Population-Based Case-Control Study of *CYP11A* Gene Polymorphism and Breast Cancer Risk

Wei Zheng,[1,3] Yu-Tang Gao,[2] Xiao-Ou Shu,[1,3] Wanqing Wen,[1] Qiuyin Cai,[1] Qi Dai,[1] and Jeffrey R. Smith[1,3]

[1]Department of Medicine and Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, Tennessee; [2]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China; and [3]Medical Research Service, VA Tennessee Valley Healthcare System, Nashville, Tennessee

Abstract

The *CYP11A* gene encodes the cholesterol side-chain cleavage enzyme (P450scc) that catalyzes the first and rate-limiting step for the biosynthesis of sex hormones. A pentanucleotide repeat [(*TAAAA*)n] polymorphism in the 5′ of the *CYP11A* gene has been reported to be related to the risk of polycystic ovary syndrome, an inherited endocrine disorder characterized by hyperandrogenemia. We investigated the association of this polymorphism with breast cancer risk in a population-based case-control study conducted among Chinese women in Shanghai. Genotype assays were completed for 1015 incident breast cancer cases and 1082 community controls. Three common alleles with 4, 6, or 8 *TAAAA* repeats were identified in the study population. The frequency of the 8 repeat allele was more common in cases (12.6%) than controls (8.5%) (odds ratio = 1.6, 95% confidence interval = 1.3–1.9; $P < 0.0001$). Compared to subjects who did not carry this allele, adjusted odds ratios were 1.5 (95% confidence interval = 1.2–1.9) and 2.9 (1.3–6.7) (P for trend, <0.001), respectively, for those who carried one and two copies of this allele. This positive association was observed in both pre- and postmenopausal women and all strata defined by major breast cancer risk factors, including years of menstruation, body mass index, and waist-to-hip ratio. The results from this study indicate that the *TAAAA* repeat polymorphism near the promoter region of the *CYP11A* gene may be an important susceptibility factor for breast cancer risk. (Cancer Epidemiol Biomarkers Prev 2004;13(5):709–14)

Introduction

The *CYP11A* gene encodes cholesterol side-chain cleavage enzyme (P450scc) that catalyzes the conversion of cholesterol to pregnenolone (1). This is the first and rate-limiting step for the biosynthesis of all steroid hormones, including estrogens, progesterone, and androgens (1). These sex hormones, particularly estrogens, have been shown in both *in vitro* and *in vivo* experiments to play an important role in the pathogenesis of breast cancer (2). High levels of sex hormones, including estrogens and androgens, have been found in epidemiological studies, including large prospective studies, to be associated with an increased risk of breast cancer (2, 3). The *CYP11A* gene is comprised of nine exons spanning 29,864 bp on chromosome 15q24.1, a region harboring additional P450 genes. It is primarily expressed in the steroidogenic tissues: adrenal cortex, gonads, and placenta. The promoter region of this gene is believed to contain multiple cAMP-regulated elements that are responsible for increasing basal transcriptional activity (1, 4–7). Recently, a pentanucleotide repeat [(*TAAAA*)n] polymorphism was identified (reported initially as a [(*TTTTA*)n] polymorphism) in the promoter region of the *CYP11A* gene, located at 529 bp upstream from the translation start site (487 bp upstream from exon 1) (4). This polymorphism was found in several studies to be associated with the risk of polycystic ovary syndrome (PCOS) (4, 8–11), a suspected risk factor for breast cancer (12), and has been correlated with the level of total serum testosterone (4, 8–10). Given the pivotal role of the *CYP11A* gene in steroid sex hormone synthesis, it is conceivable that genetic polymorphisms of this gene may be related to the risk of breast cancer. We evaluated this hypothesis in a large population-based case-control study, the Shanghai Breast Cancer Study.

Materials and Methods

Included in this study were subjects recruited from 1996 to 1998 for the Shanghai Breast Cancer Study. Detailed study methods have been published elsewhere (13, 14). Briefly, this study included 1459 incident breast cancer cases diagnosed at an age between 25 and 64 years and 1556 age frequency-matched community controls. All subjects were permanent residents of urban Shanghai. They had no prior history of any cancer and were alive at the time of interview. Cancer cases were identified through a rapid case ascertainment system, supplemented by the population-based Shanghai Cancer Registry. A total of 1602 eligible breast cancer cases were identified during the study period, of which 1459 cases (91.1%) completed in-person interviews. Cancer diagnoses for all

Received 10/3/03; revised 12/18/03; accepted 1/16/04.

