**A**/G	Exon 1	49,332,746	4	38.5	22.7	31.5	45.8	<0.01
rs727479	**C**/A	Intron 1	49,321,839	4	27.7	6.7	42.2	51.2	0.09
rs2414096	**A**/G	Intron 1	49,317,071	4	45.7	19.3	52.7	28.0	0.04
rs700519	**T**/C	Cys264Arg (exon 8)	49,295,260	4	15.1	2.1	26.0	71.9	0.58
rs10046	**G**/A	3′-UTR (exon 10)	49,290,278	4	45.8	21.3	48.9	29.8	0.61
rs4646	**A**/C	3′-UTR (exon 10)	49,290,136	4	28.6	7.9	41.4	50.7	0.63

Abbreviation: UTR, untranslated region.
*Minor allele is in boldface.
†Chromosome position based on National Center for Biotechnology Information build 35.
‡Haplotype block that the SNPs belonged to based on Haiman et al. (10).
§Minor allele frequency for each SNP.
‖For each SNP, AA, minor allele homozygote; AB, heterozygote; BB, major allele homozygote.
¶P value is the probability of the χ^2 test for Hardy-Weinberg disequilibrium.

remaining 30 patients had no match in the death registry and were assumed to be still living on December 2004, 6 months before our last search of the vital statistics registry, to allow for a possible delay of entry of death certificates into the registry. Breast cancer relapse information was not available for these 30 patients and the one person who died from other diseases. Including the three women who lacked detailed information on cause of death through active follow-up, 34 women in total were excluded from the disease-free survival analysis. The study was approved by the institutional review boards of all participating institutes. Informed consent was obtained from each participant.

Genetic Marker Selection. The genetic markers of the *CYP19A1* gene included in this study were selected based on a report from the Multiethnic Cohort Study (10). Based on haplotype analyses of 74 densely spaced single-nucleotide polymorphisms (SNP) across the gene, four haplotype blocks (blocks 1-4) and 19 haplotype tagging SNPs were identified in a Japanese population. Because of the similarity in linkage disequilibrium patterns between Chinese and Japanese populations (18), we used these 19 haplotype tagging SNPs in the present study.

Genotyping. Genomic DNA was extracted from buffy coats with a Puregene DNA Purification Kit (Gentra Systems, Minneapolis, MN) following the protocol of the manufacturer. SNP rs1004984 was genotyped with a high-throughput Masscode assay at BioServe Biotechnologies, Ltd. (Laurel, MD). The nonsynonymous SNP, rs700519, was genotyped by PCR-RFLP and confirmed by direct sequencing with BigDye Terminator Chemistry on an ABI 3700 (ABI, Applied Biosystems, Foster City, CA). Genotyping of the other 17 SNPs was done by running the 5′nuclease TaqMan allelic discrimination assay using an ABI 7900 (ABI). Details about assays, primers, probes, and procedures are available on request.

Laboratory staff were blind to the identity of the subjects. Quality control samples were included in the genotyping assays. Each 96-well plate of genomic DNA contained multiple quality controls, including one water, two samples of CEPH 1347-02, two known duplicates, and two blinded duplicates. The average agreement of the genotypes for these 19 markers determined for the blinded quality control samples was 98.7%.

Statistical Analyses. Included in the present study were 1,136 patients with both survival status and genotype information. The primary outcomes were disease-free survival and overall survival. The endpoints included cancer recurrence/metastasis or death due to breast cancer for the analysis of disease-free survival, and death from any cause for the analysis of overall survival. Survival time was calculated as the time from cancer diagnosis until the occurrence of the study endpoints, censoring at the date of last contact or noncancer death (for disease-free survival only). The Kaplan-Meier

Table 3. Association of SNPs located in haplotype block 2 and the nonsynonymous SNP in haplotype block 4 with breast cancer survival in Chinese women

