Research Article

Cancer Epidemiology, Biomarkers & Prevention

Genome-Wide Association Study Identifies a Possible Susceptibility Locus for Endometrial Cancer

Jirong Long[1], Wei Zheng[1], Yong-Bing Xiang[2], Felicity Lose[4], Deborah Thompson[6], Ian Tomlinson[7], Herbert Yu[8], Nicolas Wentzensen[9], Diether Lambrechts[10], Thilo Dörk[11], Natalia Dubrowinskaja[11], Marc T. Goodman[12], Helga B. Salvesen[13,14], Peter A. Fasching[15,17], Rodney J. Scott[18,19], Ryan Delahanty[1], Ying Zheng[3], Tracy O'Mara[4,5], Catherine S. Healey[6], Shirley Hodgson[21], Harvey Risch[8], Hannah P. Yang[9], Frederic Amant[10], Nurzhan Turmanov[11], Anita Schwake[11], Galina Lurie[12], Jone Trovik[13,14], Matthias W. Beckmann[15], Katie Ashton[18], Bu-Tian Ji[9], Ping-Ping Bao[3], Kimberly Howarth[7], Lingeng Lu[8], Jolanta Lissowska[22], Lieve Coenegrachts[10], Dilyara Kaidarova[23], Matthias Dürst[24], Pamela J. Thompson[12], Camilla Krakstad[13,14], Arif B. Ekici[16], Geoffrey Otton[20], Jiajun Shi[1], Ben Zhang[1], Maggie Gorman[7], Louise Brinton[9], An Coosemans[10], Rayna K. Matsuno[12], Mari K. Halle[13,14], Alexander Hein[15], Anthony Proietto[20], Hui Cai[1], Wei Lu[3], Alison Dunning[6], Douglas Easton[6], Yu-Tang Gao[2], Qiuyin Cai[1], Amanda B. Spurdle[4] and Xiao-Ou Shu[1]

Abstract

Background: Genome-wide association studies (GWAS) have identified more than 100 genetic loci for various cancers. However, only one is for endometrial cancer.

Methods: We conducted a three-stage GWAS including 8,492 endometrial cancer cases and 16,596 controls. After analyzing 585,963 single-nucleotide polymorphisms (SNP) in 832 cases and 2,682 controls (stage I) from the Shanghai Endometrial Cancer Genetics Study, we selected the top 106 SNPs for *in silico* replication among 1,265 cases and 5,190 controls from the Australian/British Endometrial Cancer GWAS (stage II). Nine SNPs showed results consistent in direction with stage I with $P < 0.1$. These nine SNPs were investigated among 459 cases and 558 controls (stage IIIa) and six SNPs showed a direction of association consistent with stages I and II. These six SNPs, plus two additional SNPs selected on the basis of linkage disequilibrium and P values in stage II, were investigated among 5,936 cases and 8,166 controls from an additional 11 studies (stage IIIb).

Results: SNP rs1202524, near the *CAPN9* gene on chromosome 1q42.2, showed a consistent association with endometrial cancer risk across all three stages, with ORs of 1.09 [95% confidence interval (CI), 1.03–1.16] for the *A/G* genotype and 1.17 (95% CI, 1.05–1.30) for the *G/G* genotype ($P = 1.6 \times 10^{-4}$ in combined analyses of all samples). The association was stronger when limited to the endometrioid subtype, with ORs (95% CI) of 1.11 (1.04–1.18) and 1.21 (1.08–1.35), respectively ($P = 2.4 \times 10^{-5}$).

Conclusions: Chromosome 1q42.2 may host an endometrial cancer susceptibility locus.

Impact: This study identified a potential genetic locus for endometrial cancer risk. *Cancer Epidemiol Biomarkers Prev; 21(6); 980–7. ©2012 AACR.*

Authors' Affiliations: [1]Division of Epidemiology, Department of Medicine and Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee; [2]Department of Epidemiology, Shanghai Cancer Institute; [3]Shanghai Institute of Preventive Medicine, Shanghai, China; [4]Division of Genetics and Population Health, Queensland Institute of Medical Research, [5]Cancer Program, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia; [6]Department of Oncology, Strangeways Research Laboratory, University of Cambridge, Cambridge; [7]Wellcome Trust Centre for Human Genetics and NIHR Comprehensive Biomedical Research Centre, University of Oxford, Oxford, United Kingdom; [8]Department of Epidemiology and Public Health, Yale Cancer Center, Yale University School of Medicine, New Haven, Connecticut; [9]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland; [10]Division Gynaecological Oncology, Katholieke Universiteit Leuven, Leuven, Belgium; [11]Gynaecology Research Unit, Hannover Medical School, Hannover, Germany; [12]University of Hawaii Cancer Center, University of Hawaii, Honolulu, Hawaii; [13]Department of Obstetrics and Gynecology, Haukeland University Hospital; [14]Department of Clinical Medicine, University of Bergen, Bergen, Norway; [15]Department of Gynecology and Obstetrics, University Hospital Erlangen, [16]Institute of Human Genetics, Friedrich Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-Nuremberg, Nuremberg, Erlangen, Germany; [17]Division of Hematology and Oncology, Department of Medicine, University of California at Los Angeles, Los Angeles, California; [18]University of Newcastle; [19]The Hunter Area Pathology Service, [20]Hunter New England Health Service, Newcastle, New South Wales, Australia; [21]Clinical Genetics, St George's Hospital Medical School, London, United Kingdom; [22]Department of Cancer Epidemiology and Prevention, The M. Sklodowska-Curie Cancer Center and Institute of Oncology, Warsaw, Poland; [23]Almaty Oncology Center, State Oncology Institute, Almaty, Kazakhstan; and [24]Department of Gynecology, Jena University Hospital - Friedrich Schiller University Jena, Germany