Grant support: USPHS Grants RO1CA64277 and RO1CA90899 from the National Cancer Institute.

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, Center for Health Services Research, 1215 21st Avenue South, Medical Center East, Suite 6000, Vanderbilt University Medical Center, Nashville, TN 37232-8300. Phone: (615) 936-0682; Fax: (615) 936-1269. E-mail: wei.zheng@vanderbilt.edu

Downloaded from cebp.aacrjournals.org on October 20, 2021. © 2004 American Association for Cancer Research.

patients were reviewed and confirmed by two senior pathologists. Controls were randomly selected from the general population of Shanghai using the Shanghai Resident Registry, a population registry containing demographic information for all residents of urban Shanghai, and were frequency matched on age (5-year intervals) to the expected age distribution of the case subjects in a 1:1 ratio. The inclusion criteria for controls were identical to those for cases, with the exception of a breast cancer diagnosis. Of the 1724 eligible women, 1556 (90.3%) completed in-person interviews. The major reason for non-participation was refusal, accounting for 6.8% (n = 109) for cases and 9.6% (n = 166) of controls identified for the study.

A structured questionnaire was used to elicit detailed information on demographic factors, menstrual and reproductive history, hormone use, dietary habits, prior disease history, physical activity, tobacco and alcohol use, weight, and family history of cancer. All participants were measured for their current weight and circumferences of waist and hips. Blood samples (10 ml from each woman) were obtained from 1193 (82%) cases and 1310 (84%) controls who completed the in-person interviews. These samples were processed on the same day, typically within 6 hours after sample collection, and stored at −70°C until relevant bioassays.

Genomic DNA was extracted from buffy coat fractions. Genotyping for the *CYP11A* (*TAAAA*)n polymorphism was performed by detecting fluorescent amplimers on an ABI PRISM 3700 automated DNA Analyzer. The primers were: F: 5′-*GAGCTATCTTGCCAGCTTG*-3′ and R: 5′-*gtgtCTCTGAGTCAGCTGTACTG*-3′. The reverse primer was designed using a tailing strategy (presented in lower case) to promote full non-templated nucleotide addition by AmpliTaq Gold DNA polymerase (ABI, Foster City, CA), which provided unambiguous detection of alleles separated by one base pair (15). The forward primer was labeled with FAM. Each 2.2 μl of PCR mixture included 0.1 unit AmpliTaq Gold DNA polymerase, 1× Buffer II, 2.5 mM $MgCl_2$, 0.25 mM dNTPs, 335 nM of each primer, and 1 ng DNA. Thermal cycling conditions were as follows: 95°C×10 min followed by 10 cycles of 94°C×15 s, 55°C × 15 s, and 72°C × 30 s, 20 cycles of 89°C × 15 s, 55°C × 15 s, and 72°C × 30 s with a final extension step of 72°C × 10 min. Allele fragment size estimation was accomplished using the internal size standard Genescan 400HD ROX and the Local Southern algorithm of GENESCAN software. Editing of alleles was performed in GENOTYPER. Allele binning and adjustment of run mobility according to the control alleles of CEPH 1347-02 were accomplished by custom software. The number of repeats within each allele was confirmed by direct sequencing using BigDye Terminator Chemistry on an ABI PRISM 3700 automated DNA Analyzer.

Each 96-well plate of genomic DNA contained multiple controls, including a water blank, two samples of CEPH 1347-02, two known study control duplicates, and two blinded study control duplicates. Duplicates were distributed across 96-well plates. Samples within a 96-well plate were arranged such that laser attenuation across the capillary array of a 3700 DNA Analyzer would not result in lower average signal among case or control groups. Genotyping data were obtained from 1015 (85.1%) cases and 1082 (82.6%) controls who provided a blood sample to the study. The major reasons for incomplete genotyping were insufficient DNA or unsuccessful PCR amplification. Of the 3918 alleles determined in duplicated samples, only 85 alleles were inconsistent, yielding a precision rate of 97.83%.

Fasting blood samples were measured for plasma levels of testosterone, estradiol, estrone, estrone sulfate, and DHEA-S for 115 postmenopausal controls, as part of an ancillary case-control study (3). Measurement of steroids was performed in a reference laboratory at Diagnosed Systems Laboratory, Inc. (DSL, Webster, TX) using commercial RIA kits. This laboratory specializes in *in vitro* diagnostic testing of endocrine markers and is certified by Clinical Laboratory Improvement Amendments (CLIA) and the International Standard ISO 9002. Each sample was measured in duplicate. The coefficients of variation (CV) for the intra- and interassay variability were mostly below 10%.