SNP	Overall survival					Disease-free survival*				
	Events	5-y survival (%)†	P‡	HR (95% CI)§	P	Events	5-y survival (%)†	P‡	HR (95% CI)§	P
hCV1664178										
AA	102	85.4	0.63	1.0 (reference)	0.43	135	76.3	0.07	1.0 (reference)	0.24
AC	109	82.9		1.1 (0.8-1.4)		127	77.7		0.9 (0.7-1.2)	
CC	27	80.5		1.2 (0.8-1.8)		41	69.1		1.4 (1.0-2.0)	
AA/AC	202	84.1	0.47	1.0 (reference)	0.51	262	77.1	0.02	1.0 (reference)	0.02
CC	27	80.5		1.2 (0.8-1.7)		41	69.1		1.5 (1.1-2.1)	
rs12900137										
GG	159	84.9	0.23	1.0 (reference)	0.32	206	76.2	0.02	1.0 (reference)	0.24
CG	63	82.4		1.0 (0.8-1.4)		77	76.9		1.0 (0.7-1.3)	
CC	10	73.3		1.7 (0.9-3.2)		14	56.9		2.1 (1.2-3.6)	
GG/CG	213	84.2	0.09	1.0 (reference)	0.10	283	76.7	0.01	1.0 (reference)	0.01
CC	10	73.3		1.7 (0.9-3.2)		14	56.9		2.1 (1.2-3.6)	
rs730154										
TT	103	84.9	0.58	1.0 (reference)	0.46	136	75.8	0.05	1.0 (reference)	0.35
CT	108	83.5		1.0 (0.8-1.3)		127	78.2		0.9 (0.7-1.1)	
CC	29	79.3		1.2 (0.8-1.8)		42	68.9		1.4 (1.0-2.0)	
TT/CT	202	84.2	0.31	1.0 (reference)	0.34	263	76.7	0.02	1.0 (reference)	0.02
CC	29	79.3		1.2 (0.8-1.8)		42	68.9		1.5 (1.1-2.0)	
rs936306										
CC	101	85.2	0.60	1.0 (reference)	0.42	134	76.0	0.12	1.0 (reference)	0.38
CT	111	83.0		1.1 (0.8-1.4)		130	77.7		0.9 (0.7-1.2)	
TT	28	79.8		1.2 (0.8-1.8)		40	70.2		1.4 (1.0-1.9)	
CC/CT	203	84.1	0.39	1.0 (reference)	0.42	264	76.5	0.05	1.0 (reference)	0.04
TT	28	79.8		1.2 (0.8-1.7)		40	70.2		1.4 (1.0-2.0)	
rs1902586										
GG	102	85.3	0.52	1.0 (reference)	0.29	135	76.3	0.15	1.0 (reference)	0.23
AG	110	82.9		1.1 (0.8-1.4)		130	77.0		1.0 (0.8-1.2)	
AA	27	79.1		1.3 (0.8-1.9)		39	69.4		1.4 (1.0-2.0)	
GG/AG	203	84.1	0.39	1.0 (reference)	0.38	265	76.4	0.05	1.0 (reference)	0.04
AA	27	79.1		1.2 (0.8-1.8)		39	69.4		1.4 (1.0-2.0)	
rs700519										
CC	174	84.1	0.03	1.0 (reference)	0.48	220	75.6	0.06	1.0 (reference)	0.49
CT	58	83.2		0.9 (0.7-1.2)		76	77.3		1.0 (0.7-1.2)	
TT	10	65.2		2.2 (1.1-4.1)		10	55.0		2.1 (1.1-3.9)	
CC/CT	232	83.8	0.01	1.0 (reference)	0.02	296	76.3	0.02	1.0 (reference)	0.02
TT	10	65.2		2.2 (1.2-4.1)		10	55.0		2.1 (1.1-3.9)	

*Subjects with missing information on disease relapse (n = 34) were excluded from disease-free survival analysis.
†Survival rate derived from Kaplan-Meier analysis.
‡Log-rank test for P value.
§Adjusted for age at diagnosis.

