American Journal of Epidemiology
Vol. 166, No. 12
DOI: 10.1093/aje/kwm242
Advance Access publication September 7, 2007

Original Contribution

Interaction of Soy Food and Tea Consumption with *CYP19A1* Genetic Polymorphisms in the Development of Endometrial Cancer

Wang Hong Xu[1], Qi Dai[2], Yong Bing Xiang[1], Ji Rong Long[2], Zhi Xian Ruan[1], Jia Rong Cheng[1], Wei Zheng[2], and Xiao Ou Shu[2]

[1] Department of Epidemiology, Shanghai Cancer Institute and Cancer Institute of Shanghai Jiao Tong University, Shanghai, People's Republic of China.
[2] Vanderbilt Epidemiology Center, Vanderbilt Institute of Medicine and Public Health, Department of Medicine, Vanderbilt University Medical Center and Vanderbilt-Ingram Cancer Center, Nashville, TN.

Received for publication April 5, 2007; accepted for publication July 27, 2007.

Certain polyphenols inhibit the activity of aromatase, a critical enzyme in estrogen synthesis that is coded by the *CYP19A1* gene. Consumption of polyphenol-rich foods and beverages, thus, may interact with *CYP19A1* genetic polymorphisms in the development of endometrial cancer. The authors tested this hypothesis in the Shanghai Endometrial Cancer Study (1997–2003), a population-based case-control study of 1,204 endometrial cancer cases and 1,212 controls. Dietary information was obtained by use of a validated food frequency questionnaire. Genotypes of *CYP19A1* at rs28566535, rs1065779, rs752760, rs700519, and rs1870050 were available for 1,042 cases and 1,035 controls. Unconditional logistic regression models were used to calculate odds ratios and their 95% confidence intervals after adjustment for potential confounding factors. Higher intake of soy foods and tea consumption were both inversely associated with the risk of endometrial cancer, with odds ratios of 0.8 (95% confidence interval: 0.6, 1.0) for the highest versus the lowest tertiles of intake of soy and 0.8 (95% confidence interval: 06, 0.9) for ever tea consumption. The association of single nucleotide polymorphisms rs1065779, rs752760, and rs1870050 with endometrial cancer was modified by tea consumption ($p_{interaction} < 0.05$) but not by soy isoflavone intake. The authors' findings suggest that tea polyphenols may modify the effect of *CYP19A1* genetic polymorphisms on the development of endometrial cancer.

aromatase; endometrial neoplasms; polymorphism, genetic; soy foods; tea

Abbreviations: QC, quality control; SECS, Shanghai Endometrial Cancer Study; SNP, single nucleotide polymorphism.

The incidence rate of endometrial cancer in Asian countries, such as China, is substantially lower than that in Western countries, and incidence of the disease has been found to increase when Asian women emigrate to the United States (1). One possible explanation is that certain lifestyle factors, particularly diets commonly consumed in Asian countries that include soy and tea, foods and beverages rich in certain polyphenols, may confer some protection against endometrial cancer (2, 3).

Animal and cell culture studies have suggested that dietary polyphenols may have a preventive effect on cancers (4, 5). Furthermore, certain polyphenols, such as isoflavones from soy foods, are classified as phytoestrogens because they have a structure similar to steroid hormones and weak estrogenic or antiestrogenic activity. Evidence from many in vitro and in vivo observations suggests that phytoestrogens may reduce the risk of hormone-related cancers (2, 6), possibly by interfering with the synthesis, metabolism, and signal transduction of steroid hormones (2, 7, 8). Epidemiologic studies have suggested that polyphenols from soy foods and tea may reduce the risk of some cancers (9–15), including endometrial cancer (12–15).

Correspondence to Dr. Xiao Ou Shu, Vanderbilt Epidemiology Center, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 600, Nashville, TN 37203-1738 (e-mail: Xiao-ou.shu@vanderbilt.edu).

Recent studies have found that tea polyphenols, as well as isoflavones, moderately inhibit the activity of aromatase (CYP19A1) (16, 17), the crucial enzyme converting androstenedione and testosterone to estrone and estradiol, respectively. Genetic variations in the *CYP19A1* gene have been shown to alter aromatase activity and affect hormone levels (18, 19), possibly influencing the pathogenesis of endometrial cancer (20–23). We hypothesized that isoflavones and tea polyphenols may interact with polymorphisms in the *CYP19A1* gene in the development of endometrial cancer and tested this hypothesis in the Shanghai Endometrial Cancer Study (SECS).

MATERIALS AND METHODS

The SECS is a population-based case-control study of 1,204 incident cases diagnosed between the ages of 30 and 69 years from 1997 to 2003 and 1,212 age frequency-matched controls. Cancer cases were identified through the population-based Shanghai Cancer Registry. A total of 1,454 eligible endometrial cancer cases were identified during the study period, and 1,204 (82.8 percent) completed an in-person interview. Cases were confirmed by medical chart review. The median interval between diagnosis and interview for cases was 5.6 months. Controls were randomly selected from the general population of Shanghai by use of the Shanghai Resident Registry and matched to cases according to the age distribution (in 5-year intervals) of endometrial cancer cases in 1996. Women with a history of cancer or hysterectomy were not eligible. Of the 1,629 eligible women identified, 1,212 (74.4 percent) participated in the study. The study protocols were approved by the institutional review boards of all institutes involved in the study, and written informed consent was obtained from all subjects.