Chi-squared statistics were used to evaluate case-control differences in the distribution of genotypes. To accommodate the age frequency-matched study design, we used logistic regression models conditioned on age to estimate odds ratios (ORs) and their 95% confidence intervals (95% CIs) as a measure of the strength of the association between gene polymorphism and breast cancer risk (16, 17). Analyses stratified by menopausal status were conducted to evaluate whether the association differed for pre- and postmenopausal women. Analyses were also performed to evaluate the potential modifying effect of the *CYP11A* genotypes on the associations between breast cancer risk and surrogate measures of endogenous sex hormone exposure, such as years of menstruation, body mass index, and waist-to-hip ratio. The relationship between *CYP11A* genotypes and blood sex hormone levels was evaluated using data from postmenopausal controls. Because the distributions of blood sex hormone levels mostly skew to the high value, geometric means were estimated and compared using t tests between two genotype groups. P values of less than 0.05 (two-sided probability) were interpreted as statistically significant.

Results

Table 1 compares breast cancer patients and controls with respect to selected demographic characteristics and major risk factors for breast cancer. These data were presented separately for subjects included in the whole study and those with genotyping data. The latter group was very similar to the former with regard to virtually all variables compared, although small P values were noted for the mean comparisons of age and WHR in the two case groups. For subjects included in the *CYP11A* genotype study, cases and controls were comparable in age. With the exception of a family history of breast cancer, statistically significant associations were observed for education level and all major risk factors of breast cancer. More cases than controls had a family history of breast cancer, although the difference was not statistically significant, due to a low number of subjects with a positive family history. Detailed results for breast cancer risk factors by menopausal status in this study have been published elsewhere (18–20). Briefly, with the exception of

Downloaded from cebp.aacrjournals.org on October 20, 2021. © 2004 American Association for Cancer Research.

Table 1. Comparisons of cases and controls by selected demographic characteristics and major risk factors, the Shanghai Breast Cancer Study

Characteritics[§]	Subjects in whole study		Subjects with genotyping data		*P* values for differences[‡]		
	Cases (n = 1459)	Controls (n = 1556)	Cases (n = 1015)	Controls (n = 1082)	Cases	Controls	Cases *versus* Controls
	(1)	(2)	(3)	(4)	(3 *versus* 1)	(4 *versus* 2)	(3 *versus* 4)
Demographic factors							
Age (years)	47.93 ± 7.99	47.25 ± 8.79	47.51 ± 7.95	47.13 ± 8.75	0.045	0.607	0.294
Education, lower than middle school (%)	12.13	13.88	11.63	14.70	0.656	0.466	0.038
Major risk factors							
Breast cancer in first-degree relatives (%)	3.70	2.44	3.35	2.40	0.610	0.988	0.194
Ever had breast fibroadenoma (%)	9.60	5.01	9.86	5.08	0.828	0.971	<0.001
Age at menarche	14.48 ± 1.62	14.70 ± 1.71	14.48 ± 11.62	14.74 ± 1.73	0.937	0.388	<0.001
Age at first live birth[*]	26.77 ± 4.15	26.21 ± 3.86	26.87 ± 4.09	26.17 ± 3.88	0.389	0.741	<0.001
Age at menopause[†]	48.13 ± 4.58	47.45 ± 4.89	48.18 ± 4.64	47.46 ± 4.94	0.796	0.981	0.047
Physically active in past 10 years (%)	18.72	25.21	19.43	26.09	0.593	0.529	<0.001
Body mass index	23.53 ± 3.40	23.15 ± 3.38	23.48 ± 3.35	23.14 ± 3.38	0.569	0.952	0.022
Waist-to-hip ratio	0.81 ± 0.06	0.80 ± 0.06	0.81 ± 0.06	0.80 ± 0.06	0.008	0.244	0.005

Note: [‡]From χ^2 test (categorical variables) or t test (continuous variables).
[*]Among parous women.
[†]Among postmenopausal women.
[§]Unless otherwise noticed, $\bar{x}$ ± SD are presented.

body mass index, all other risk factors listed in Table 1 were associated with the risk of both pre- and postmenopausal breast cancer. Body mass index was primarily associated with breast cancer risk in postmenopausal women.