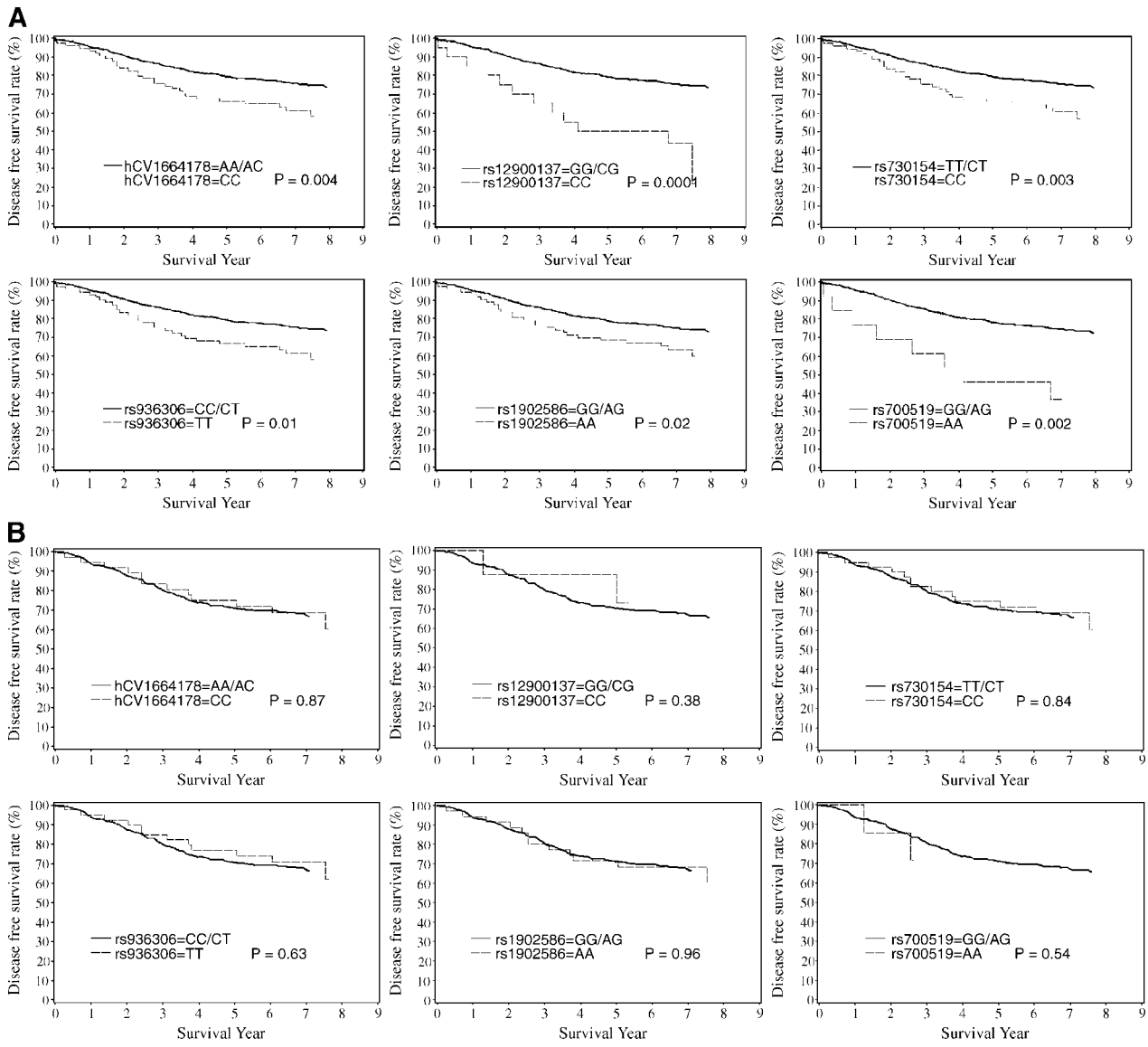

Figure 1. Survival of breast cancer patients after diagnosis stratified by genotypes for the five SNPs located in haplotype block 2 and the nonsynonymous SNP in block 4. **A,** premenopausal women; **B,** postmenopausal women.

method was used to estimate the survival function, and differences in survival across groups defined by genotypes were examined using the log-rank test. The Cox regression model was employed to compute hazard ratios (HR).

The program PHASE, which is based on a Bayesian statistical model (19), was used to reconstruct haplotypes within each block of the *CYP19A1* gene. The association of haplotypes and breast cancer survival was evaluated using a Cox model by treating each haplotype as a continuous variable with probabilistically assigned values, as described by Zaykin et al. (20). All statistical analyses were done with SAS version 9.1 (SAS Institute, Cary, NC) and all tests were based on two-sided probability.

Results

Table 1 shows selected demographic factors and TNM stages for breast cancer and the corresponding 5-year overall survival rate. The descriptive characteristics were similar for the entire cohort of 1,455 breast cancer patients and the 1,136 subjects who were included in the present study. Most of the patients received surgery, adjuvant chemotherapy, and tamoxifen treatment. As expected, TNM stage was the major prognostic factor for survival.

Genotype and allele distributions for the 19 SNPs are summarized in Table 2. With the exception of *rs6493494, rs12907866*, and *rs2414096,* all other SNPs were consistent with the Hardy-Weinberg equilibrium distribution ($P > 0.05$). Table 3 and Fig. 1 show the results of association analyses for the genetic polymorphisms and breast cancer survival. Significant associations were observed for disease-free survival for each of the five SNPs located in haplotype block 2. Patients homozygous for the minor alleles had reduced 5-year disease-free survival rates compared with those carrying the major allele (69.1% versus 77.1%, 56.9% versus 76.7%, 68.9% versus 76.7%, 70.2% versus 76.5%, and 69.4% versus 76.4% for *hCV1664178, rs12900137, rs730154, rs936306*, and *rs1902586,* respectively). The age-adjusted HRs and 95% confidence intervals (95% CI) were 1.5 (1.1-2.1), 2.1 (1.2-3.6), 1.5 (1.1-2.0), 1.4 (1.0-2.0), and 1.4 (1.0-2.0) for *hCV1664178, rs12900137, rs730154, rs936306*, and *rs1902586,* respectively. The nonsynonymous SNP, *rs700519* (*Arg264Cys*), located in haplotype block 4, was also associated with breast cancer survival. The age-adjusted HR (95% CI) for

the *Cys/Cys* (*T/T*) genotype was 2.2 (1.2-4.1) for overall survival and 2.1 (1.1-3.9) for disease-free survival compared with those carrying the major allele, *Arg* (*C*). The Kaplan-Meier survival curves presented in Fig. 1 show apparent associations of disease-free survival after breast cancer diagnosis with the genotypes defined by these six SNPs. The distributions of the genotypes of these SNPs were similar across TNM stage and estrogen receptor/progesterone receptor status. Additional adjustment for these tumor characteristics did not materially change the genotype-survival associations (data not shown). The observed associations were similar between tamoxifen users and nonusers (data not shown). Analysis stratified by menopause status indicated that the above associations were mainly evident in premenopausal women. No association was observed in postmenopausal women (Table 4). No clear association was observed for the other 10 SNPs in relation to either overall survival or disease-free survival (data not shown).