Corresponding Authors: Xiao-Ou Shu, Professor of Medicine, Vanderbilt Epidemiology Center, Institute for Medicine and Public Health, Vanderbilt University School of Medicine, 2525 West End Avenue, Suite 600, Nashville, TN 37203. Phone: 615-936-0713; Fax: 615-936-8291; E-mail: xiao-ou.shu@vanderbilt.edu

doi: 10.1158/1055-9965.EPI-11-1160

Introduction

Endometrial cancer is the second most common gynecologic cancer in the world (1). Early age at menarche, late age at menopause, nulliparity, obesity, use of estrogen-alone hormone replacement therapy, use of tamoxifen, and family history of endometrial, breast, ovarian, or colorectal cancer are established risk factors for endometrial cancer (2). Over the past 10 years, a number of studies have reported that certain genetic variants are associated with endometrial cancer risk, including polymorphisms in genes involved in estrogen metabolism (3). However, very few of these associations have been confirmed (4–7). Given that almost all previous studies used the candidate gene approach, in which only a limited number of genetic variants are investigated, and the choice of genes is driven by our limited knowledge of cancer biology, more comprehensive genetic investigations of endometrial cancer risk are urgently needed. Genome-wide association studies (GWAS) have successfully identified more than 100 genetic loci at $P < 5 \times 10^{-8}$ for common cancers (8). Many GWAS-identified markers are located in genes or nongene regions that had not previously been associated with cancer susceptibility. Very recently, the first endometrial cancer GWAS, conducted by Spurdle and colleagues (9), reported a novel endometrial cancer susceptibility marker, rs4430796. This marker is located on chromosome 17q, very close to the *HNF1B* gene, and has also been associated with an increased risk of prostate cancer (10, 11) and decreased risk of type II diabetes (12). To identify additional endometrial cancer risk variants, we conducted a 3-stage GWAS among 24,071 patients with endometrial cancer and controls.

Materials and Methods

An overview of the study design, including study stage, study population, and single-nucleotide polymorphism (SNP) selection, are presented in Fig. 1.

Study population

Fourteen studies contributed a total of 8,492 endometrial cancer cases and 16,596 controls to this study. Detailed descriptions of participating studies are included in the Supplementary text. The discovery stage (stage I) was conducted among participants of the Shanghai Endometrial Cancer Genetics Study (SECGS) and included 832 endometrial cancer cases from the Shanghai Endometrial Cancer Study (SECS) and 2,682 controls from the Shanghai Breast Cancer Study (SBCS) and Shanghai Women's Health Study (SWHS). Data for stage II were extracted from the Australian/British Endometrial Cancer GWAS (9), which included 599 cases from the Australian National Endometrial Cancer Study (ANECS), 666 cases from the Study of Epidemiology and Risk Factors in Cancer Heredity (SEARCH), and 5,190 controls from the Wellcome Trust Case–Control Consortium (WTCCC). All cases in stage II were endometrioid subtype cases. Data for stage IIIa came from the Polish Endometrial Cancer Study (PECS) and included 459 cases and 558 controls. The final replication stage (stage IIIb) included 5,936 cases and 8,166 controls from the Bavarian Endometrial Cancer Study (BECS), Connecticut Endometrial Cancer Study (CECS), Hannover-Almaty Endometrial Cancer Study (HAECS), Hawaii Endometrial Cancer Study (HECS), Hannover-Jena Endometrial Cancer Study (HJECS), Leuven Endometrial Cancer Study (LES), Molecular Markers in Treatment of Endometrial Cancer (MoMaTEC), National Study of Endometrial Cancer Genetics (NSECG), additional samples from SEARCH that were not included in stage II, additional samples from ANECS that were not included in stage II analyzed together with Australian cases recruited into the Newcastle Endometrial Cancer Study (NECS) and Australian controls recruited via the Australian Red Cross Blood Service (ARCBS), and additional samples from the SECGS (SECGS-II) that were not included in stage I (Supplementary Table S1).

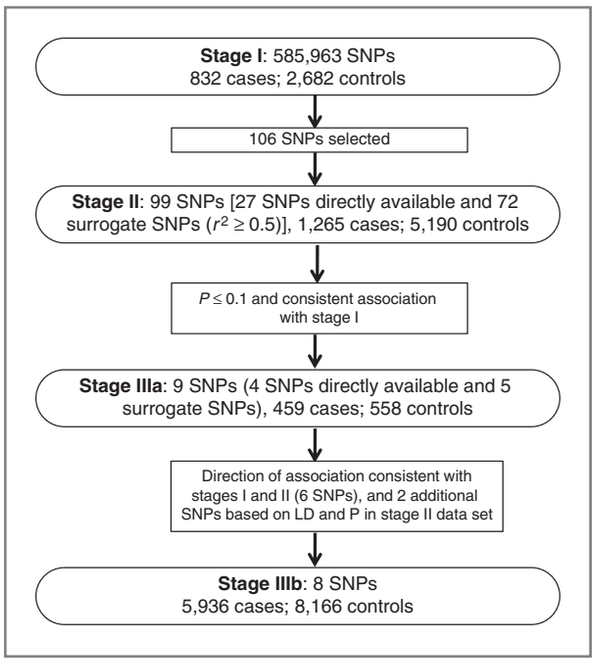

Figure 1. Overview of the study design.