The allele distribution by cases and controls is presented in Table 2. Three common alleles with 4, 6, or 8 (*TAAAA*) repeats were identified, accounting for nearly 99% of total alleles detected in the study population. The genotype distributions of these alleles was consistent with the Hardy-Weinberg equilibrium in controls (P = 0.34). Cases and controls differed in the frequency of the 6-repeat allele (62.86% in cases V.S. 65.9% in controls, P = 0.04) or the 8-repeat allele (12.56% in cases V.S. 8.46% in controls, P < 0.0001). ORs for specific allele combinations are presented in Table 3. When compared to subjects homozygous for the 4-repeat allele, women carrying the 8-repeat allele had an elevated risk of breast cancer (OR = 1.8, 95% CI = 1.2–2.7) and the risk was further increased among those homozygous for this risk allele (OR = 3.4, 95% CI = 1.4–8.4). All genotypes containing the 8-repeat allele were associated with an elevated risk of breast cancer. Therefore, subsequent analyses were focused on evaluating the association between the 8-repeat allele and breast cancer. When compared to subjects who did not carry the 8-repeat allele, the ORs were 1.5 (95% CI = 1.2–1.9) and 2.9 (95%

Table 2. Allele frequency of *CYP11A* (*TAAAA*)n polymorphism in cases and controls, the Shanghai Breast Cancer Study

Number of *TAAAA* repeats	Length of PCR product (bp)	Cases		Controls		*P* value
		N	%	*N*	%	
Four	260	477	23.50	524	24.21	0.586
Six	270	1276	62.86	1426	65.90	0.040
Seven	275	3	0.15	8	0.37	0.160
Eight	280	255	12.56	183	8.46	<0.0001
Nine	285	7	0.34	11	0.51	0.418
Ten	290	9	0.44	7	0.32	0.529
Others	Others	3	0.15	5	0.23	0.537
No. of chromosomes		2030	100.00	2164	100.00	
		χ^2 = 22.02, df = 6, P = 0.001				

Table 3. Association of *CYP11A* (*TAAAA*)n polymorphism with breast cancer risk, the Shanghai Breast Cancer Study

CYP11A genotypes	Cases		Controls		OR (95% CI)[*]
	No.	%	No.	%	
4 repeat/4 repeat	57	5.62	74	6.84	1.0 (reference)
4 repeat/6 repeat	302	29.75	325	30.04	1.3 (0.9–1.8)
6 repeat/6 repeat	403	39.70	481	44.45	1.1 (0.8–1.7)
8 repeat/any	235	23.15	175	16.17	1.8 (1.2–2.7)
8 repeat/4 repeat	56	5.52	45	4.16	1.6 (1.0–2.8)
8 repeat/6 repeat	156	15.37	120	11.09	1.8 (1.2–2.7)
8 repeat/8 repeat	20	1.97	8	0.74	3.4 (1.4–8.4)
8 repeat/others	3	0.30	2	0.18	1.9 (0.3–11.7)
Others	18	1.77	27	2.50	1.0 (0.5–1.9)
Presence of the 8 repeat allele					
No	780	76.85	907	83.83	1.0 (reference)
Yes	235	23.15	175	16.17	1.6 (1.3–1.9)
One copy	215	21.18	167	15.43	1.5 (1.2–1.9)
Two copies	20	1.97	8	0.74	2.9 (1.3–6.7)
Trend test					P < 0.0001

[*]Adjusted for age and education.

Downloaded from cebp.aacrjournals.org on October 20, 2021. © 2004 American Association for Cancer Research.

CI = 1.3–6.7), respectively, for those who carried one and two copies of this allele (P for trend, <0.0001). Additional adjusting for all known risk factors presented in Table 1 did not appreciably change the risk estimate.

The positive association between *CYP11A* genotypes and breast cancer risk was observed in both pre- and postmenopausal women (Table 4) and tests for gene-dose effect were statistically significant. To evaluate the potential modifying effects of the *CYP11A* genotypes on the associations between breast cancer risk and surrogate measures of endogenous sex hormone exposure, such as years of menstruation, body mass index, waist-to-hip ratio, additional analyses were performed by joint distribution of *CYP11A* genotypes and these variables (Table 5). Although there was some variation in the point estimate and none of the tests for multiplicative interaction were statistically significant, an elevated risk of breast cancer was associated with the 8-repeat allele in all strata defined by these major breast cancer risk factors. In particular, there was a nearly 9-fold elevated risk of breast cancer among women who had a long duration of menstruation and carried two copies of the 8 repeat allele.