Haplotype analyses were conducted to evaluate the combined effect of the SNPs located within each haplotype block (Table 5). The genotype distributions of three SNPs, *rs6493494, rs12907866*, and *rs2414096,* deviated from Hardy-Weinberg equilibrium (Table 2) and, thus, they were eliminated from the haplotype reconstruction. The frequencies of major haplotypes are similar to those observed in a Japanese population (10). An association was observed for the haplotype *CCCTA* (minor alleles for *hCV1664178, rs12900137, rs730154, rs936306*, and *rs1902586* in block 2) for disease-free survival with an age-adjusted HR of 1.9 (95% CI, 1.1-3.3). The association, however, was not statistically significant for overall survival. Analysis stratified by menopause status indicated that the association of haplotype *CCCTA* with survival was mainly evident among premenopausal women, with HRs being 2.3 (95% CI, 1.1-4.6) for overall survival and 2.8 (95% CI, 1.5-5.2) for disease-free survival. No association was observed among postmenopausal women. The haplotype *AAGC* in block 4 was associated with both overall survival (HR, 1.7; 95% CI, 0.9-3.5) and disease-free survival (HR, 2.6; 95% CI, 1.2-5.6) among premenopausal women. We did not find any associations of other haplotypes with either disease-free survival or overall survival. Including the three SNPs that deviated from the Hardy-Weinberg equilibrium did not affect the association of the haplotypes with breast cancer survival described above (data not shown).

Discussion

This is one of the first studies to evaluate the associations of *CYP19A1* genetic polymorphisms with breast cancer survival. We have identified associations between all of the five SNPs located in haplotype block 2 and survival. Women who are homozygous for the minor allele at any of the five SNPs located in block 2 have a decreased rate of disease-free survival compared with those carrying the major allele. Haplotype analyses confirmed the above associations. Haplotype *CCCTA*

Table 4. Association of SNPs located in haplotype block 2 and the nonsynonymous SNP with breast cancer survival, stratified by menopause status

Marker	Overall survival						Disease-free survival[b]					
	Premenopausal women			Postmenopausal women			Premenopausal women			Postmenopausal women		
	Event	HR (95% CI)[a]	*P*[a]	Event	HR (95% CI)[a]	*P*[a]	Event	HR (95% CI)[a]	*P*[a]	Event	HR (95% CI)[a]	*P*[a]
hCV1664178												
AA	63	1.0 (reference)	0.29	38	1.0 (reference)	0.93	87	1.0 (reference)	0.12	48	1.0 (reference)	0.92
AC	59	1.0 (0.7-1.5)		50	1.2 (0.8-1.8)		71	0.9 (0.7-1.2)		56	1.0 (0.7-1.5)	
CC	19	1.4 (0.8-2.3)		8	0.9 (0.4-1.8)		29	1.7 (1.1-2.6)		12	1.0 (0.5-1.8)	
AA/AC	122	1.0 (reference)	0.21	88	1.0 (reference)	0.52	158	1.0 (reference)	0.004	104	1.0 (reference)	0.87
CC	19	1.4 (0.8-2.2)		8	0.8 (0.4-1.6)		29	1.8 (1.2-2.7)		12	1.0 (0.5-1.7)	
rs12900137												
GG	97	1.0 (reference)	0.57	61	1.0 (reference)	0.44	130	1.0 (reference)	0.35	76	1.0 (reference)	0.65
CG	29	0.8 (0.5-1.2)		34	1.4 (0.9-2.2)		39	0.8 (0.6-1.1)		38	1.3 (0.9-1.9)	
CC	9	2.4 (1.2-4.8)		1	0.4 (0.1-2.8)		12	2.9 (1.6-5.3)		2	0.6 (0.1-2.3)	
GG/CG	126	1.0 (reference)	0.01	95	1.0 (reference)	0.29	169	1.0 (reference)	0.0001	114	1.0 (reference)	0.38
CC	9	2.6 (1.3-5.0)		1	0.3 (0.1-2.5)		12	3.1 (1.7-5.6)		2	0.5 (0.1-2.2)	
rs730154												
TT	62	1.0 (reference)	0.25	40	1.0 (reference)	0.91	86	1.0 (reference)	0.12	50	1.0 (reference)	0.61
CT	61	1.1 (0.7-1.5)		47	1.0 (0.7-1.6)		73	0.9 (0.7-1.2)		54	0.9 (0.6-1.3)	
CC	19	1.4 (0.9-2.4)		10	0.9 (0.5-1.9)		29	1.7 (1.1-2.6)		13	0.9 (0.5-1.7)	
TT/CT	123	1.0 (reference)	0.18	87	1.0 (reference)	0.80	159	1.0 (reference)	0.003	104	1.0 (reference)	0.84
CC	19	1.4 (0.9-2.3)		10	0.9 (0.5-1.8)		29	1.8 (1.2-2.7)		13	0.9 (0.5-1.7)	
rs936306												
CC	62	1.0 (reference)	0.24	38	1.0 (reference)	0.96	86	1.0 (reference)	0.15	48	1.0 (reference)	0.64
CT	62	1.1 (0.7-1.5)		49	1.1 (0.7-1.7)		74	0.9 (0.7-1.2)		56	1.0 (0.7-1.4)	
TT	19	1.5 (0.9-2.4)		9	0.9 (0.4-1.8)		28	1.7 (1.1-2.6)		12	0.9 (0.5-1.6)	
CC/CT	124	1.0 (reference)	0.16	87	1.0 (reference)	0.57	160	1.0 (reference)	0.01	104	1.0 (reference)	0.63
TT	19	1.4 (0.9-2.3)		9	0.8 (0.4-1.6)		28	1.8 (1.2-2.7)		12	0.9 (0.5-1.6)	
rs1902586												
GG	61	1.0 (reference)	0.18	40	1.0 (reference)	0.91	85	1.0 (reference)	0.10	50	1.0 (reference)	0.86
AG	63	1.2 (0.8-1.6)		47	1.1 (0.7-1.7)		76	1.0 (0.7-1.4)		54	1.0 (0.7-1.4)	
AA	18	1.4 (0.9-2.4)		9	1.0 (0.5-2.0)		27	1.6 (1.1-2.5)		12	1.0 (0.5-1.8)	
GG/AG	124	1.0 (reference)	0.25	87	1.0 (reference)	0.84	161	1.0 (reference)	0.02	104	1.0 (reference)	0.96
AA	18	1.3 (0.8-2.2)		9	0.9 (0.5-1.9)		27	1.6 (1.1-2.5)		12	1.0 (0.5-1.8)	
rs700519												
GG	101	1.0 (reference)	0.22	73	1.0 (reference)	0.53	135	1.0 (reference)	0.27	85	1.0 (reference)	0.68
AG	35	0.9 (0.6-1.4)		22	0.9 (0.5-1.4)		46	0.9 (0.7-1.3)		30	1.0 (0.7-1.5)	
AA	8	2.9 (1.4-6.0)		2	0.9 (0.2-3.6)		8	3.1 (1.5-6.3)		2	0.6 (0.2-2.7)	
GG/AG	136	1.0 (reference)	0.003	95	1.0 (reference)	0.88	181	1.0 (reference)	0.002	115	1.0 (reference)	0.54
AA	8	3.0 (1.5-6.1)		2	0.9 (0.2-3.8)		8	3.1 (1.5-6.4)		2	0.6 (0.2-2.7)	