Study participants were interviewed in person by trained interviewers. Weight, height, and circumferences of the waist and hips were measured according to a standardized protocol at the time of interview. A structured questionnaire was used to elicit detailed information on demographic factors, menstrual and reproductive history, hormone use, prior disease history, physical activity, tobacco and alcohol use, weight history, and family history of cancer. Regular smokers were defined as women who had ever smoked at least one cigarette per day for 6 months or more, while alcohol drinkers were defined as women who had ever drunk alcoholic beverages at least three times a week for at least 6 months. Similarly, tea drinkers were defined as women drinking tea at least three times per week for 6 months or longer. Tea drinkers were also asked to provide the age at which they started to drink tea regularly, the total number of years they drank tea, and the usual frequency, amount, and major types of tea consumed. Participants were asked if they had engaged in regular exercise/sports (at least once a week for at least 3 months) during the preceding 5 years. Usual dietary habits over the preceding 5 years were assessed by use of a validated, quantitative food frequency questionnaire, which included 71 food items and covered more than 85 percent of the commonly consumed foods in Shanghai (24). During the interview, each participant was first asked how frequently she consumed a specific food or group of foods (per day, week, month, year, or never), followed by a question on how many liang (50 g) were consumed per unit of time (day, week, month, or year) during the preceding 5 years, ignoring any recent dietary changes. For seasonal foods, each participant was asked to describe her consumption during the month(s) when the food was available. The average daily intake of each seasonal item was estimated by calculating the percentage of months that the food was on the market over a 1-year period. Total soy food intake was measured by soy protein intake and was estimated by multiplying the amount of soy food consumed with the amount of protein in that food according to the Chinese food composition tables (25).

Of the study participants, 857 cases and 837 controls donated a blood sample, and 282 cases and 286 controls provided a buccal cell sample. Among those who provided a buccal cell sample, 189 cases and 198 controls provided samples using a mouthwash method, and 93 cases and 88 controls provided samples using a buccal swab method. Because of the very low DNA yield of the buccal swab method, we did not include buccal swab DNA samples in the genotyping. In addition, DNA from blood samples donated by 19 control subjects was not available because of their use in previous studies. Thus, DNA samples from 1,046 (857 blood and 189 buccal cell) cases (86.9 percent) and 1,035 (837 blood and 198 buccal cell) controls (85.4 percent) were included in the genotyping study. Genotyping success rates were between 98.1 percent and 99.6 percent.

The samples were processed on the same day as collection, typically within 6 hours, at the Shanghai Cancer Institute. The buffy coat (white blood cell) and the buccal cell pellet samples were stored at −70°C. Genomic DNA was extracted from buffy coat fractions or buccal cells by using a QIAmp DNA mini kit (Qiagen, Inc., Valencia, California) following the manufacturer's protocol. The allelic discrimination of the *CYP19A1* polymorphisms at rs28566535 (originally named hcv1664178), rs1065779, rs752760, and rs1870050 was assessed with the ABI PRISM 7900 Sequence Detection System (Applied Biosystems, Inc., Foster City, California) using TaqMan Assay-on-Demand kits also obtained from Applied Biosystems, Inc. The *CYP19A1* rs700519 polymorphism was genotyped using an MGB Eclipse (3′ hybridization-triggered fluorescence reaction) assay (Epoch Biosciences, Bothell, Washington). (A detailed lab protocol has been described elsewhere (23).)

The laboratory staff was blind to the identity of the subjects. Quality control (QC) samples were included in the genotyping assays. Each 384-well plate contained four water, eight Centre d'Etude du Polymorphisme Humain (CEPH) (http://www.cephb.fr/cephdb/) 1347-02 DNA, eight blinded QC DNA, and eight unblinded QC DNA samples. The concordance rates for the quality control samples were 97.4 percent for rs752760 and rs1870050, 98.7 percent for rs700519, and 100 percent for rs1065779 and rs28566535. In addition, the DNA of Chinese samples that were used in the HapMap ($n = 45$) and Perlegen ($n = 24$) projects was purchased from Coriell Cell Repositories (Coriell Institute, Inc., Camden, New Jersey) and genotyped for single nucleotide polymorphisms (SNPs) rs752760, rs1870050, and rs700519. The consistency rate was 100.0 percent for all

TABLE 1. Comparison of cases and controls by selected demographic factors, major risk factors, and *CYP19A1* genotypes, Shanghai Endometrial Cancer Study, 1997–2003

	Subjects with genotyping data		*p* value*
	Cases ($n = 1,042$)	Controls ($n = 1,035$)	
Demographic factors			
Age at diagnosis (years)†	54 (48, 61)	54 (48, 62)	0.71
Education $\geq$ high school (%)	40.5	39.7	0.41
Marital status, married (%)	87.0	87.6	0.94
Nongenetic risk factors			
Regular smoker (%)	3.2	3.5	0.69
Regular alcohol consumption (%)	3.1	5.5	<0.01
First-degree relative with any cancer (%)	35.3	29.1	<0.01
Family history of hormone-related female cancers (%)‡	5.3	3.4	0.03
Menopausal status (%)	56.7	61.7	0.02
Age at menarche (years)†	14 (13, 16)	15 (13, 16)	<0.01
Age at menopause (years)†	50.3 (48.6, 52.5)	49.5 (47.3, 51.1)	<0.01
Years of menstruation†	33.2 (30.1, 36.0)	31.4 (28.1, 34.2)	<0.01
No. of pregnancies†	3 (2, 4)	3 (2, 4)	<0.01
Oral contraceptive use, ever (%)	18.3	25.0	<0.001
Hormone replacement therapy use, ever (%)	4.7	4.3	0.62
No regular exercise (%)	71.8	66.3	0.05
Intensity of physical activities (metabolic equivalent tasks)†	9.49 (7.30, 12.95)	10.82 (7.85, 13.65)	<0.001
Diagnosis of diabetes (%)	15.0	6.9	<0.01
Body mass index (kg/m^2)†	25.3 (22.9, 28.2)	23.4 (21.4, 25.9)	<0.0001
Total caloric intake (kcal/day)†	1,739 (1,473, 2,048)	1,711 (1,424, 2,14)	0.03
Total fruit and vegetable intake (g/day)†	509.8 (341.6, 699.0)	488.3 (338.3, 684.1)	0.75
CYP19A1 genotypes			
rs28566535 (%)			
AA	41.9	43.8	
AC	46.0	45.6	
CC	12.1	10.6	0.46
rs1065779 (%)			
GG	31.9	30.8	
GT	51.5	48.1	
TT	16.6	21.2	0.03
rs752760 (%)			
TT	35.5	39.3	
TC	49.5	45.0	
CC	15.0	15.7	0.12
rs1870050 (%)			
AA	55.8	47.5	
AC	37.7	42.3	
CC	6.5	10.2	<0.01
rs700519 (%)			
CC	75.4	75.4	
CT	23.1	22.6	
TT	1.5	2.0	0.68

* For a χ^2 test (categorical variables) or a *t* test (continuous variables).
† Median (25th, 75th percentiles).
‡ Breast, endometrial, and ovarian cancers.