Table 6 compares the geometric means of major steroid sex hormones measured in blood by *CYP11A* genotypes. This analysis was performed only among postmenopausal controls. Because only three women were homozygous for the 8-repeat allele, they were combined with the heterozygous group. Overall, there was no statistical difference between women with and without the 8-repeat allele. The only marginally significant difference was for blood estradiol, for which a very low level was detected in postmenopausal women.

Discussion

Over the last 10 years, a large number of epidemiological studies have been conducted to investigate the association between breast cancer risk and genetic polymorphisms in the genes involved in estrogen biosynthesis and metabolism (21). To our knowledge, none of these studies, however, has evaluated P450scc (encoded by the *CYP11A* gene), one of the key enzymes in steroid hormone synthesis. In this large population-based case-control study, we found that a pentanucleotide repeat [(*TAAAA*)n] polymorphism in the promoter region of the *CYP11A* gene was associated with an increased risk of breast cancer. Women who carried the 8-repeat allele had a greater than 50% elevated risk of breast cancer, while the risk was increased nearly 3-fold in those who were homozygous for this allele. This finding is new and consistent with the important role of P450scc in the biosynthesis of steroid sex hormones.

Table 4. Associations of *CYP11A* genotypes and breast cancer risk, stratified by menopausal status, in the Shanghai Breast Cancer Study

Presense of the 8 (TAAAA) repeat allele	Cases		Controls		OR[†]
	No.	%	No.	%	
Premenopausal women					
None	521	75.95	576	83.97	1.0 (reference)[*]
One copy	152	22.16	105	15.31	1.7 (1.3–2.3)
Two copies	13	1.90	5	0.73	2.6 (0.9–7.5)
Trend test					$P < 0.001$
Postmenopausal women					
None	254	78.40	329	83.90	1.0 (reference)[*]
One copy	63	19.44	60	15.31	1.5 (1.0–2.2)
Two copies	7	2.16	3	0.77	3.4 (0.8–14.1)
Trend test					$P = 0.018$

Note: [†]Adjusted for age and education.
[*]Reference group.

The conversion of cholesterol into pregnenolone is the common pathway leading to the production of progesterone, androgens, estrogen, and other steroid hormones (1). This rate-limiting step is carried out in mitochondria and includes three distinct sequential reactions, all catalyzed by P450scc (1). Studies from rabbits with a spontaneous *CYP11A* gene deletion show that P450scc is the only enzyme that can convert cholesterol to pregnenolone (22). A linkage analysis showed that the *CYP11A* gene may be the major susceptibility locus for PCOS (4), an inherited endocrine disorder characterized by hyperandrogenemia (10). In a sibling-pair analysis, the *CYP11A* gene was again implicated for PCOS, although the test was no longer statistically significant after correction for multiple comparisons (23). In human ovarian cell cultures, thecal cells from women with polycystic ovaries produce approximately 20-fold more androstenedione than do cells from women with normal ovaries (24). Intriguingly, thecal cells from PCOS patients also express a substantially higher level of *CYP11A* mRNA than those from controls (11), indicating that increased transcription of the gene is responsible for the excess androgen production observed in the PCOS patients.

The promoter region of the *CYP11A* gene from a number of different species contains conserved transcription factor binding sites within 200 bp upstream of the initiator methionine that play a role in regulating *CYP11A* gene expression (1, 4–7). However, the sequence in the 5′ regions is not well conserved across the well-studied mouse, bovine, and ovine *CYP11A* genes. The [(*TAAAA*)n] repeat is only observed in the human orthologue. By sequencing the *CYP11A* promoter in a 1.85-kb region 5′ to the translation start site, Gharani *et al*. (4) identified a highly polymorphic pentanucleotide repeat [(*TAAAA*)n] polymorphism. Four alleles with 4, 6, 8, and 9 *TAAAA* repeats were found, and the 4-repeat allele accounted for nearly 60% of the total identified in the study population. Because of the small sample size, alleles with 6 repeats or longer were combined into one group in the data analysis, and this group of high-order alleles was associated with an elevated risk of PCOS and hirsutism, as well as an increased level of serum testosterone among women with clinical evidence of hyperandrogenism (4). Similar results were reported from three subsequent studies (8–10).