NOTE: a, adjusted for age at diagnosis; b, subjects with missing information on disease relapse (n = 34 in total) were excluded from disease-free survival analysis.

Table 5. The *CYP19A1* gene haplotypes in association with breast cancer survival

Block	Haplotype[a]	Haplotype[b] frequency	All women				Premenopausal	
			Overall survival		Disease-free survival		Overall survival	
			HR (95% CI)[c,d]	P^d	HR (95% CI)[c,d]	P^d	HR (95% CI)[c,d]	P^d
Block 1	AGTGA	0.379	0.9 (0.6-1.4)	0.69	1.1 (0.8-1.5)	0.71	0.7 (0.4-1.3)	0.28
	TGCGA	0.244	0.5 (0.3-1.2)	0.12	0.6 (0.3-1.2)	0.15	0.8 (0.4-1.6)	0.50
	TCCAT	0.140	1.4 (0.6-3.0)	0.46	1.1 (0.5-2.5)	0.77	1.9 (0.8-4.3)	0.12
	AGCAA	0.093	0.6 (0.1-4.1)	0.62	1.0 (0.2-3.8)	0.90	0.8 (0.1-5.5)	0.79
	TCCAA	0.085	0.0 (0-249.1)	0.47	0.1 (0.0-25.8)	0.37		
Block 2	AGTCG	0.662	0.9 (0.7-1.2)	0.50	1.0 (0.8-1.2)	0.90	0.8 (0.6-1.2)	0.33
	CCCTA	0.157	1.5 (0.8-3.0)	0.22	1.9 (1.1-3.3)	0.03	2.3 (1.1-4.6)	0.03
	CGCTA	0.156	0.6 (0.2-1.8)	0.34	0.8 (0.3-1.9)	0.55	1.3 (0.4-4.2)	0.67
Block 3	AA	0.465	1.0 (0.7-1.4)	0.99	1.1 (0.8-1.4)	0.63	1.0 (0.7-1.5)	0.97
	GG	0.292	1.1 (0.7-1.8)	0.75	1.1 (0.7-1.6)	0.76	0.9 (0.5-1.6)	0.63
	GA	0.240	1.1 (0.6-1.8)	0.85	1.2 (0.8-2.0)	0.37	1.1 (0.6-2.3)	0.69
Block 4	ACAC	0.523	0.9 (0.7-1.2)	0.58	0.9 (0.7-1.2)	0.64	1.1 (0.8-1.6)	0.50
	CCGA	0.252	1.0 (0.6-1.7)	0.89	1.0 (0.6-1.6)	0.89	0.8 (0.4-1.6)	0.49
	ATGC	0.143	1.9 (0.9-3.8)	0.09	1.7 (0.9-3.5)	0.13	2.6 (1.2-5.6)	0.01