The study protocol was approved by the Institutional Review Boards at Vanderbilt University Medical Center (Nashville, TN) and each of the collaborating institutes. Informed consent was obtained from all participants.

Stage I genotyping

Genotyping was conducted by the Affymetrix Genome-Wide Human SNP Array 6.0. All genotyping protocols and quality control (QC) procedures were identical to those used in our breast cancer GWAS (13). We included one negative and 3 positive QC samples from Coriell Cell Repositories (14) in each of the 96-well plates for GWAS genotyping. The average concordance rate among QC samples was 99.85% with a median value of 100%. The sex of all study samples was confirmed to be female.

Multidimensional scaling analyses of the pooled data set, which included data from 210 unrelated HapMap samples and stage I of the present study, showed that all study participants clustered closely with HapMap Asians. All close relatives with a pairwise proportion of identity by descent (IBD) estimate of PI_HAT $\geq$ 0.25 were excluded. SNPs that fell into any of the following categories were excluded: (i) minor allele frequency (MAF) <5%; (ii) call rate <95%; (iii) Hardy–Weinberg equilibrium (HWE) test $P < 1.0 \times 10^{-4}$; (iv) concordance rate <95% among QC samples; and (v) significantly different missing rates for cases and controls ($P < 1.0 \times 10^{-4}$). After these exclusion criteria were applied, 585,963 SNPs remained for the stage I analyses.

Stage II SNP selection and genotyping

We selected 106 SNPs for *in silico* replication in stage II according to the following criteria: (i) MAF $\geq$10%; (ii) $P <$ 0.001 in stage I; (iii) high genotyping quality, as indicated by very clear genotyping clusters; and (iv) included in the Illumina 610K array (stage II samples were genotyped using this platform). For SNPs not included in this array, we used a surrogate SNP in linkage disequilibrium (LD; $r^2 \geq 0.5$) with the originally selected SNP based on HapMap Caucasian data. In addition, if multiple SNPs in the same region met the above criteria, an SNP with a high functional score was chosen. The functional score was derived from SNP factors available through the NIEHS SNP Function Prediction (FuncPred) tool (15). Potential functional SNPs are those causing amino acid changes, alterations in splicing regulation, stop codons, structural changes, changes in transcription factor–binding sites, and changes in miRNA-binding site. Last, we confirmed that there was no LD between the selected SNPs ($r^2 < 0.2$).

Stage II genotyping

Methods have been described in detail elsewhere (9). In brief, endometrioid endometrial cancer subtype cases were genotyped by the Illumina Human 610K array; control data were extracted for these 610K SNPs from existing Illumina 1.2M genome-wide scan data for the WTCCC2 (16). Of the top 106 SNPs selected from the current stage I, data were available for 99 SNPs in the stage II data set, including 27 directly genotyped SNPs and 72 surrogate SNPs (Supplementary Table S2; imputation data were not available at the time of study implementation). Surrogate SNPs were selected on the basis of the following criteria: (i) data available in stage II; (ii) in LD with the index SNP ($r^2 > 0.5$ in Caucasians); (iii) if multiple SNPs were in LD with the index SNP, the SNP with the strongest LD with the index SNP was selected; and (iv) if multiple SNPs in similar LD with the index SNP in Caucasians, the SNP with the strongest LD in Asians was selected.

Stage III SNP selection and genotyping

Among the 99 SNPs investigated in stage II, 9 SNPs showed a direction of association consistent with stage I and were further investigated for replication in stage IIIa. Stage IIIa data were extracted from the GWAS data set genotyped by Illumina 660W-Quad from the PECS. We used directly genotyped data for 4 SNPs and used surrogates for the other 5 SNPs (Supplementary Table S2) in an approach similar to that used for stage II. Among these 9 SNPs, 6 SNPs (rs1202529, rs17556883, rs10795561, rs11663212, rs12208947, and rs190262) showed a direction of association consistent with stages I and II and were selected for genotyping in stage IIIb samples. After careful reevaluation of these 6 loci in the Australian/British GWAS data (9), rs9397178 was more significant than the index SNP rs12208947 ($P = 0.009$ and 0.075, respectively). These 2 SNPs are in strong LD in Asians ($r^2 = 0.93$) and moderate LD in Caucasians ($r^2 = 0.43$) based on HapMap data. Similarly, the association of SNP rs1202524 was slightly more significant than rs3828128, a proxy for the index SNP rs1202529 ($P = 0.02$ and 0.045, respectively). Thus, these 2 SNPs, rs9397178 and rs1202524, were also included in the stage IIIb validation.

Stage IIIb genotyping was conducted by the Sequenom iPLEX MassArray or TaqMan platforms. Genotyping assay protocols were developed and validated at the Vanderbilt Molecular Epidemiology Laboratory. TaqMan genotyping assay reagents were provided to investigators of 3 studies, HAECS, HJCES, and SEARCH. SNP rs450712 was genotyped to serve as a surrogate for the index SNP rs190262 ($r^2 = 1.00$ in both Asians and Caucasians based on HapMap data) for TaqMan genotyping. Sequenom assay reagents were provided to 5 studies, including the ANECS, BECS, MoMaTEC, NSECG (genotyped in Brisbane), and LES (genotyped in Leuven). Samples from the SECGS-II, CECS, and HECS were genotyped using Sequenom at the Vanderbilt Molecular Epidemiology Laboratory. Another 435 cases from the SECGS-II were genotyped using the same Sequenom SNP pool at Fudan University in Shanghai, China. All studies complied with QC standards by including 2 or more non-DNA template controls per 384-well assay plate, assaying 2% or more duplicate samples, and achieving a genotyping call rate of >95% and concordance rate between duplicate samples of $\geq$95% for each SNP assay. Studies with HWE $P < 1 \times 10^{-4}$ were automatically excluded from the association analysis. Data for NSECG controls were drawn from the UK1/CORGI colorectal cancer GWAS, which used the Illumina 550K platform (17). SNPs not on the Illumina 550K were imputed using IMPUTE for the NSECG controls.

