ARTICLE

Genome-wide association study identifies new susceptibility loci for epithelial ovarian cancer in Han Chinese women

Kexin Chen[1,*], Hongxia Ma[2,*], Lian Li[1,*], Rongyu Zang[3,*], Cheng Wang[2], Fengju Song[1], Tingyan Shi[3,4], Dianke Yu[5], Ming Yang[6], Wenqiong Xue[7], Juncheng Dai[2], Shuang Li[8], Hong Zheng[1], Chen Wu[5], Ying Zhang[7], Xiaohua Wu[3], Dake Li[9], Fengxia Xue[10], Haixin Li[1], Zhi Jiang[11], Jibin Liu[12], Yuexin Liu[13], Pei Li[1], Wen Tan[5], Jing Han[2], Jiang Jie[2], Quan Hao[14], Zhibin Hu[2], Dongxin Lin[5], Ding Ma[8], Weihua Jia[7], Hongbing Shen[2] & Qingyi Wei[4,15]

Ovarian cancer is the leading cause of death from gynaecological malignancies worldwide. Here we perform a three-stage genome-wide association study (GWAS) in Han Chinese women to identify risk genetic variants for epithelial ovarian cancer (EOC). We scan 900,015 single-nucleotide polymorphisms (SNPs) in 1,057 EOC cases and 1,191 controls in stage I, and replicate 41 SNPs ($P_{meta} < 10^{-4}$) in 960 EOC cases and 1,799 controls (stage II), and an additional 492 EOC cases and 1,004 controls (stage III). Finally, we identify two EOC susceptibility loci at 9q22.33 (rs1413299 in *COL15A1*, $P_{meta} = 1.88 \times 10^{-8}$) and 10p11.21 (rs1192691 near *ANKRD30A*, $P_{meta} = 2.62 \times 10^{-8}$), and two consistently replicated loci at 12q14.2 (rs11175194 in *SRGAP1*, $P_{meta} = 1.14 \times 10^{-7}$) and 9q34.2 (rs633862 near *ABO* and *SURF6*, $P_{meta} = 8.57 \times 10^{-7}$) ($P < 0.05$ in all three stages). These results may advance our understanding of genetic susceptibility to EOC.

[1] Department of Epidemiology and Biostatistics, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy of Tianjin, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China. [2] State Key Laboratory of Reproductive Medicine, Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, School of Public Health, Nanjing Medical University, Nanjing 211166, China. [3] Department of Gynecologic Oncology, Fudan University Shanghai Cancer Center, Shanghai 200032, China. [4] Cancer Institute, Fudan University Shanghai Cancer Center, Shanghai 200032, China. [5] Department of Etiology and Carcinogenesis, Cancer Institute and Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China. [6] College of Life Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China. [7] State Key Laboratory of Oncology in Southern China, Cancer Center, Sun Yat-sen University, Guangzhou, Guangdong 510060, China. [8] Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430030, China. [9] Department of Gynecology, Jiangsu Province Hospital of TCM, Affiliated Hospital of Nanjing University of TCM, Nanjing 210029, China. [10] Department of Gynecology and Obstetrics, Tianjin Medical University General Hospital, Tianjin 300052, China. [11] Department of Gynecologic Surgery, Jiangsu Cancer Hospital, Nanjing Medical University, Nanjing, Jiangsu 210009, China. [12] Department of Tissue Bank, Nantong Tumor Hospital, Nantong, Jiangsu 226361, China. [13] Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [14] Department of Urology, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China. [15] Duke Cancer Institute, Duke University Medical Center, Durham, North Carolina 27710, USA. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to K.C. (email: chenkexin@tijmu.edu.cn) or to D.M. (email: dingma424@126.com) or to Q.W. (email: weiqingyi@yahoo.com or qingyi.wei@duke.edu) or to H.S. (email: hbshen@njmu.edu.cn) or to W.J. (email: jiaweih@mail.sysu.edu.cn).

Ovarian cancer is one of the most common cancers among women and the leading cause of deaths from gynaecological malignancies in the world[1]. In addition to familial *BRCA1* and *BRCA2* mutations, there are other kinds of genetic risk factors, including common genetic variants of lower penetrance[2,3]. Molecular epidemiological studies have been conducted with the candidate gene approach to identify low penetrance susceptibility genes for ovarian cancer, many of which have showed inconsistent results[4–11]. Recent genome-wide association studies (GWASs) have reported 11 new single-nucleotide polymorphisms (SNPs) that are associated with ovarian cancer risk in European populations[12–17]. However, the results from the published GWASs have also demonstrated race- and ethnicity-specific cancer susceptibility[18]. Therefore, additional GWASs are needed to identify new ovarian cancer susceptibility genes, especially in non-European populations.