NOTE: a, SNPs were arranged in the order of rs2446405-rs2445765-rs2470144-rs1004984-rs1902584 for block 1; hCV1664178-rs12900137-rs730154-rs936306-rs1902586 for block 2; rs749292-rs1008805 for block 3; and rs727479-rs700519-rs10046-rs4646 for block 4; b, haplotype frequency was derived using the program PHASE; c, HRs were derived from the Cox regression model by treating the probability of each haplotype as a continuous independent variable-recessive model; d, adjusted for age at diagnosis.

(all minor alleles of the five SNPs in block 2) was associated with decreased disease-free survival. We also found associations of the minor allele homozygote of *rs700519* (*Cys264Arg*) with both overall survival and disease-free survival. These findings are biologically plausible, given the pivotal role of the *CYP19A1* gene in estrogen metabolism, its potential role in tumor growth and progression, and the potential functional significance of these polymorphisms.

It is widely accepted that estrogen plays an important role in the growth and progression of human breast cancer by disrupting the normal balance between cell differentiation and proliferation (21). Aromatase catalyzes the biosynthesis of estrogen in the adipose tissues through the conversion of androgens (22). Currently, aromatase inhibitors are regularly used in the treatment of postmenopausal breast cancer (6). Thus, it is biologically plausible that the polymorphisms in the *CYP19A1* gene, which encodes the aromatase, may be associated with breast cancer survival. In our study, the association was observed in premenopausal women only. The reasons for this specific association, however, were unclear. Menopause status for our study subjects was recorded at the time of cancer diagnosis, and it is likely that the majority of the premenopausal women included in our study had menopause not long after cancer diagnosis due to breast cancer adjuvant therapies. The treatment-induced menopause suddenly cuts off the supply of ovary-synthesized estrogen to breast cancer cells that initially grow in a high-estrogen environment (premenopausal breast tissues). These cancers may have a greater need of estrogen for their growth than those developed in postmenopausal women, and after menopause, most of the estrogens were synthesized in adipose tissues by aromatase. Therefore, it is conceivable that aromatase activity (and thus *CYP19A1* gene variants) may be more closely related to breast cancer survival in premenopausal than in postmenopausal women.

The *CYP19A1* gene, located at 15q21.2, is composed of nine coding exons (II-X) covering ~30 kb and at least 10 different untranslated first exons that are regulated by tissue-specific promoters (23). *CYP19A1* gene expression has been suggested to play a role in neoplastic proliferation in human breast carcinomas. Elevated levels of *CYP19A1* mRNA have been observed in breast cancer tissue as compared with normal breast tissue (5). The increased expression of *CYP19A1* in breast cancer tissues was associated with a switch in the promoter utilization. Adipose tissue, in general, including the adipose tissue of a normal breast, maintains low levels of aromatase expression primarily via promoter I.4, whereas promoters I.3, I.7, and II are used only minimally (24). However, in breast cancer, activities of the latter three promoters are strikingly increased (5, 24).

Promoter I.7, recently cloned by Sebastian et al. (25), is located ~36 kb upstream of the coding region. It is unique in that it is a GATA-2-regulated endothelial promoter and its usage is correlated with the extent of angiogenesis in breast cancer tissue (25). Interestingly, angiogenesis has been postulated to be one of the mechanisms for resistance to hormone therapy (26). Consistent with the potential role of this promoter in breast cancer pathology, we found an association between minor alleles of all five SNPs located around this promoter (in block 2) and decreased disease-free survival of breast cancer, both by individual locus analysis and by haplotype analysis. To our knowledge, our study is the first to identify the association of these polymorphisms around promoter I.7 and breast cancer survival. In line with our results, the haplotype *CCCTA* was also found to be associated with an elevated risk of breast cancer with an odds ratio of 1.2 (95% CI, 1.1-1.4) in the Multiethnic Cohort Study (10).

Promoters II and I.3 are in block 4, where we genotyped six SNPs. Association was observed at the SNP rs700519. The *T/T* (*Cys/Cys*) genotype was associated with a lower survival rate for both overall survival and disease-free survival. This SNP results in amino acid changes from *Arg* to *Cys*, which may lead to a difference in enzyme activity. Although no reports are available about the relationship between this polymorphism and breast cancer progression, several studies have reported a relationship between the *Cys* allele and increased breast cancer risk, with odds ratios (95% CIs) of 2.4 (1.0-5.5) in Hawaiians (10), 1.4 (1.1-1.9) in Japanese women (10), and 1.5 (1.1-2.2) in Korean women (27), although no association was observed in Caucasians (28) or another Japanese population (29). We did not find any apparent associations with breast cancer survival through haplotype analyses for the SNPs located in block 4, nor did the Multiethnic Cohort Study for breast cancer risk (10).