On the other hand, no association was found between *CYP11A* genotypes and blood sex hormone levels in two studies (25, 26), including a recent study conducted among normal pre-menopausal women (25). Because of small sample sizes, most previous studies have combined all alleles with 6 repeats or longer into one group in the data analysis. In our study, conducted among Chinese women, we found that the 6-repeat allele had, in fact, a slightly lower frequency in breast cancer cases

Downloaded from cebp.aacrjournals.org on October 20, 2021. © 2004 American Association for Cancer Research.

Table 5. Associations of breast cancer risk with join distributions of *CYP11A* genotypes and selected surrogate measures of endogenous sex hormone exposure, the Shanghai Breast Cancer Study

Stratified variables (by median)	Presence of the 8 (*TAAAA*) repeat allele						*P* value for trend test
	No		One copy		Two copies		
	Cases/ Controls	OR (95% CI)	Cases/ Controls	OR (95% CI)*	Cases/ Controls	OR (95% CI)	
Years of menstruation							
<27	295/438	1.0 (reference)†	87/86	1.5 (1.1–2.1)	5/5	1.5 (0.4–5.2)	0.024
≥27	480/467	1.8 (1.4–2.3)	128/79	2.8 (2.0–4.1)	15/3	8.9 (2.5–31.4)	<0.001
					P for interaction = 0.381		
Body mass index							
<22.75	344/448	1.0 (reference)†	101/90	1.5 (1.1–2.1)	7/3	3.3 (0.8–13.0)	0.004
≥22.75	436/459	1.3 (1.0–1.6)	114/77	1.9 (1.4–2.7)	13/5	3.4 (1.2–9.7)	0.002
					P for interaction = 0.964		
Waist-to-hip ratio							
<0.798	351/463	1.0 (reference)†	93/83	1.5 (1.1–2.0)	10/2	7.3 (1.6/33.5)	<0.001
≥0.798	429/444	1.3 (1.1–1.6)	122/54	2.0 (1.4–2.7)	10/6	2.2 (0.8–6.0)	0.007
					P for interaction = 0.251		

*Adjusted for age and education.
†Reference group.

than controls, and only the 8-repeat allele was associated with an elevated risk of breast cancer. Interestingly, in the study by Daneshmand *et al*. (11), the only previous study that evaluated each major allele individually, less than 6% of subjects carried the 8-repeat allele, and only the 9-repeat allele was overrepresented in PCOS patients. Although that study did not find a correlation of blood sex hormones with the 9 repeat allele, *CYP11A* mRNA expression appeared to increase with the copy number of the 9-repeat allele in controls. We did not find any apparent difference in blood sex hormone levels between women with and without the 8-repeat allele among postmenopausal healthy women. Because of the small sample for this component of the study, we could not evaluate the association of blood sex hormone levels with homozygosity for the 8-repeat allele. Furthermore, we could not rule out a possible association of *CYP11A* genotypes with blood sex hormone levels in premenopausal women, an exposure that might be more relevant to breast cancer risk.

Because several previous studies did not present data for the frequencies of all common alleles identified for the (*TAAAA*)n polymorphism in the *CYP11A* gene, we were unable to compare the allele frequencies of this polymorphism across various ethnic groups. Several previous studies, conducted mostly in Caucasian women, reported a higher frequency of the 4-repeat allele (4, 25, 26) than our study. This appears contradictory to the observation that Caucasians have a higher risk of breast cancer than their Chinese counterparts. However, given the multifactorial nature of breast cancer etiology, it would be prudent not to make an etiological inference from such an ecological comparison of the frequencies of a single risk allele across various ethnic groups.

No study has directly evaluated how these short-tandem repeats affect the expression of the *CYP11A* gene. Similar (*TAAAA*)n repeat polymorphisms have been found to be associated with increased transcriptional activity and increased plasma levels of apolipoprotein(a) (27). Furthermore, as suggested for the insulin gene, variable number tandem repeat polymorphisms may play an important role in the regulation of transcriptional activity (28). We identified 11,618 [(*TAAAA*)n] short tandem repeat polymorphisms within the current draft of the human genome sequence (April 2003 release), and 86 within 1 kb 5′ of the transcription start site of known genes, including the *SULT1A1* and *GSTP1* genes that have been linked to breast cancer risk (21, 29). Recently, Calvo performed a heteroduplex analysis of exons and intron-exon boundaries of the *CYP11A* gene and identified a single missense variant that was presented in one of the 29 hirsute patients and none of the 50 controls (30). It is possible that other polymorphisms in linkage disequilibrium with the [(*TAAAA*)n] marker,