TABLE 2. Association of soy food and tea consumption with endometrial cancer risk, Shanghai Endometrial Cancer Study, 1997–2003

	No. of cases (n = 1,204)	No. of controls (n = 1,212)	Odds ratio*	95% confidence interval	Odds ratio†	95% confidence interval
Soy protein (g/day)‡						
<7.1	384	404	1.0		1.0	
7.1–13.3	429	405	1.0	0.8, 1.2	1.0	0.8, 1.2
>13.3	391	403	0.8	0.6, 1.0	0.8	0.6, 1.0
p_{trend}				0.06		0.07
Tea consumption						
Never	845	834	1.0		1.0	
Ever	359	378	0.8	0.6, 0.9	0.8	0.6, 1.0
Primarily green tea	322	343	0.8	0.6, 0.9	0.8	0.6, 0.9
Primarily black tea	21	25	0.7	0.4, 1.3	0.7	0.4, 1.3
Other types of tea	16	10	1.5	0.6, 3.6	1.5	0.6, 3.5
Amount of tea consumed (g/year)						
Never	845	834	1.0		1.0	
≤1,500	180	189	0.8	0.6, 1.0	0.8	0.6, 1.0
>1,500	167	171	0.8	0.6, 1.0	0.8	0.6, 1.0
p_{trend}				0.03		0.03
Years of tea consumption						
Never	845	834	1.0		1.0	
1–10	78	118	0.6	0.4, 0.8	0.6	0.4, 0.8
11–20	106	96	1.0	0.7, 1.3	1.0	0.7, 1.3
>20	175	164	0.8	0.7, 1.1	0.8	0.6, 1.1
p_{trend}				0.13		0.11
Frequency of tea consumption (times/week)						
Never	845	834	1.0		1.0	
1–6	46	32	1.2	0.8, 2.0	1.2	0.8, 2.0
7	294	328	0.7	0.6, 0.9	0.7	0.6, 0.9
>7	19	17	0.8	0.4, 1.7	0.8	0.4, 1.7
p_{trend}				0.0063		0.0062

* Adjusted for age, education, menopausal status, years of menstruation, number of pregnancies, diagnosis of diabetes, alcohol consumption, body mass index, physical activity, energy intake, and total fruit and vegetable intake.

† Additionally adjusted for tea consumption (for soy protein intake) or soy protein intake (for tea-related items).

‡ Soy protein: protein from soymilk, tofu, soy bean sprouts, fresh soybeans, and other soy products.

three SNPs when compared with the data from HapMap (http://www.hapmap.org) and/or Perlegen (http://genome.perlegen.com). The other two SNPs, rs1065779 and rs28566535, were not genotyped in either of these two databases.

For statistical analysis, chi-squared tests were used to evaluate case-control differences in the distribution of genotypes. Multivariate analyses were performed to adjust for potential confounding variables, including age (continuous variable), education (no formal education, elementary, middle school, high school, college), menopausal status, years of menstruation (<25, <30, <35, ≥35 years), number of pregnancies (0, 1, 2, 3, 4, ≥5), body mass index (by quintile), alcohol consumption (never/ever), diagnosis of diabetes (never/ever), physical activity in metabolic equivalent tasks (METs) (by quintile), total energy intake (by quintile), and total fruit and vegetable intake (by quintile). Logistic regression models were used to estimate odds ratios and their 95 percent confidence intervals. Haplotypes for all

TABLE 3. Association of *CYP19A1* genotypes with endometrial cancer risk, stratified by dietary soy protein and tea intake status, Shanghai Endometrial Cancer Study, 1997–2003

CYP19A1 genotype	Low intake				High intake				$p_{interaction}$
	No. of cases	No. of controls	Odds ratio*	95% confidence interval	No. of cases	No. of controls	Odds ratio*	95% confidence interval	
Soy protein†									
Rs28566535									
AA	139	138	1.0		297	314	1.0		
AC	146	154	1.0	0.7, 1.4	332	316	1.1	0.9, 1.4	
CC	45	44	1.1	0.6, 1.8	81	65	1.3	0.9, 1.9	0.46
p_{trend}				0.86				0.21	
Rs1065779									
GG	101	107	1.0		229	210	1.0		
GT	167	154	1.3	0.9, 1.9	365	341	0.9	0.7, 1.2	
TT	59	74	0.9	0.6, 1.5	112	144	0.7	0.5, 1.0	0.45
p_{trend}				0.92				0.05	
Rs752760									
TT	115	126	1.0		250	265	1.0		
TC	157	146	1.1	0.8, 1.6	351	302	1.2	0.9, 1.6	
CC	54	50	1.0	0.6, 1.7	100	106	0.9	0.6, 1.3	0.85
p_{trend}				0.87				0.89	
Rs1870050									
AA	179	155	1.0		387	323	1.0		
AC	127	138	0.8	0.6, 1.2	256	288	0.8	0.6, 1.0	
CC	19	30	0.5	0.3, 1.0	47	73	0.6	0.4, 0.9	0.92
p_{trend}				0.04				<0.01	
Rs700519									
CC	245	244	1.0		528	530	1.0		
CT	73	78	1.0	0.7, 1.5	164	154	1.1	0.8, 1.4	
TT	5	11	0.3	0.1, 1.1	10	9	1.2	0.4, 3.1	0.40
p_{trend}				0.09				0.64	