Statistical analyses

In stage I, PLINK (version 1.06) was used to analyze the genome-wide data. A set of 4,305 SNPs with a MAF $\geq$35% and a distance of $\geq$100 kb between 2 adjacent SNPs were selected to evaluate the population structure. The inflation factor λ was estimated to be 1.045, suggesting that any population substructure, if present, should not have any appreciable effect on the results.

Associations between SNPs and endometrial cancer risk were assessed using ORs and 95% confidence intervals (CI) derived from logistic regression models. ORs were estimated for heterozygotes and homozygotes for the variant allele compared with homozygotes for the common allele in stage I. ORs were also estimated for the variant allele based on a log-additive model. Meta-analyses for stages II and III were conducted using a weighted average method with the inverse-variance weights, $w = 1/se^2$. An overall z-statistic and P value was calculated from the weighted average of the individual statistics. All P values presented are based on 2-tailed tests.

Results

In stage I which was conducted in the Chinese population, multiple genomic locations were revealed as potentially related to endometrial cancer risk (Supplementary Fig. S1), and the observed number of SNPs with a small P value was larger than that expected by chance (Supplementary Fig. S2). P values presented in Supplementary Fig. S1 are derived from trend tests using logistic regression [degrees of freedom (df), 1].

Of the 99 SNPs selected for stage II *in silico* replication using data from the Australia/British GWAS (9), 13 SNPs had an association with endometrial cancer ($P < 0.1$) and of these SNPs, 9 had the same direction of association as in stage I (rs10795561, rs11663212, rs1202529, rs12208947, rs12516561, rs17556883, rs190262, rs2876846, and rs4729568; Supplementary Table S3). These 9 SNPs were further investigated for replication in the PECS (stage IIIa). Of these, 6 SNPs (rs10795561, rs11663212, rs1202529, rs12208947, rs17556883, and rs190262) had the same direction of association as in stages I and II (Supplementary

Table 1. Association results for the 8 SNPs included in all 3 stages of the study

			N		Frequency		OR (95% CI)		
SNP	Effect/reference allele	Study	Cases	Controls	Cases	Controls	Heterozygotes	Homozygotes	*P*
rs1202524	G/A	Stage I	832	2,682	0.12	0.09	1.47 (1.21–1.78)	2.25 (0.97–5.21)	2.3E-05
		Stage II	1,265	5,189	0.35	0.32	1.07 (0.94–1.22)	1.31 (1.07–1.59)	1.4E-02
		Stage III	6,371	8,686	0.28	0.26	1.06 (0.98–1.13)	1.10 (0.97–1.25)	7.2E-02
		Combined	8,468	16,557	0.27	0.25	1.09 (1.03–1.16)	1.17 (1.05–1.30)	1.6E-04
rs1202529	G/C	Stage I	824	2,668	0.12	0.09	1.45 (1.20–1.76)	1.97 (0.87–4.48)	5.5E-05
		Stage II	1,265	5,189	0.35	0.32	1.07 (0.94–1.22)	1.31 (1.07–1.59)	1.4E-02
		Stage III	6,366	8,684	0.28	0.26	1.06 (0.99–1.14)	1.09 (0.96–1.23)	8.1E-02
		Combined	8,455	16,541	0.27	0.25	1.10 (1.03–1.16)	1.16 (1.04–1.29)	2.4E-04
rs17556883	T/C	Stage I	832	2,680	0.35	0.40	0.77 (0.65–0.91)	0.67 (0.52–0.85)	1.9E-04
		Stage II	1,264	5,188	0.54	0.56	0.88 (0.75–1.03)	0.83 (0.70–1.00)	5.3E-02
		Stage III	6,361	8,663	0.52	0.52	0.94 (0.87–1.03)	0.96 (0.87–1.06)	3.3E-01
		Combined	8,457	16,531	0.51	0.52	0.90 (0.84–0.96)	0.90 (0.83–0.97)	3.1E-03
rs9397178	G/A	Stage I	832	2,682	0.41	0.37	1.16 (0.98–1.38)	1.41 (1.12–1.78)	3.6E-03
		Stage II	1,265	5,189	0.09	0.07	1.14 (0.96–1.36)	2.45 (1.37–4.40)	8.1E-03
		Stage III	6,369	8,683	0.13	0.15	0.97 (0.88–1.06)	1.07 (0.87–1.33)	8.8E-01
		Combined	8,466	16,554	0.15	0.16	1.03 (0.96–1.10)	1.28 (1.10–1.49)	1.9E-02
rs12208947	A/G	Stage I	831	2,534	0.41	0.37	1.16 (0.97–1.38)	1.48 (1.17–1.88)	1.3E-03
		Stage II	1,265	5,187	0.15	0.13	1.09 (0.94–1.26)	1.34 (0.89–2.04)	1.0E-01
		Stage III	6,367	8,684	0.17	0.19	0.96 (0.89–1.04)	0.97 (0.80–1.16)	3.0E-01
		Combined	8,463	16,405	0.19	0.20	1.01 (0.95–1.08)	1.15 (1.01–1.32)	2.0E-01
rs11663212	A/C	Stage I	832	2,679	0.30	0.25	1.34 (1.14–1.57)	1.42 (1.03–1.94)	3.0E-04
		Stage II	1,265	5,190	0.12	0.11	1.17 (1.00–1.36)	0.92 (0.50–1.68)	1.0E-01
		Stage III	6,358	8,657	0.14	0.14	1.02 (0.94–1.11)	1.08 (0.86–1.36)	4.2E-01
		Combined	8,455	16,526	0.15	0.15	1.10 (1.03–1.17)	1.16 (0.97–1.39)	3.7E-03
rs10795561	T/A	Stage I	832	2,681	0.29	0.26	1.16 (0.99–1.37)	1.55 (1.16–2.08)	2.1E-03
		Stage II	1,145	4,705	0.36	0.34	1.12 (0.98–1.29)	1.16 (0.94–1.44)	7.4E-02
		Stage III	6,353	8,653	0.34	0.34	1.00 (0.93–1.08)	1.08 (0.97–1.21)	3.2E-01
		Combined	8,330	16,039	0.34	0.33	1.04 (0.98–1.11)	1.14 (1.04–1.25)	8.2E-03
rs190262	T/G	Stage I	832	2,681	0.16	0.13	1.25 (1.05–1.50)	1.84 (1.12–3.01)	1.2E-03
		Stage II	1,265	5,190	0.13	0.11	1.19 (1.02–1.38)	1.30 (0.81–2.11)	1.5E-02
		Stage III	6,356	8,663	0.11	0.12	0.93 (0.86–1.01)	1.06 (0.79–1.42)	2.1E-01
		Combined	8,453	16,534	0.12	0.12	1.02 (0.96–1.09)	1.24 (0.99–1.55)	1.7E-01