Here we report a three-stage GWAS that identifies new common genetic variants associated with ovarian cancer risk in Han Chinese women. We identify two SNPs that are significantly associated with ovarian cancer risk (rs1413299 in *COL15A1*, $P_{meta} = 1.88 \times 10^{-8}$ and rs1192691 near *ANKRD30A*, $P_{meta} = 2.62 \times 10^{-8}$) in Chinese women and two other consistently replicated loci (rs11175194 in *SRGAP1*, $P_{meta} = 1.14 \times 10^{-7}$ and rs633862 near *ABO* and *SURF6*, $P_{meta} = 8.57 \times 10^{-7}$). In addition, we also confirm rs9303542 at 17q21 from those 11 SNPs in previously reported GWASs of the Europeans.

Results

Association analyses. We conduct a three-stage GWAS in Han Chinese women (Table 1 and Supplementary Table 1). In the initial discovery stage, we perform a GWA scan in 1,057 epithelial ovarian cancer (EOC) cases and 1,191 age-matched healthy controls using the Illumina HumanOmniZhongHua-8 BeadChip, which contains 900,015 SNPs. After the standard quality control (QC) filtering, 710,714 SNPs and 2,216 individuals (1,044 cases and 1,172 controls) are included in the follow-up analyses (Supplementary Fig. 1). Principal component analysis (PCA) shows little evidence of population stratification in each (northern and southeastern) of the GWA scan populations (see Methods and Supplementary Fig. 2). We use logistic regression with adjustment for age and the first three eigenvectors to test the additive effect of a minor allele of each SNP. Meta-analysis was used to combine the results of the two populations in the GWA scan.

For the replication in stage II, we genotype 41 SNPs in an independent sample set of 960 EOC cases and 1,799 controls (Fig. 1 and Supplementary Data 1). The selection criteria of 41 SNPs were described in Methods. As a result, four SNPs associated with a significant cancer risk in stage II (rs1413299 ($P_{meta} = 4.00 \times 10^{-3}$) at 9q22.33, rs1192691 ($P_{meta} = 3.00 \times 10^{-3}$) at 10p11.21, rs11175194 ($P_{meta} = 3.58 \times 10^{-3}$) at 12q14.2 and rs633862 ($P_{meta} = 2.38 \times 10^{-2}$) at 9q34.2) ($P_{meta} < 0.05$) were further subjected to replication in stage III and consistently validated (Table 2). In the combined analysis of the five substudies (that is, Ia, Ib, IIa, IIb and III) in all the three stages, two SNPs (rs1192691, odds ratio (OR) = 0.81, 95% confidence interval (CI) = 0.75–0.87, $P_{meta} = 2.62 \times 10^{-8}$ and rs1413299, OR = 1.24, 95% CI = 1.15–1.33, $P_{meta} = 1.88 \times 10^{-8}$) reached the genome-wide significance ($P < 5 \times 10^{-8}$) without significant heterogeneity among all the stages. In addition, the other two consistently validated ($P < 0.05$ in all the three stages) SNPs also showed strong associations but short of the genome-wide significance (rs11175194, OR = 0.82, 95% CI = 0.76–0.88, $P_{meta} = 1.14 \times 10^{-7}$ and rs633862, OR = 0.83, 95% CI = 0.77–0.89, $P_{meta} = 8.57 \times 10^{-7}$) (Table 2 and Supplementary Fig. 3). These four SNPs have not reported in the previous GWASs of the Europeans, possibly due to racial diversity and differences in study power. For example, the minor allele frequency (MAF) of these four SNPs was quite different between European and Asian populations in the HapMap project (http://hapmap.ncbi.nlm.nih.gov/). In the previous GWASs, only a limited number of top-hit SNPs were validated in a large sample size. Therefore, additional GWASs need to be performed in other non-European populations.

Stratification analyses by histological subtype. We then evaluated the associations between these four SNPs and EOC risk after stratifying cases by histological subtype. We found no evidence of significant heterogeneity among the subtypes (Table 3). Using the imputation analyses based on data from the 1,000 Genomes Project (Phase I integrated variant set release, v3, http://www.1000genomes.org/), we tested for the risk associations of the SNPs (imputed info > 0.5, MAF ≥ 0.05) surrounding the lead SNPs in a 500-kb window, and we observed a series of significant signals around rs1413299 at 9q22.33 and rs11175194 at 12q14.2 (Supplementary Table 2 and Fig. 2).