Table continues

five SNPs were constructed on the basis of their chromosomal position (SNP1, rs1065779; SNP2, rs700519; SNP3, rs28566535; SNP4, rs752760; SNP5, rs1870050) via a Bayesian approach using PHASE software (26, 27). Because of the significant interaction of SNPs rs1065779, rs752760, and rs1870050 with tea consumption in the risk of endometrial cancer, haplotypes were reconstructed on the basis of these three SNPs (rs1065779, rs752760, rs1870050). The differences of haplotype frequencies between cases and controls were tested with 100 times of permutation (28). The haplotype data were analyzed by weighted haplotype probability. Logistic regression was used to assess the association of endometrial cancer risk with each haplotype under dominant, recessive, and additive genetic models. Stratified and joint association analyses were performed to evaluate whether intakes of soy food and tea modified the associations of *CYP19A1* genotypes with endometrial cancer risk. The likelihood ratio test was conducted to formally test multiplicative interactions. p values of less than 0.05 (two-sided probability) were interpreted as statistically significant.

RESULTS

Selected demographic and risk factors were compared between cases and controls as shown in table 1. Cases and controls were similar in age, marital status, and education. There were no significant differences between cases and controls with respect to use of hormone replacement therapy, cigarette smoking, or total fruit and vegetable intake. Compared with controls, cases were significantly more likely to have a younger age at menarche, an older age at menopause, a greater number of years of menstruation, a first-degree relative with any cancer or a hormone-related female cancer, and a higher body mass index; to be

TABLE 3. Continued

CYP19A1 genotype	Low intake				High intake				$p_{interaction}$
	No. of cases	No. of controls	Odds ratio*	95% confidence interval	No. of cases	No. of controls	Odds ratio*	95% confidence interval	
Tea‡									
Rs28566535									
AA	301	307	1.0		132	144	1.0		
AC	331	321	1.0	0.8, 1.3	146	147	1.2	0.9, 1.8	
CC	90	75	1.3	0.9, 1.9	36	34	1.1	0.6, 2.0	0.60
p_{trend}				0.35				0.40	
Rs1065779									
GG	211	226	1.0		117	90	1.0		
GT	382	322	1.3	1.0, 1.7	148	171	0.6	0.4, 0.9	
TT	126	153	0.9	0.6, 1.2	45	65	0.5	0.3, 0.8	0.01
p_{trend}				0.75				<0.01	
Rs752760									
TT	261	257	1.0		103	134	1.0		
TC	344	319	1.0	0.8, 1.3	161	127	1.7	1.2, 2.6	
CC	108	102	0.9	0.6, 1.3	46	53	1.0	0.6, 1.7	0.04
p_{trend}				0.63				0.29	
Rs1870050									
AA	389	326	1.0		176	149	1.0		
AC	266	303	0.8	0.6, 1.0	115	123	0.9	0.6, 1.3	
CC	50	61	0.8	0.5, 1.2	15	42	0.2	0.1, 0.5	0.01
p_{trend}				0.04				<0.01	
Rs700519									
CC	533	527	1.0		240	247	1.0		
CT	170	159	1.1	0.8, 1.5	67	73	0.8	0.6, 1.3	
TT	12	15	0.6	0.3, 1.4	3	5	0.7	0.1, 3.7	0.89
p_{trend}				0.93				0.39	

* Adjusted for age, education, menopausal status, years of menstruation, number of pregnancies, diagnosis of diabetes, alcohol consumption, body mass index, physical activity, caloric intake, and total fruit and vegetable intake.

† High and low levels of intake were classified by the 33rd percentile of intake.

‡ Never and ever.

premenopausal, nulliparous, or diagnosed with diabetes; and to consume more total energy. Cases were less likely to be physically active, to drink alcohol, or to use oral contraceptives. There were no appreciable differences between subjects included in the present study and those included in the parent study (data not shown).

The distribution of all five SNPs was consistent with Hardy-Weinberg equilibrium among controls (data not shown in the table). The frequencies of the genotypes were significantly different between cases and controls for SNPs rs1065779 and rs1870050 (table 1).

High intake of soy protein was associated with a decreased risk of endometrial cancer, with an odds ratio of 0.8 (95 percent confidence interval: 0.6, 1.0) for the highest versus the lowest tertile of intake (table 2). Compared with non-tea drinkers, tea drinkers had a 20 percent (odds ratio = 0.8, 95 percent confidence interval: 0.6, 0.9) reduced risk of endometrial cancer. Of 737 tea drinkers, 665 (90.2 percent) drank primarily green tea and 46 (6.2 percent) drank mainly black tea. Excluding the 72 women who drank primarily non-green tea did not change the results substantially. Among tea drinkers, we did not observe a significant dose-response association when frequency, amount, or duration of tea consumption was considered. Soy intake and tea consumption were not significantly correlated in this population (data not shown).