Table 2. Association of rs1202524 and 1202529 with type I endometrial cancer, for all participants and by ethnicity

		N		Frequency		OR (95% CI)		
Population	**SNP**	**Cases**	**Controls**	**Cases**	**Controls**	**Heterzygotes**	**Homozygotes**	***P***
All women: type I endometrial cancer cases vs. controls								
	rs1202524	7,558	16,557	0.27	0.25	1.11 (1.04–1.18)	1.21 (1.08–1.35)	2.4E-05
	rs1202529	7,546	16,541	0.27	0.25	1.11 (1.04–1.18)	1.19 (1.07–1.33)	3.2E-05
Asian women: type I endometrial cancer cases vs. controls								
	rs1202524	1,689	3,829	0.11	0.10	1.22 (1.05–1.42)	1.60 (0.86–2.96)	3.0E-03
	rs1202529	1,681	3,812	0.11	0.10	1.22 (1.05–1.42)	1.45 (0.78–2.69)	4.6E-03
European-ancestry women: type I endometrial cancer cases vs. controls								
	rs1202524	6,646	11,797	0.32	0.31	1.07 (1.00–1.15)	1.18 (1.06–1.32)	1.7E-03
	rs1202529	6,641	11,796	0.32	0.31	1.07 (1.00–1.15)	1.17 (1.05–1.31)	2.2E-03

Table S3). After further evaluation of these 6 loci in the stage II data (9), we found that the association for surrogate SNPs rs9397178 and rs1202524 were more significant than the index SNPs rs12208947 and rs1202529. Thus, the 2 surrogate SNPs, rs9397178 and rs1202524, were also included in the stage IIIb validation.

Among the 8 SNPs included in the stage IIIb replication, an association was observed for rs1202524 ($P = 0.072$). The ORs (95% CI) were 1.06 (0.98–1.13) for the A/G genotype and 1.10 (0.97–1.25) for G/G genotype versus A/A (Table 1). Meta-analyses of all studies combined resulted in a P value of 1.6×10^{-4}. Restricting analyses to type I endometrial cancer (endometrioid) cases (7,558 of the 8,468 cases genotyped for this SNP) strengthened the association, with ORs (95% CI) of 1.11 (1.04–1.18) for the A/G genotype and 1.21 (1.08–1.35) for the G/G genotype ($P = 2.4 \times 10^{-5}$; Table 2). Analyses stratified by ethnicity showed that this SNP was associated with ORs of 1.22 (1.05–1.42) for the A/G genotype and 1.60 (0.86–2.96) for the G/G genotype in Asians and 1.07 (1.00–1.15) for the A/G genotype and 1.18 (1.06–1.32) for the G/G genotype among European-ancestry participants. Similar results were observed for SNP rs1202529, which is in strong LD with rs1202524 ($r^2 = 1.00$ in both Asian and Caucasian HapMap data), with P values of 2.4×10^{-4} and 3.2×10^{-5}, respectively, among all women and limited to type I cases (Table 2). Additional adjustment for body mass change did not change the results appreciably (data not shown). None of the other SNPs were associated with endometrial cancer risk in stage IIIb (Supplementary Table S4).