We also evaluated the associations for SNPs in stage I with serous ovarian cancer risk, the dominant histologic subtype. We found that 104 SNPs were significantly associated with cancer risk among 594 patients with serous ovarian cancers and 1,172 controls ($P_{meta} < 1.00 \times 10^{-4}$). Of these 104 SNPs, 21 SNPs were also significantly associated with overall risk of ovarian cancer, and rs6784988 ($P_{serous} = 3.57 \times 10^{-7}$; $P_{total} = 2.23 \times 10^{-6}$) at 3q26.33 and rs7420064 ($P_{serous} = 3.87 \times 10^{-7}$; $P_{total} = 4.04 \times 10^{-6}$)

Table 1 | Three-stage study design of ovarian cancer GWAS in Chinese women.

	No. of SNPs	Genotyping method	Source of population	Controls		Cases	
				N	**Age (mean ± s.d.)**	**N**	**Age (mean ± s.d.)**
Stage I	900,015	Illumina HumanOmniZhongHua-8 BeadChip	Northern*	425	55.40 ± 9.49	429	54.80 ± 10.37
			Southeasthern†	747	52.86 ± 10.81	615	54.10 ± 10.43
			Stage I total	1,172	52.86 ± 10.81	1,044	54.39 ± 10.40
Stage II	51‡	Sequenom MassArray iPlex	Southeasthern†	899	54.75 ± 10.97	552	53.33 ± 10.18
			Southern§	900	51.87 ± 13.38	408	50.70 ± 12.49
			Stage II total	1,799	53.31 ± 12.32	960	52.21 ± 11.29
Stage III	51‡	Sequenom MassArray iPlex /Taqman	Northern‖	1,004	52.50 ± 11.44	492	53.69 ± 11.79

GWAS, genome-wide association study; SNP, single-nucleotide polymorphism.
*Northern Han population in stage I was recruited from Tianjin and Beijing in northern China.
†Southeastern Han population in stage I and stage II was recruited from Shanghai and Nanjing in southeastern China.
‡Forty-one newly identified SNPs and ten previously reported SNPs.
§Southern Han population in stage II was recruited from Guangzhou in southern china.
‖Northern Han in stage III was recruited from Tianjin and Shandong province in northern China.
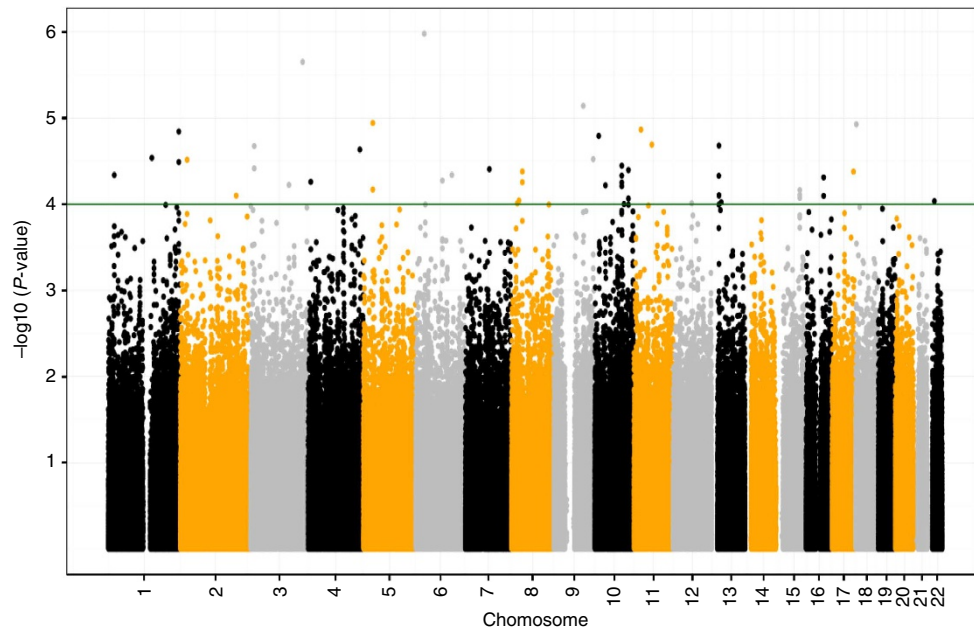
Figure 1 | Genome-wide association results for EOC in Han Chinese women. Manhattan plot showing the genome-wide association results for ovarian cancer. The association between 710,714 genotyped SNPs and ovarian cancer risk (1,044 ovarian cancer cases and 1,172 controls) was analysed. The green horizontal line represents $P = 1.0 \times 10^{-4}$.

Table 2 | Summary of GWAS scan and replication studies for the four SNPs.