Table 6. Geometric means (95% CI) of blood steroid sex hormones by *CYP11A* genotypes among postmenopausal controls, the Shanghai Breast Cancer Study

Blood hormones	Presence of the 8 (*TAAAA*) allele		*P* value from *t* tests
	No (*N* = 92)	Yes (*n* = 23)	
Testosterone (pg/ml)	149.9 (130.3–174.2)	141.2 (107.8–183.1)	0.67
Estradiol (pg/ml)	8.0 (6.8–9.5)	5.9 (4.4–7.1)	0.06
Estrone (pg/ml)	16.0 (13.9–18.4)	13.1 (9.7–17.6)	0.22
Estrone-S (pg/ml)	1261.4 (897.9–1772.2)	1299.8 (735.1–2298.5)	0.93
DHEA-S (ng/ml)	566.8 (478.2–665.1)	533.8 (437.0–645.5)	0.66

Downloaded from cebp.aacrjournals.org on October 20, 2021. © 2004 American Association for Cancer Research.

in either the *CYP11A* gene or a neighboring gene, may be responsible for the association observed in the study. In such a case, the degree of disequilibrium could also affect the assessment of ORs. Nonetheless, we present here the identification of the [(*TAAAA*)n] repeat in the *CYP11A* gene as a potential marker of elevated risk for breast cancer that warrants further investigation in future studies.

Acknowledgments

The authors thank Drs. Qi Dai, Fan Jin, and Jia-Rong Cheng for their contributions in coordinating data and specimen collection in Shanghai; Dr. Herbert Yu for his contribution in measuring blood sex hormones; and Bethanie Hull for technical assistance in the preparation of this manuscript. This study would not have been possible without the support of all of the study participants and research staff of the Shanghai Breast Cancer Study.