This is the first study to comprehensively evaluate the association of *CYP19A1* genetic polymorphisms with breast cancer survival. The study hypotheses were based on sound biologic plausibility: the critical role of the *CYP19A1* gene in estrogen biosynthesis. Selection bias was minimized given the population-based study design, high response rate, and high follow-up rate. In addition, >98% of the patients in this

Table 5. The *CYP19A1* gene haplotypes in association with breast cancer survival (Cont'd)

Premenopausal		Postmenopausal			
Disease-free survival		Overall survival		Disease-free survival	
HR (95% CI)[c,d]	P[d]	HR (95% CI)[c,d]	P[d]	HR (95% CI)[c,d]	P[d]
0.9 (0.6-1.4)	0.72	1.1 (0.7-1.9)	0.69	1.2 (0.7-1.9)	0.46
0.8 (0.4-1.6)	0.54			0.1 (0.0-3.0)	0.18
1.6 (0.7-3.7)	0.23	0.1 (0.0-17.8)	0.41	0.1 (0.0-15.3)	0.35
1.3 (0.3-5.2)	0.73	0.2 (0.0-414.2)	0.69	0.2 (0.0-427)	0.65
		0.1 (0.0-1,137.5)	0.07	0.1 (0.0-67.3)	0.52
0.9 (0.7-1.2)	0.57	1.0 (0.7-1.5)	0.96	1.1 (0.8-1.6)	0.63
2.8 (1.5-5.2)	0.001	0.4 (0.05-2.6)	0.31	0.5 (0.1-2.2)	0.39
1.4 (0.5-3.9)	0.53	0.0 (0.0-7,000)	0.60	0.2 (0.0-1.8)	0.17
1.1 (0.7-1.5)	0.72	1.0 (0.6-1.5)	0.84	1.0 (0.7-1.5)	0.96
0.8 (0.5-1.4)	0.48	1.6 (0.7-3.2)	0.24	1.8 (0.9-3.5)	0.08
1.3 (0.7-2.3)	0.35	0.8 (0.3-2.2)	0.72	1.0 (0.4-2.2)	0.98
1.1 (0.8-1.6)	0.43	0.7 (0.4-1.1)	0.12	0.7 (0.4-1.0)	0.07
0.8 (0.4-1.5)	0.48	1.3 (0.6-3.0)	0.51	1.4 (0.6-2.9)	0.43
2.7 (1.3-5.8)	0.01	0.5 (0.1-3.7)	0.50	0.4 (0.1-2.6)	0.31

study belong to a single ethnic group; thus, the potential confounding effect due to ethnicity is limited. The large sample size and long follow-up period provided high statistical power. A potential concern for the study is that as relapse and cause of death information was collected on the basis of self-reports or death certificate information, some errors may exist. However, these errors are likely to be random, which is likely to attenuate the association observed in the study. Three tagging SNPs selected initially for the study were excluded from the data analyses because of their deviation from Hardy-Weinberg equilibrium. This may affect the accuracy of the study in predicting common haplotypes. However, the results from our study about the association between *CYP19A1* genotypes and breast cancer survival should not be affected by these three SNPs. The haplotype tagging SNPs were selected based on the extensive work conducted by Haimen et al. (10). Using the latest data from the HapMap project, a panel of 38 tagging SNPs was required to capture all 170 common variants for the *CYP19A1* gene and its adjacent region with pairwise $r^2 \geq 0.80$. Among our 19 genotyped SNPs, 15 SNPs were included in the HapMap project and they can capture 91 of the 170 common variants with $r^2 \geq 0.80$. These data provide additional assurance that the total SNP panel used in our study should be able to capture most of the common variants.

In summary, this study provides evidence for associations of breast cancer survival with SNPs located in haplotype block 2 and the nonsynonymous SNP in the *CYP19A1* gene. These results are novel and, if confirmed, may have significant clinical implications for breast cancer treatment.

Acknowledgments

We thank Dr. Qi Dai, Dr. Fan Jin, and Jia-Rong Cheng for their contributions in coordinating data and specimen collection in Shanghai; Qing Wang and Regina Courtney for technique assistance in genotyping assays; Bethanie Hull for technical assistance in the preparation of this manuscript; Dr. Hui Cai for statistical consultation in data analysis; and Dr. Christopher A. Haiman from the University of Southern California for sharing the information for primers and probes of several SNPs.