We evaluated the joint effect of diet and *CYP19A1* genetic polymorphisms on endometrial cancer risk (table 3). Soy protein intake did not significantly interact with any of the five *CYP19A1* SNPs. We estimated the amount of soy

TABLE 4. Odds ratios of *CYP19A1* haplotypes with endometrial cancer risk, stratified by tea consumption, Shanghai Endometrial Cancer Study, 1997–2003

Haplotypes (no. of minor alleles)	Cases (%) (n = 1,042)	Controls (%) (n = 1,035)	All subjects						Non-tea drinkers						Tea drinkers					
			Dominant		Recessive		Additive		Dominant		Recessive		Additive		Dominant		Recessive		Additive	
			Odds ratio*	95% confidence interval	Odds ratio*	95% confidence interval	Odds ratio*	95% confidence interval	Odds ratio*	95% confidence interval	Odds ratio*	95% confidence interval	Odds ratio*	95% confidence interval	Odds ratio*	95% confidence interval	Odds ratio*	95% confidence interval	Odds ratio*	95% confidence interval
Haplotype 1†																				
GCACA (1)	26.2	27.8	1.2	1.0, 1.4	0.7	0.5, 1.0	1.0	0.9, 1.2	1.1	0.8, 1.4	0.7	0.5, 1.2	1.0	0.8, 1.2	1.5	1.1, 2.2	0.8	0.4, 1.5	1.2	0.9, 1.6
GCCTA (1)	14.1	16.0	1.2	1.0, 1.5	2.2	1.1, 4.6	1.2	1.0, 1.5	1.1	0.8, 1.4	2.3	0.9, 6.0	1.1	0.9, 1.5	1.7	1.2, 2.6	2.3	0.7, 7.6	1.7	1.2, 2.4
GCATC (1)	9.6	8.7	0.9	0.7, 1.2	0.8	0.3, 2.3	0.9	0.7, 1.2	0.9	0.6, 1.2	1.6	0.4, 6.9	0.9	0.7, 1.2	1.0	0.6, 1.7	0.1	0.01, 1.4	0.9	0.6, 1.4
TCATC (2)	15.6	12.3	0.7	0.6, 0.9	0.4	0.2, 0.8	0.7	0.6, 0.9	0.8	0.6, 1.1	0.6	0.2, 1.7	0.8	0.6, 1.1	0.6	0.4, 0.9	0.2	0.1, 0.7	0.6	0.4, 0.8
TCACA (2)	8.3	8.5	1.1	0.8, 1.4	0.7	0.2, 2.0	1.0	0.8, 1.3	1.1	0.8, 1.6	0.5	0.1, 2.0	1.0	0.8, 1.4	0.9	0.5, 1.5	0.8	0.1, 6.7	0.9	0.5, 1.4
TTCTA (3)	7.6	8.8	1.3	1.0, 1.6	0.7	0.2, 2.2	1.2	0.9, 1.6	1.3	0.9, 1.7	1.0	0.2, 4.2	1.2	0.9, 1.7	1.2	0.8, 2.0			1.1	0.7, 1.8
Haplotype 2‡																				
GTA (0)	18.6	15.8	1.4	1.1, 1.7	1.7	0.9, 3.0	1.3	1.1, 1.6	1.3	0.9, 1.7	1.5	0.8, 3.1	1.2	1.0, 1.6	2.0	1.3, 3.0	2.6	0.9, 7.8	1.8	1.3, 2.6
GCA (1)	29.2	27.8	1.2	1.0, 1.4	0.7	0.5, 1.0	1.0	0.9, 1.2	1.0	0.8, 1.3	0.8	0.5, 1.2	1.0	0.8, 1.2	1.5	1.0, 2.2	0.7	0.4, 1.4	1.2	0.9, 1.6
TTA (1)	16.4	15.0	1.1	0.9, 1.4	1.8	1.0, 3.2	1.2	1.0, 1.6	1.2	0.9, 1.6	2.3	1.1, 4.7	1.3	1.0, 1.6	1.0	0.6, 1.5	0.9	0.3, 2.8	1.0	0.7, 1.4
GTC (1)	9.8	11.1	0.8	0.6, 1.1	0.8	0.3, 2.2	0.9	0.7, 1.1	0.8	0.6, 1.1	1.4	0.4, 4.9	0.8	0.6, 1.1	1.0	0.6, 1.6	0.1	0.02, 1.4	0.9	0.5, 1.4
TTC (2)	15.5	20.1	0.7	0.6, 0.9	0.3	0.2, 0.6	0.7	0.6, 0.8	0.8	0.6, 1.1	0.4	0.2, 0.9	0.8	0.6, 1.0	0.5	0.3, 0.8	0.2	0.1, 0.5	0.5	0.4, 0.7
TCA (2)	10.3	10.0	1.1	0.8, 1.4	0.7	0.3, 1.7	1.0	0.8, 1.3	1.1	0.8, 1.5	0.5	0.2, 1.4	1.0	0.7, 1.3	0.9	0.6, 1.6	1.8	0.3, 11.5	1.0	0.6, 1.6

* Adjusted for age, education, menopausal status, years of menstruation, number of pregnancies, diagnosis of diabetes, alcohol consumption, body mass index, physical activity, caloric intake, soy protein intake, and total fruit and vegetable intake.

† In the order of single nucleotide polymorphisms rs1065779, rs700519, rs28566535, rs752760, and rs1870050 based on their chromosome position.

‡ In the order of single nucleotide polymorphisms rs1065779, rs752760, and rs1870050.

TABLE 5. Association of number of minor alleles in *CYP19A1* diplotypes with endometrial cancer risk, the Shanghai Endometrial Cancer Study, 1997–2003

No. of minor alleles in diplotypes	All subjects				Non-tea drinkers				Tea drinkers			
	No. of cases	No. of controls	Odds ratio*	95% confidence interval	No. of cases	No. of controls	Odds ratio*	95% confidence interval	No. of cases	No. of controls	Odds ratio*	95% confidence interval
Diplotype 1†												
0–2	362	323	1.0		239	228	1.0		123	95	1.0	
3	342	354	0.9	0.7, 1.1	238	235	1.0	0.8, 1.3	104	119	0.6	0.4, 0.9
4	209	195	0.9	0.7, 1.2	153	128	1.1	0.8, 1.5	56	67	0.6	0.4, 1.0
>4	129	163	0.8	0.6, 1.0	98	115	0.9	0.6, 1.2	31	48	0.5	0.3, 0.8
p_{trend}				0.08				0.64				<0.01
$p_{interaction}$								0.14				
Diplotype 2‡												
0–1	254	199	1.0		158	139	1.0		96	60	1.0	
2	407	404	0.8	0.6, 1.0	300	274	1.0	0.7, 1.3	107	130	0.5	0.3, 0.7
3	317	339	0.7	0.6, 0.9	221	232	0.8	0.6, 1.1	96	107	0.5	0.3, 0.8
4–5	64	93	0.5	0.4, 0.8	49	61	0.7	0.4, 1.1	15	32	0.2	0.1, 0.5
p_{trend}				<0.001				0.08				<0.001
$p_{interaction}$								0.02				

* Adjusted for age, education, menopausal status, years of menstruation, number of pregnancies, diagnosis of diabetes, alcohol consumption, body mass index, physical activity, caloric intake, soy protein intake, and total fruit and vegetable intake.