SNP	Locus	Allele*	Proposed candidate	Stage	Population	MAF (cases/controls)	OR (95% CI)	P-value†	P_{meta}‡	P_{meta}§
rs1413299	9q22.33	A/C	COL15A1/ TGFBR1	Ia	N	0.48/0.38	1.53 (1.25–1.86)	2.47×10^{-5}	7.20×10^{-6}	1.88×10^{-8}
				Ib	SE	0.46/0.42	1.21 (1.04–1.41)	1.60×10^{-2}		
				IIa	S	0.44/0.41	1.14 (0.97–1.36)	1.21×10^{-1}	4.00×10^{-3}	
				IIb	SE	0.46/0.42	1.22 (1.04–1.41)	1.16×10^{-2}		
				III	N	0.42/0.38	1.20 (1.03–1.41)	1.89×10^{-2}		
rs1192691	10p11.21	C/A	NAMPTL/ ANKRD30A	Ia	N	0.43/0.46	0.89 (0.73–1.07)	2.20×10^{-1}	6.02×10^{-5}	2.62×10^{-8}
				Ib	SE	0.38/0.46	0.71 (0.60–0.83)	2.48×10^{-5}		
				IIa	S	0.38/0.41	0.89 (0.75–1.06)	1.82×10^{-1}	3.00×10^{-3}	
				IIb	SE	0.38/0.44	0.80 (0.68–0.93)	4.50×10^{-3}		
				III	N	0.41/0.46	0.81 (0.70–0.95)	8.74×10^{-3}		
rs11175194	12q14.2	G/A	SRGAP1	Ia	N	0.32/0.40	0.72 (0.59–0.88)	1.32×10^{-3}	9.75×10^{-5}	1.14×10^{-7}
				Ib	SE	0.34/0.38	0.82 (0.70–0.96)	1.61×10^{-2}		
				IIa	S	0.38/0.43	0.83 (0.70–0.98)	3.06×10^{-2}	3.58×10^{-3}	
				IIb	SE	0.36/0.40	0.86 (0.73–1.00)	5.01×10^{-2}		
				III	N	0.37/0.41	0.84 (0.72–0.98)	2.84×10^{-2}		
rs633862	9q34.2	A/G	ABO/SURF6	Ia	N	0.40/0.48	0.72 (0.59–0.87)	8.05×10^{-4}	2.99×10^{-5}	8.57×10^{-7}
				Ib	SE	0.41/0.45	0.80 (0.68–0.95)	8.35×10^{-3}		
				IIa	S	0.50/0.55	0.85 (0.72–1.01)	5.75×10^{-2}	2.38×10^{-2}	
				IIb	SE	0.41/0.43	0.90 (0.77–1.05)	1.92×10^{-1}		
				III	N	0.42/0.46	0.84 (0.72–0.98)	2.89×10^{-2}		

CI, confidence interval; GWAS, genome-wide association study; MAF, minor allele frequency; N, Northern population; OR, odds ratio; S, Southern population; SE, Southeastern population; SNP, single-nucleotide polymorphism.
*Major allele/minor allele.
†P-value, OR and 95% CI were assessed under an additive model using logistic regression adjusted for age and the first three principal components of population stratification in stage I, and adjusted for age in stages II and III.
‡Meta-analysis P-value of each stage.
§Meta-analysis P-value for all three stages.

at 2q13 were the top hits (Supplementary Data 2). Further validation studies are warranted to confirm these SNPs in large populations.

Association analyses of previously reported SNPs. Previous GWASs of ovarian cancer have reported a number of SNPs that were associated with ovarian cancer risk among populations with a European ancestry[12–17]. The present study in Han Chinese populations also tested the risk associations with the 11 previously reported SNPs from those GWAS analyses, but only 8 SNPs were included in our GWA scan or imputation analyses, of which four SNPs (rs2665390 at 3q25, rs9303542 at 17q21, rs8170 at 19p13.11 and rs757210 at 17q12) were found to be significantly associated with EOC risk in our stage I study.

Table 3 | Ovarian cancer risk associated with the four identified SNPs in stratified analysis by histological subtype.