Discussion

In this 3-stage GWAS of endometrial cancer conducted among 8,492 cases and 16,596 controls, we found that SNPs rs1202524 and rs1202529, which are in strong LD, were associated with endometrial cancer risk among both Chinese women and women of European ancestry. Although neither association reached the conventionally used GWAS significance level of 5×10^{-8}, the associations were consistent across 9 of 14 contributing studies and did not vary by ethnicity. These 2 SNPs are located on chromosome 1q42.2, approximately 13 and 17 kb downstream of the *CAPN9* gene (Fig. 2). The *CAPN9* gene belongs to the calpain gene family, encoding a ubiquitous, well-conserved family of calcium-dependent, cysteine proteases. Multiple, alternatively spliced transcript variants encoding different isoforms have been found for this gene. A growing body of literature has implicated the role of calpain in various aspects of carcinogenesis, including cell-cycle progression, cellular differentiation, and apoptosis (18). TP53, a substrate of calpains, is known to play an important role in the pathogenesis of various tumors (19). Inhibitors of calpains have a stabilizing effect on TP53 and can increase TP53 levels and reduce tumor growth. It has been reported that calpain may play an important role in the regulation of estrogen receptor (ER) function. Shiba and colleagues (20, 21) found that calpain activity was significantly higher in breast cancer tissue compared with normal breast tissue, in addition to being higher in ER-positive tumors compared with ER-negative tumors. A very recent study showed that calpain 1, calpain 2, and the inhibitor calpastatin are expressed in endometrial tissue and that endometrial carcinoma has higher expression of calpastatin and calpain 2 than benign endometrial tissue (22). A number of oncogenic and tumor suppressor proteins are substrates of calpain family enzymes, including c-fos, c-jun, p53, the ER, and integrin (23). Calpain 9, expressed predominantly in the digestive tract, has been associated with suppression of gastric cancer (19) and is downregulated in gastric cancer tissue and both differentiated and poorly differentiated cell lines (24, 25). It has been suggested that calpain 9 might act as a tumor suppressor through degradation of oncogenic proteins (15).

There exists limited evidence for the functional significance of SNPs rs1202524 and rs1202529. Vasan and colleagues (26) found rs1202524 to be associated with left ventricular diastolic dimension ($P = 0.03$). Both rs1202529

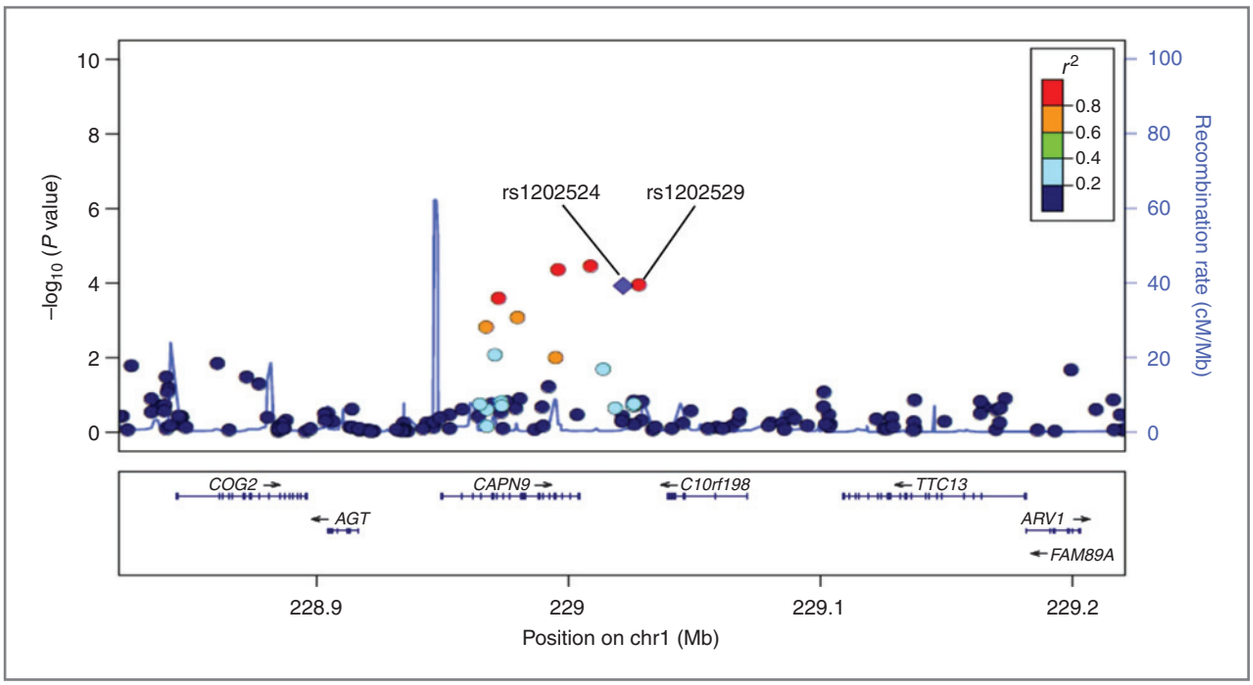

Figure 2. A regional plot of the $-\log_{10} P$ values for SNPs at 1q42.2. The LD is estimated using data from the HapMap Asian population. Also shown are recombination rates in centimorgans (cM) per megabase (Mb) and genes in the region (below) based on the March 2006 UCSC Genome Browser assembly.

and rs1202524 are in LD with 2 synonymous SNPs in the *CAPN9* gene, rs3828126 ($r^2 = 0.68/0.57$ in CEU/Asian data) and rs2282319 (0.48/0.10). Both synonymous SNPs are predicted to be exon splice enhancers by the ESE finder (27).

To date, only one SNP (rs4430796) has been identified by a GWAS of endometrial cancer (9). This SNP is located in the *HNF1B* gene on chromosome 17q. The per *G* allele OR (95% CI) was 0.84 (0.79–0.89; $P = 7.1 \times 10^{-10}$). This SNP was not significantly associated with endometrial cancer among Chinese women based on our stage I data (OR per *G* allele: 0.97; 95% CI, 0.84–1.12), but the direction of the association was consistent with what has been observed among European-ancestry women.