† Diplotype 1: in the order of single nucleotide polymorphisms rs1065779, rs700519, rs28566535, rs752760, and rs1870050 based on their chromosome position.

‡ Diplotype 2: in the order of single nucleotide polymorphisms rs1065779, rs752760, and rs1870050.

isoflavone intake by using the Chinese food composition tables (25) and assessed its association with endometrial cancer risk and its potential modifying effect on the gene-cancer association. We found that intake of isoflavones was inversely related to endometrial cancer risk but that it did not modify the gene-disease association (data not shown).

We found that tea consumption significantly modified the association of endometrial cancer risk with three SNPs: rs1065779 ($p_{interaction} = 0.01$), rs752760 ($p_{interaction} = 0.04$), and rs1870050 ($p_{interaction} = 0.01$). The most notable finding was that SNPs rs1065779 and rs1870050 were related to a reduced risk of endometrial cancer only among women who drank tea regularly. Additionally adjusting for soy protein intake did not materially change the results (data not shown).

We present in table 4 the results of association analyses among the common (frequency: >5 percent) *CYP19A1* haplotypes (for all five SNPs and only those three SNPs that significantly interacted with tea consumption—rs1065779, rs752760, and rs1870050). A global significant difference in haplotype frequencies was found for both the five-SNP and the three-SNP haplotypes ($p < 0.05$). Haplotypes TCATC and TTC were related to a reduced risk of endometrial cancer, and the associations were more pronounced among tea drinkers. For haplotypes GCCTA and GTA, conversely, a positive association with endometrial cancer was observed (table 4).

We conducted further exploratory analyses by grouping women together by the number of minor alleles of these SNPs that they carried. In a comparison of women with no or one minor allele in the diplotype derived from all five SNPs or three SNPs (rs1065779, rs752760, and rs1870050), the risk of endometrial cancer decreased with an increasing number of minor alleles (table 5). This reduction in risk was stronger among tea drinkers than among non-tea drinkers, although the interaction tests were significant only for the three-SNP–based diplotype ($p_{interaction} = 0.02$).

DISCUSSION

In this population-based case-control study, we found that dietary consumption of soy foods and tea was inversely associated with endometrial cancer risk. Three polymorphisms (rs1065779, rs752760, and rs1870050), in both single marker and haplotype analyses, were found to be related to the risk of endometrial cancer, and their effects were modified by tea consumption.

Our findings of an inverse association of endometrial cancer with intakes of soy food and tea are consistent with those of previous studies (12–15, 28), including our earlier reports from subsets of the SECS (12–15). Soy foods are rich in isoflavones, which have estrogenic or antiestrogenic properties due to the similarity of their structure to that of estradiol (7, 8). It has been suggested that these phytoestrogens, as well as tea polyphenols, decrease estrogen biosynthesis and produce antiestrogenic effects through lowering the activity of aromatase (16, 17, 29, 30). In addition to this antiestrogenic effect, polyphenols also have other anticarcinogenic properties, such as an antioxidant effect (31).

Aromatase is a crucial enzyme catalyzing three consecutive hydroxylation reactions converting C_{19} androgens to aromatic C_{18} estrogenic steroids, namely, androstenedione and testosterone to estrone and estradiol, respectively. The *CYP19A1* gene codes aromatase, and its transcription is regulated by tissue-specific promoters (32). The SNPs rs1870050 and rs752760 are located in the first exon, close to promoter I.1, the major promoter for the placenta. The SNP rs28566535 is located in the first exon close to promoter I.4, which is the promoter for adipose tissue, bone, and skin. The SNP rs1065779 is located in intron 9, 53 base pairs upstream of exon 10, which may affect transcription or expression of aromatase. The SNP rs700519 (Arg264Cys) is located in exon 7. By causing an amino acid change from arginine to cysteine, this SNP may result in a change of enzyme activity. Previous studies have found that variants in the *CYP19A1* gene lead to altered aromatase activity and estrogen levels (18, 19, 33), which are related to risk of hormone-related cancers (34, 35), including endometrial cancer (21, 22). We also found that several polymorphisms, SNPs rs1065779, rs752760, and rs1870050, were associated with risk of endometrial cancer (23), suggesting that these polymorphisms may play a critical role in the development of endometrial cancer.

In this study, we observed significant interactions between three *CYP19A1* genetic polymorphisms and tea consumption. Although these findings are new, they are biologically possible. It has been shown that green tea catechins and black tea polyphenol theaflavins decrease aromatase enzyme activity (17, 31). Our finding that the associations of rs1065779, rs752760, and rs1870050 with endometrial cancer are more pronounced in tea drinkers is in agreement with these results. It is possible that SNPs rs1065779 and rs752760, two polymorphisms close to the promoter for the placenta, may play an important role in the pathogenesis of endometrial cancer. SNP rs1065779, through its effect on transcription or the expression of aromatase, may also be involved in the carcinogenesis of endometrial cancer. The tea and single SNP interactions were also confirmed in the haplotype analyses. Of note, these interactions were most evident under the recessive model. Given that the functionality of SNPs and associated haplotypes is not well understood and that the genetic regulation of aromatase has not been extensively studied in endometrial carcinoma, further studies with a more comprehensive SNP coverage of the *CYP19A1* gene are needed to evaluate our findings.