References

1. Toniolo PG, Levitz M, Zeleniuch-Jacquotte A, et al. A prospective study of endogenous estrogens and breast cancer in postmenopausal women. J Natl Cancer Inst 1995;87:190–7.
2. Pasqualini JR, Chetrite G, Blacker C, et al. Concentrations of estrone, estradiol, and estrone sulfate and evaluation of sulfatase and aromatase activities in pre and postmenopausal breast cancer patients. J Clin Endocrinol Metab 1996;81:1460–4.
3. Bulun SE, Price TM, Aitken J, Mahendroo MS, Simpson ER. A link between breast cancer and local estrogen biosynthesis suggested by quantification of breast adipose tissue aromatase cytochrome *P*450 transcripts using competitive polymerase chain reaction after reverse transcription. J Clin Endocrinol Metab 1993;77:1622–8.
4. Maggiolini M, Bonofiglio D, Pezzi V, et al. Aromatase overexpression enhances the stimulatory effects of adrenal androgens on MCF7 breast cancer cells. Mol Cell Endocrinol 2002;193:13–8.
5. Irahara N, Miyoshi Y, Taguchi T, Tamaki Y, Noguchi S. Quantitative analysis of aromatase mRNA expression derived from various promoters (I.4, I.3, PII and I.7) and its association with expression of TNF-α, IL-6 and COX-2 mRNAs in human breast cancer. Int J Cancer 2006;118:1915–21.
6. Mouridsen HT, Robert NJ. Benefit with aromatase inhibitors in the adjuvant setting for postmenopausal women with breast cancer. Med Gen Med 2005;7:20.
7. Dowsett M, Cuzick J, Wale C, Howell T, Houghton J, Baum M. Retrospective analysis of time to recurrence in the ATAC trial according to hormone receptor status: an hypothesis-generating study. J Clin Oncol 2005;23:7512–7.
8. 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.
9. Haiman CA, Hankinson SE, Spiegelman D, Brown M, Hunter DJ. No association between a single nucleotide polymorphism in CYP19 and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:215–6.
10. Haiman CA, Stram DO, Pike MC, et al. A comprehensive haplotype analysis of CYP19 and breast cancer risk: the Multiethnic Cohort. Hum Mol Genet 2003;12:2679–92.
11. Baxter SW, Choong DY, Eccles DM, Campbell IG. Polymorphic variation in CYP19 and the risk of breast cancer. Carcinogenesis 2001;22:347–9.
12. Ahsan H, Whittemore AS, Chen Y, et al. Variants in estrogen-biosynthesis genes CYP17 and CYP19 and breast cancer risk: a family-based genetic association study. Breast Cancer Res 2005;7:R71–81.
13. 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.
14. Goode EL, Dunning AM, Kuschel B, et al. Effect of germ-line genetic variation on breast cancer survival in a population-based study. Cancer Res 2002;62:3052–7.
15. Gao YT, Shu XO, Dai Q, et al. Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. Int J Cancer 2000;87:295–300.
16. Lu H, Shu XO, Cui Y, et al. Association of genetic polymorphisms in the VEGF gene with breast cancer survival. Cancer Res 2005;65:5015–9.
17. Shu XO, Gao YT, Cai Q, et al. Genetic polymorphisms in the TGF-beta 1 gene and breast cancer survival: a report from the Shanghai Breast Cancer Study. Cancer Res 2004;64:836–9.
18. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. Nature 2005;437:1299–320.
19. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001;68:978–89.
20. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. Hum Hered 2002;53:79–91.
21. Bouchard MF, Taniguchi H, Viger RS. Protein kinase A-dependent synergism between gata factors and the nuclear receptor, liver receptor homolog-1, regulates human aromatase (CYP19) PII promoter activity in breast cancer cells. Endocrinology 2005;146:4905–16.

22. Paridaens R, Thomas J, Wildiers J, et al. Safety, activity and estrogen inhibition by exemestane in postmenopausal women with advanced breast cancer: a phase I study. Anticancer Drugs 1998;9:675–83.
23. Simpson ER. Aromatase: biologic relevance of tissue-specific expression. Semin Reprod Med 2004;22:11–23.
24. Bulun SE, Takayama K, Suzuki T, Sasano H, Yilmaz B, Sebastian S. Organization of the human aromatase p450 (CYP19) gene. Semin Reprod Med 2004;22:5–9.
25. Sebastian S, Takayama K, Shozu M, Bulun SE. Cloning and characterization of a novel endothelial promoter of the human CYP19 (aromatase P450) gene that is up-regulated in breast cancer tissue. Mol Endocrinol 2002;16:2243–54.
26. Foekens JA, Peters HA, Grebenchtchikov N, et al. High tumor levels of vascular endothelial growth factor predict poor response to systemic therapy in advanced breast cancer. Cancer Res 2001;61:5407–14.
27. Lee KM, Abel J, Ko Y, et al. Genetic polymorphisms of cytochrome P450 19 and 1B1, alcohol use, and breast cancer risk in Korean women. Br J Cancer 2003;88:675–8.
28. Hefler LA, Tempfer CB, Grimm C, et al. Estrogen-metabolizing gene polymorphisms in the assessment of breast carcinoma risk and fibroadenoma risk in Caucasian women. Cancer 2004;101:264–9.
29. Miyoshi Y, Iwao K, Ikeda N, Egawa C, Noguchi S. Breast cancer risk associated with polymorphism in CYP19 in Japanese women. Int J Cancer 2000;89:325–8.