Several limitations of this study should be noted. First, although the present study is one of the largest epidemiologic studies of endometrial cancer, the sample size in the discovery stage is still small and mixed-ethnicity groups were included in the replication study. Unlike the genetic studies of more common cancers, such as the one we conducted on breast cancer (13), conducting a GWAS of endometrial cancer among Chinese populations alone or even among all Asians is currently unfeasible simply due to the rareness of this cancer and the limited amount of epidemiologic research conducted on it. Therefore, the validation study had to rely mainly on women of European ancestry. Because the allele frequencies of some SNPs differ across populations and LD patterns of adjacent SNPs also vary considerably by population, underlying differences in the genetic architecture of Asian- and European-ancestry populations could have resulted in false-negative results (i.e., failed replication) for some of the 99 SNPs prioritized for replication. Another limitation of this study is that some surrogate SNPs used in the replication stages showed only moderate LD with the index SNPs, which could have reduced the statistical power of the replication study. Nevertheless, finding a consistent association across different populations adds strong evidence for a true association. The lack of direct evidence on the functional significance of the 2 SNPs near the *CAPN9* gene is another limitation. Additional functional studies are needed to understand the nature of these SNP associations. Finally, cases with different subtypes of endometrial cancer were included in the present study, which may have biased the results toward the null. In fact, analyses restricted to women with endometrioid histology resulted in more significant associations with the 2 top SNPs.

In summary, in this large GWAS of endometrial cancer, we found one potential locus downstream of the *CAPN9* gene associated with endometrial cancer risk. Further research on this genetic region, including functional analyses, would shed light on the biologic role of *CAPN9* in endometrial carcinogenesis.

Disclosure of Potential Conflicts of Interest

The authors have no conflicts of interest to disclose.

Acknowledgments

The SECS GWAS thank Drs. Wanghong Xu, Fan Jin, and other research staff for their contributions to the field operation, Regina Courtney for

DNA preparation, and Bethanie Rammer for editorial support in the preparation of the manuscript. The HJECS thank Dr. Wen Zheng, Professor Peter Hillemanns, and Professor Ingo Runnebaum for their support in patient recruitment. The HAECS thank Professor Tjoung-Won Park-Simon for her support; the use of DNA from the British 1958 Birth Cohort collection; study participants and collaborators, and the research teams involved in design and implementation of the individual studies included; ANECS for the contributions of Study Investigator Penelope Webb; The Newcastle Endometrial Cancer Study (NECS) thank for the contributions from the University of Newcastle, The NBN Children's Cancer Research Group, Ms. Jennie Thomas, and the Hunter Medical Research Institute. PECS thank Drs. Montserrat Garcia-Closas and Mark Sherman and the Core Genotyping Facility at the Division of Cancer Epidemiology and Genetics of the National Cancer Institute, USA; Dr. Witold Zatonski of the Department of Cancer Epidemiology and Prevention, The M. Sklodowska-Curie Cancer Center and Institute of Oncology (Warsaw, Poland); Drs. Neonila Szeszenia-Dabrowska and Beata Peplonska of the Nofer Institute of Occupational Medicine (Lodz, Poland); and Pei Chao and Michael Stagner from Information Management Services (Silver Spring, MD), for their valuable contributions to the study (see Supplementary text for full list of collaborators and specific acknowledgements).