We did not observe a significant interaction between intake of soy foods and any of the five polymorphisms in the *CYP19A1* gene. These results are not surprising, because previous studies have relatively consistently found that the inhibitory activity of isoflavones on aromatase, if any (30, 36–38), was much weaker than that of other polyphenols (30, 37–40). On the other hand, in vitro studies have found that soy isoflavones can inhibit the activity of 17β-hydroxysteroid dehydrogenase type I (17β-HSD1), a key enzyme in catalyzing estrone to the biologically more active estradiol (2, 41, 42). We have previously reported from the same study that soy consumption may interact with polymorphisms in the *17β-HSD1* gene in relation to endometrial cancer risk (43). Taken together, these findings suggest that isoflavones and tea polyphenols may modify the associations of genetic polymorphisms in different estrogen-related genes with endometrial cancer risk.

As with all case-control studies, the possibility for recall bias could not be completely eliminated. We tried to minimize recall bias by shortening the interval between diagnosis and interview for cases and by asking participants to ignore any dietary change over the preceding year. Furthermore, it is unlikely that recall bias would depend on an individual's genotype. In this study, we collected information on tea consumption according to the amount of tea leaves used, because Chinese people drink tea by putting loose tea leaves in a cup and repeatedly adding water using the same tea leaves. Variability in the amount of water added to the tea may have introduced measurement errors with regard to tea consumption. Likewise, soy intake was also subject to measurement errors. Misclassification of tea and soy consumption, which is most likely to be nondifferential, may bias the results toward the null. Although we adjusted for many potential confounding factors, we still cannot exclude the possibility that residual confounding or related dietary patterns may partially explain our results. Finally, we included only five polymorphisms in this study. These SNPs were chosen on the basis of literature review of published data and the potential functionality of the SNPs. Thus, our study is not as comprehensive as other studies that have applied the haplotype-tagging SNP approach (44, 45). Therefore, we cannot exclude the possibility that there may be other polymorphisms that interact with dietary polyphenols and that these unstudied polymorphisms may be responsible for the gene-tea interaction observed in this study.

Strengths of this study included the population-based study design, high response rate, high DNA sample donation rate, and the low frequency of hysterectomy in the population (3.6 percent), which minimized the selection bias. The homogeneous ethnic background (>98 percent Han Chinese) of the study participants avoided potential confounding from ethnicity. The large sample size enabled us to explore potential gene-environment interactions.

In summary, this population-based study suggests that consumption of soy food and tea was related to a reduced risk of endometrial cancer. Tea consumption may interact with genotypes of the *CYP19A1* gene in the etiology of endometrial cancer.

ACKNOWLEDGMENTS

This work was supported by US Public Health Service grant R01CA92585 from the National Cancer Institute.

The authors thank Dr. Fan Jin for her contributions to implementing the study in Shanghai; Drs. Qiuyin Cai, Hongmei Wu, and Regina Courtney for their contributions to the genotyping; and Bethanie Hull for her assistance in the preparation of this manuscript. This study would not have been possible without the support of all of the research staff of the Shanghai Endometrial Cancer Study.

Conflict of interest: none declared.