References

1. Norman AW, Litwack G, editors. Hormones, 2nd ed. San Diego: Academic Press; 1997.
2. Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women. Reanalysis of nine prospective studies. J Natl Cancer Inst 2002;94:606-16.
3. Yu H, Shu XO, Shi R, et al. Plasma sex steroid hormones and breast cancer risk in Chinese women. Int J Cancer 2003;105:92-7.
4. Gharani N, Waterworth DM, Batty S, et al. Association of the steroid synthesis gene CYP11a with polycystic ovary syndrome and hyperandrogenism. Hum Mol Genet 1997;6:397-402.
5. Pestell RG, Hammond VE, Crawford RJ. Molecular cloning and characterization of the cyclic AMP-responsive ovine CYP11A1 (cholesterol side-chain cleavage) gene promoter: DNase 1 protection of conserved consensus elements. J Mol Endocrinol 1993;10:297-311.
6. Guo IC, Chung BC. Cell-type specificity of human CYP11A1 TATA box. J Steroid Biochem Mol Biol 1999;69:329-34.
7. Huang Y, Hu M, Hsu N, Wang CL, Chung B. Action of hormone responsive sequence in 2.3 kb promoter of CYP11A1. Mol Cell Endocrinol 2001;175:205-10.
8. Pugeat M, Cousin P, Nicolas MH, et al. Pentanucleotide repeat polymorphism in the promoter region of the human CYP11a gene and polycystic ovaries in hirsute women. Proceeding of the Endocrine Society, 81st Annual Meeting, San Diego, California. Abstract OR29-1, 1999.
9. Diamanti-Kandarakis E, Bartzis MI, Bergiele AT, Tsianateli TC, Kouli CR. Microsatellite polymorphism (tttta)(n) at −528 base pairs of gene CYP11aα influences hyperandrogenemia in patients with polycystic ovary syndrome. Fertil Steril 2000;73:735-41.
10. Franks S, Gharani N, McCarthy M. Candidate genes in polycystic ovary syndrome. Hum Reprod Updat 2001;7:405-10.
11. Daneshmand S, Weitsman SR, Navab A, Jakimiuk AJ, Magoffin DA. Overexpression of theca-cell messenger RNA in polycystic ovary syndrome does not correlate with polymorphisms in the cholesterol side-chain cleavage and 17α-hydroxylase/C(17-20) lyase promoters. Fertil Steril 2002;77:274-80.
12. Balen A. Polycystic ovary syndrome and cancer. Rev Hum Reprod Updat 2001;7:522-5.
13. 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.
14. Zheng W, Xie DW, Jin F, et al. Genetic polymorphism of cytochrome P450-1B1 and risk of breast cancer. Cancer Epidemiol Biomark Prev 2000;9:147-50.
15. Brownstein MJ, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 1996;20:1004-10.
16. Breslow NE, Day NE, editors. Statistical methods in cancer research. Vol. 1. The analysis of case-control studies. Lyon, France: International Agency for Research on Cancer; 1980.
17. Rothman KJ, Greenland S. Modern Epidemiology. 2nd ed. Philadelphia, PA: Lippincott-Williams & Wilkins; 1998.
18. 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.
19. Shu XO, Jin F, Dai Q, et al. Association of body size and fat distribution with risk of breast cancer among Chinese women. Int J Cancer 2001;94:449-55.
20. Matthews CE, Shu XO, Jin F, et al. Lifetime physical activity and breast cancer risk in the Shanghai Breast Cancer Study. Br J Cancer 2001;84:994-1001.
21. Mitrunen K, Hirvonen A. Molecular epidemiology of sporadic breast cancer. The role of polymorphic genes involved in oestrogen biosynthesis and metabolism. Mutat Res 2003;544:9-41.
22. Yang X, Iwamoto K, Wang M, Artwohl J, Mason JI, Pang S. Inherited congenital adrenal hyperplasia in the rabbit is caused by a deletion in the gene encoding cytochrome P450 cholesterol side-chain cleavage enzyme. Endocrinology 1993;132:1977-82.
23. Urbanek M, Legro RS, Driscoll DA, et al. Thirty-seven candidate genes for polycystic ovary syndrome: strongest evidence for linkage is with follistatin. Proc Natl Acad Sci USA 1999;96:8573-8.
24. Gilling-Smith C, Willis DS, Beard RW, Franks S. Hypersecretion of androstenedione by isolated thecal cells from polycystic ovaries. J Clin Endocrinol Metab 1994;79:1158-65.
25. Garcia-Closas M, Herbstman J, Schiffman M, Glass A, Dorgan JF. Relationship between serum hormone concentrations, reproductive history, alcohol consumption and genetic polymorphisms in premenopausal women. Int J Cancer 2002;102:172-8.
26. San Millan JL, Sancho J, Calvo RM, Escobar-Morreale HF. Role of the pentanucleotide (tttta)(n) polymorphism in the promoter of the CYP11a gene in the pathogenesis of hirsutism. Fertil Steril 2001;75:797-802.
27. Mooser V, Mancini FP, Bopp S, et al. Sequence polymorphisms in the apo(a) gene associated with specific levels of Lp(a) in plasma. Hum Mol Genet 1995;4:173-81.
28. Waterworth DM, Bennett ST, Gharani N, et al. Linkage and association of insulin gene VNTR regulatory polymorphism with polycystic ovary syndrome. Lancet 1997;349:986-90.
29. Zheng W, Xie D, Cerhan JR, Sellers TA, Wen WQ, Folsom AR. Sulfotransferase 1A1 polymorphism, endogenous estrogen exposure, well-done meat intake and breast cancer risk. Cancer Epidemiol Biomark Prev 2001;10:89-94.
30. Calvo RM, Asuncion M, Telleria D, Sancho J, San Millan JL, Escobar-Morreale HF. Screening for mutations in the steroidogenic acute regulatory protein and steroidogenic factor-1 genes, and in CYP11A and dosage-sensitive sex reversal-adrenal hypoplasia gene on the X chromosome, gene-1 (DAX-1), in hyperandrogenic hirsute women. J Clin Endocrinol Metab 2001;86:1746-9.

Downloaded from cebp.aacrjournals.org on October 20, 2021. © 2004 American Association for Cancer Research.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Population-Based Case-Control Study of *CYP11A* Gene Polymorphism and Breast Cancer Risk

Wei Zheng, Yu-Tang Gao, Xiao-Ou Shu, et al.

Cancer Epidemiol Biomarkers Prev 2004;13:709-714.

Updated version Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/13/5/709

Cited articles This article cites 25 articles, 4 of which you can access for free at:
http://cebp.aacrjournals.org/content/13/5/709.full#ref-list-1

Citing articles This article has been cited by 8 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/13/5/709.full#related-urls

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link
http://cebp.aacrjournals.org/content/13/5/709.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC)
Rightslink site.

Downloaded from cebp.aacrjournals.org on October 20, 2021. © 2004 American Association for Cancer Research.