Grant Support

The SECS was supported by U.S. PHS grant R01 CA092585 (PI: to X.-O. Shu). Other studies that contributed to the SECS GWAS were funded by U.S. PHS grants, R01 CA064277, R01 CA090899, and R37 CA070867 (PI: to W. Zheng). The stage II ANECS/SEARCH research was supported by the National Health and Medical Research Council (ID#552402), the Wellcome Trust, and by Cancer Research UK grants C1287/A10118, C490/A1021, C8197/A10865, and C8197/A10123. A.B. Spurdle is an NHMRC Senior Research Fellow. T. O'Mara. is supported by an Australian Postgraduate Award, an Institute of Health and Biomedical Innovation PhD Top-Up, and an Qld Government Smart State PhD Award. D.F. Easton is a Principal Research Fellow of Cancer Research UK. A.M. Dunning is supported by the Joseph Mitchell Trust. I. Tomlinson is supported by Cancer Research UK and the Oxford Comprehensive Biomedical Research Centre. The ANECS/SEARCH study made use of data generated by the Wellcome Trust Case Control Consortium 2, and a full list of the investigators who contributed to the generation of these data is available from the WTCCC website (16). The use of DNA from the British 1958 Birth Cohort collection was funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Funding for this project was provided by the Wellcome Trust under award 085475. P.A. Fasching was partly funded by Dr. Mildred Scheel Stiftung of the Deutsche Krebshilfe (German Cancer Aid). F. Lose, A.B. Spurdle, and T. O'Mara represent the ANECS. ANECS was supported by recruitment by project grants from the National Health and Medical Research Council of Australia (ID#339435), The Cancer Council Queensland (ID#4196615), and Cancer Council Tasmania (ID#403031 and ID#457636). The Bavarian Endometrial Cancer Study (BECS) was partly funded by the ELAN fund of the University of Erlangen. The Leuven Endometrium Study (LES) was supported by the Verelst Foundation for endometrial cancer. MoMaTEC received financial support from a Helse Vest Grant, the University of Bergen, Melzer Foundation, the Norwegian Cancer Society (Harald Andersens legat), the Research Council of Norway, and Haukeland University Hospital. The National Study of Endometrial Cancer Genetics Group was supported principally by Cancer Research UK and by funds from the Oxford Comprehensive Biomedical Research Centre, with core infrastructure support to the Wellcome Trust Centre for Human Genetics, Oxford provided by grant 075491/Z/04. The Polish Endometrial Cancer Study (PECS) was funded by the intramural research program at the U.S. National Cancer Institute, Division of Cancer Epidemiology and Genetics in the Hormonal and Reproductive Epidemiology Branch. SEARCH is funded by a programme grant from Cancer Research UK (C490/A10124).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 21, 2011; revised February 13, 2012; accepted February 28, 2012; published OnlineFirst March 16, 2012.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.
2. Mueck AO, Seeger H, Rabe T. Hormonal contraception and risk of endometrial cancer: a systematic review. Endocr Relat Cancer 2010;17:R263–71.
3. Olson SH, Bandera EV, Orlow I. Variants in estrogen biosynthesis genes, sex steroid hormone levels, and endometrial cancer: a HuGE review. Am J Epidemiol 2007;165:235–45.
4. O'Mara TA, Fahey P, Ferguson K, Marquart L, Lambrechts D, Despierre E, et al. Progesterone receptor gene variants and risk of endometrial cancer. Carcinogenesis 2011;32:331–5.
5. Xu WH, Long JR, Zheng W, Ruan ZX, Cai Q, Cheng JR, et al. Association of the progesterone receptor gene with endometrial cancer risk in a Chinese population. Cancer 2009;115:2693–700.
6. Setiawan VW, Doherty JA, Shu XO, Akbari MR, Chen C, De Vivo I, et al. Two estrogen-related variants in CYP19A1 and endometrial cancer risk: a pooled analysis in the Epidemiology of Endometrial Cancer Consortium. Cancer Epidemiol Biomarkers Prev 2009;18:242–7.
7. Xu WH, Dai Q, Xiang YB, Long JR, Ruan ZX, Cheng JR, et al. Interaction of soy food and tea consumption with CYP19A1 genetic polymorphisms in the development of endometrial cancer. Am J Epidemiol 2007;166:1420–30.
8. GWAS catalog. Bethesda, MD: NIH; 2012. [cited 2012 Jan 26] Available from: http://www.genome.gov/26525384.
9. Spurdle AB, Thompson DJ, Ahmed S, Ferguson K, Healey CS, O'Mara T, et al. Genome-wide association study identifies a common variant associated with risk of endometrial cancer. Nat Genet 2011;43:451–4.
10. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet 2007;39:977–83.
11. Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet 2008;40:310–5.
12. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–89.
13. Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 2009;41:324–8.
14. Coriell Cell Repositories. Camden, NJ: Coriell Institute for Medical Research; 2012. [cited 2012 Jan 26]. Available from: http://ccr.coriell.org.
15. NIEHS SNP Function Prediction (FuncPred) tool. Research Triangle Park, NC: National Institute of Environmental Health Sciences; 2012. [cited 2012 Jan 26]. Available from: http://snpinfo.niehs.nih.gov/snpfunc.htm.
16. WTCCC2. London, UK: Wellcome Trust Case Control Consortium 2; 2012. [cited 2012 Jan 26]. Available from: http://www.wtccc.org.uk/ccc2.
17. Houlston RS, Cheadle J, Dobbins SE, Tenesa A, Jones AM, Howarth K, et al. Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. Nat Genet 2010;42:973–7.
18. Wang KK. Calpain and caspase: can you tell the difference? Trends Neurosci 2000;23:20–6.
19. Huang Y, Wang KK. The calpain family and human disease. Trends Mol Med 2001;7:355–62.
20. Shiba E, Kambayashi JI, Sakon M, Kawasaki T, Kobayashi T, Koyama H, et al. Ca^{2+}-dependent neutral protease (calpain) activity in breast cancer tissue and estrogen receptor status. Breast Cancer 1996;3:13–7.
21. Shiba E, Kim S, Fujitani M, Kambayashi JI, Kawamura I, Tsujimoto S, et al. Possible involvement of calpain in the growth of estrogen

receptor positive breast cancer cells. Anticancer Res 1996;16: 773–7.
22. Salehin D, Fromberg I, Haugk C, Dohmen B, Georg T, Bohle RM, et al. Immunhistochemical analysis for expression of calpain 1, calpain 2 and calpastatin in endometrial cancer. Anticancer Res 2010;30:2837–43.
23. Wang KKW, Yuen PW. Calpain: pharmacology and toxicology of calcium-dependent protease. Philadelphia, PA: Taylor & Francis; 1999.
24. Davis TL, Walker JR, Finerty PJ Jr, Mackenzie F, Newman EM, Dhe-Paganon S. The crystal structures of human calpains 1 and 9 imply diverse mechanisms of action and auto-inhibition. J Mol Biol 2007;366:216–29.
25. Yoshikawa Y, Mukai H, Hino F, Asada K, Kato I. Isolation of two novel genes, down-regulated in gastric cancer. Jpn J Cancer Res 2000; 91:459–63.
26. Vasan RS, Larson MG, Aragam J, Wang TJ, Mitchell GF, Kathiresan S, et al. Genome-wide association of echocardiographic dimensions, brachial artery endothelial function and treadmill exercise responses in the Framingham Heart Study. BMC Med Genet 2007;8 Suppl 1:S2.
27. ESE finder. 2012. [cited 2012 Jan 26]. Available from: http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home.