REFERENCES

1. Parkin DM, Whelan SL, Ferlay J, et al, eds. Cancer incidence in five continents. Vol VIII. Lyon, France: International Agency for Research on Cancer, 2002. (IARC scientific publication no. 155).
2. Adlercreutz H, Mazur W. Phyto-oestrogens and Western diseases. Ann Med 1997;29:95–120.
3. Wu AH, Yu MC. Tea, hormone-related cancers and endogenous hormone levels. Mol Nutr Food Res 2006;50:160–9.
4. Yang CS, Landau JM, Huang MT, et al. Inhibition of carcinogenesis by dietary polyphenolic compounds. Annu Rev Nutr 2001;21:381–406.
5. Kuntz S, Wenzel U, Daniel H. Comparative analysis of the effects of flavonoids on proliferation, cytotoxicity, and apoptosis in human colon cancer cell lines. Eur J Nutr 1999;38:133–42.
6. Hsieh CY, Santell RC, Haslam SZ, et al. Estrogenic effects of genistein on the growth of estrogen receptor-positive human breast cancer (MCF-7) cells in vitro and in vivo. Cancer Res 1998;58:3833–8.
7. Miksicek RJ. Interaction of naturally occurring nonsteroidal estrogens with expressed recombinant human estrogen receptor. J Steroid Biochem Mol Biol 1994;49:153–60.
8. Mueller SO, Simon S, Chae K, et al. Phytoestrogens and their human metabolites show distinct agonistic and antagonistic properties on estrogen receptor (ER) and ERß in human cells. Toxicol Sci 2004;80:14–25.
9. Rossi M, Negri E, Talamini R, et al. Flavonoids and colorectal cancer in Italy. Cancer Epidemiol Biomarkers Prev 2006;15:1555–8.
10. Yamamoto S, Sobue T, Kobayashi M, et al. Soy, isoflavones, and breast cancer risk in Japan. J Natl Cancer Inst 2003;95:906–13.
11. Lee AH, Fraser ML, Meng X, et al. Protective effects of green tea against prostate cancer. Expert Rev Anticancer Ther 2006;6:507–13.
12. Goodman MT, Wilkens LR, Hankin JH, et al. Association of soy and fiber consumption with the risk of endometrial cancer. Am J Epidemiol 1997;146:294–306.
13. Horn-Ross PL, John EM, Canchola AJ, et al. Phytoestrogen intake and endometrial cancer risk. J Natl Cancer Inst 2003;95:1158–64.
14. Xu WH, Zheng W, Xiang YB, et al. Soya food intake and risk of endometrial cancer among Chinese women in Shanghai: population based case-control study. BMJ 2004;328:1285–8.
15. Gao J, Xiang YB, Xu WH, et al. Green tea consumption and the risk of endometrial cancer: a population-based case-control study in urban Shanghai. (In Chinese). Zhonghua Liu Xing Bing Xue Za Zhi 2005;26:323–7.
16. Jeong HJ, Shin YG, Kim IH, et al. Inhibition of aromatase activity by flavonoids. Arch Pharm Res 1999;22:309–12.
17. Satoh K, Sakamoto Y, Ogata A, et al. Inhibition of aromatase activity by green tea extract catechins and their endocrinological effects of oral administration in rats. Food Chem Toxicol 2002;40:925–33.
18. Ma CX, Adjei AA, Salavaggione OE, et al. Human aromatase: gene resequencing and functional genomics. Cancer Res 2005;65:11071–82.
19. Simpson ER, Mahendroo MS, Means GD, et al. Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis. Endocr Rev 1994;15:342–55.
20. Jongen VH, Hollema H, Van Der Zee AG, et al. Aromatase in the context of breast and endometrial cancer. A review. Minerva Endocrinol 2006;31:47–60.
21. Paynter RA, Hankinson SE, Colditz GA, et al. *CYP19* (aromatase) haplotypes and endometrial cancer risk. Int J Cancer 2005;116:267–74.
22. Berstein LM, Imyanitov EN, Suspitsin EN, et al. *CYP19* gene polymorphism in endometrial cancer patients. J Cancer Res Clin Oncol 2001;127:135–8.
23. Tao MH, Cai Q, Zhang ZF, et al. Polymorphisms in the *CYP19A1* (aromatase) gene and endometrial cancer risk in Chinese women. Cancer Epidemiol Biomarkers Prev 2007;16:943–9.
24. Shu XO, Yang G, Jin F, et al. Validity and reproducibility of food frequency questionnaire used in the Shanghai Women's Health Study. Eur J Clin Nutr 2004;58:17–23.
25. Yang YX, Wang GY, Pang XC, eds. Chinese food composition tables. Beijing, China: Beijing University Medical Press, 2002.
26. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001;68:978–89.
27. Stephens M, Donnelly P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 2003;73:1162–9.
28. Zhao JH, Curtis D, Sham PC. Model-free analysis and permutation tests for allelic associations. Hum Hered 2000;50:133–9.
29. Petridou E, Kedikoglou S, Koukoulomatis P, et al. Diet in relation to endometrial cancer risk: a case-control study in Greece. Nutr Cancer 2002;44:16–22.
30. Pelissero C, Lenczowski MJ, Chinzi D, et al. Effects of flavonoids on aromatase activity, an in vitro study. J Steroid Biochem Mol Biol 1996;57:215–23.
31. Way TD, Lee HH, Kao MC, et al. Black tea polyphenol theaflavins inhibit aromatase activity and attenuate tamoxifen resistance in HER2/*neu*-transfected human breast cancer cells through tyrosine kinase suppression. Eur J Cancer 2004;40:2165–74.
32. Scalbert A, Mancah C, Morand C, et al. Dietary polyphenols and the prevention of disease. Crit Rev Food Sci Nutr 2005;45:287–306.
33. Bulun SE, Takayama K, Suzuki T, et al. Organization of the human aromatase P450 (*CYP19*) gene. Semin Reprod Med 2004;22:5–9.
34. Berstein LM, Imyanitov EN, Kovalevskij AJ, et al. CYP17 and *CYP19* genetic polymorphisms in endometrial cancer: association with intratumoral aromatase activity. Cancer Lett 2004;207:191–6.
35. Miyoshi Y, Iwao K, Ikeda N, et al. Breast cancer risk associated with polymorphism in *CYP19* in Japanese women. Int J Cancer 2000;89:325–8.
36. Ibrahim AR, Abul-Hajj YJ. Aromatase inhibition by flavonoids. J Steroid Biochem Mol Biol 1990;37:257–60.
37. Le Bail JC, Laroche T, Marre-Fournier F, et al. Aromatase and 17β-hydroxysteroid dehydrogenase inhibition by flavonoids. Cancer Lett 1998;133:101–6.
38. Kao YC, Zhou C, Sherman M, et al. Molecular basis of the inhibition of human aromatase (estrogen synthetase) by flavone and isoflavone phytoestrogens: a site-directed mutagenesis study. Environ Health Perspect 1998;106:85–92.
39. Lacey M, Bohday J, Fonseka SM, et al. Dose-response effects of phytoestrogens on the activity and expression of 3β-hydroxysteroid dehydrogenase and aromatase in human

granulosa-luteal cells. J Steroid Biochem Mol Biol 2005;96: 279–86.
40. Le Bail JC, Champavier Y, Chulia AJ, et al. Effects of phytoestrogens on aromatase, 3β and 17β-hydroxysteroid dehydrogenase activities and human breast cancer cells. Life Sci 2000;66:1281–91.
41. Makela S, Davis VL, Tally WC, et al. Dietary estrogens act through estrogen receptor-mediated processes and show no antiestrogenicity in cultured breast cancer cells. Environ Health Perspect 1994;102:572–8.
42. Brooks JD, Thompson LU. Mammalian lignans and genistein decrease the activities of aromatase and 17β-hydroxysteroid dehydrogenase in MCF-7 cells. J Steroid Biochem Mol Biol 2005;94:461–7.
43. Dai Q, Xu WH, Long JR, et al. Interaction of soy and *17β-HSD1* gene polymorphisms in the risk of endometrial cancer. Pharmacogenet Genomics 2007;17:161–7.
44. Long JR, Shu XO, Cai Q, et al. *CYP19A1* genetic polymorphisms may be associated with obesity-related phenotypes in Chinese women. Int J Obes (Lond) 2007;31:418–23.
45. Haiman CA, Dossus L, Setiawan VW, et al. Genetic variation at the *CYP19A1* locus predicts circulating estrogen levels but not breast cancer risk in postmenopausal women. Cancer Res 2007;67:1893–7.