SNP	Chr.	Proposed candidate gene	Pathology type	Case/control	MAF (cases/controls)*	OR (95% CI)†	P-value†	Q-P‡
rs1413299	9q22.33	COL15A1/TGFBR1	Serous	1,320/3,964	0.455/0.401	1.246 (1.139–1.377)	2.13E−06	1.03E−01
			Mucinous	167/3,964	0.437/0.401	1.127 (0.905–1.405)	2.97E−01	
			Endometriod	277/3,964	0.419/0.401	0.980 (0.803–1.209)	8.76E−01	
			Others	720/3,964	0.457/0.401	1.297 (1.150–1.448)	9.69E−06	
rs1192691	10p11.21	NAMPTL/ANKRD30A	Serous	1,310/3,898	0.395/0.445	0.795 (0.719–0.869)	1.13E−06	8.86E−01
			Mucinous	165/3,898	0.412/0.445	0.869 (0.691–1.094)	2.19E−01	
			Endometriod	273/3,898	0.396/0.445	0.827 (0.677–1.010)	6.90E−02	
			Others	719/3,898	0.399/0.445	0.827 (0.733–0.923)	1.16E−03	
rs11175194	12q14.2	SRGAP1	Serous	1,321/3,973	0.351/0.403	0.827 (0.748–0.905)	6.22E−05	5.91E−01
			Mucinous	167/3,973	0.359/0.403	0.827 (0.657–1.041)	1.10E−01	
			Endometriod	277/3,973	0.325/0.403	0.726 (0.589–0.896)	3.36E−03	
			Others	718/3,973	0.368/0.403	0.861 (0.763–0.970)	1.40E−02	
rs633862	9q34.2	ABO/SURF6	Serous	1,312/3,947	0.415/0.450	0.861 (0.787–0.951)	1.95E−03	4.56E−01
			Mucinous	166/3,947	0.452/0.450	0.869 (0.691–1.094)	2.24E−01	
			Endometriod	274/3,947	0.401/0.450	0.726 (0.589–0.896)	2.50E−03	
			Others	720/3,947	0.421/0.450	0.803 (0.712–0.896)	1.93E−04	

CI, confidence interval; OR, odds ratio; SNP, single-nucleotide polymorphism.
*Major allele/minor allele.
†P-value, OR and 95% CI of the meta-analysis for all populations.
‡P-value of Cochran's Q test.

We successfully re-designed the primers for 10 of the 11 SNPs and genotyped them in our stage II and III studies, and we identified 1 SNP (rs9303542 at 17q21) that was significantly associated with EOC risk (Supplementary Table 3) in both stage I and replication data sets. This rs9303542 SNP is intronic to SKAP1, which has strong homology to SRC oncogene at the carboxy-terminal end of their proteins. Although not significantly replicated, all of the tested SNPs were in the same direction of ORs published in previous GWASs. These discrepant results are probably due to the relatively lower statistical power of the present study and racial diversity.

Molecular analyses. The SNP may exert a long-range effect on the expression of genes upstream and downstream of the loci. A search of genes that were present within 1 Mb of the SNP revealed some potentially interesting genes (Fig. 2 and Supplementary Table 4). To gain further insight for the possible involvement of these genes in ovarian cancer pathophysiology, we examined whether expression levels of these genes are clinically relevant by using the comprehensive data from The Cancer Genome Atlas. Interestingly, these gene expression data suggested that COL15A1 ($P = 3.03 \times 10^{-3}$) gene expression levels were significantly different between tumours and normal tissues, and that the expression levels of TGFBR1 (transforming growth factor-β receptor type I) were also significantly associated with patient survival ($P = 7.00 \times 10^{-4}$). ANKRD30A expression data from The Cancer Genome Atlas suggested that there was no significant difference in the expression levels between ovarian cancer and normal tissues.

Discussion
In this three-stage GWAS of EOC, we identified two SNPs that were significantly associated ($P < 5 \times 10^{-8}$) with ovarian cancer risk in Chinese women and two other consistently replicated ($P < 5 \times 10^{-7}$) loci. The top-signal SNP, rs1413299, is in intron 6 of COL15A1 that encodes the α-chain of type XV collagen at 9q22.33, a member of the FACIT (fibril-associated collagens with interrupted helices) collagen family. Loss of the COL15A1 protein from the basement membranes/basement membrane zone was reported to promote tumour cell infiltration in human ductal

breast carcinoma cells and colon carcinomas[19,20]. In addition, COL15A was found to suppress tumorigenesis in a dose-dependent manner in human cervical carcinoma cell line[21]. Interestingly, rs1413299 is located in the DNase-seq, Chip-seq and Histone modification peak regions in dozens of cell lines, including lymphocyte-derived cell lines (GM12878) and breast cell lines (MCF-7; data from the Encode project, http://genome.ucsc.edu/ENCODE/)[22], indicating its potential biofeatures of an enhancer. By querying results from CTCF ChIA-PET in the Encode Project[23,24], we further found interaction signals in the region and nearby genes (TGFBR1 and GALNT2) in MCF-7, suggesting possible putative distant chromatin interactions bounded by CTCF. TGFBR1 is located 105 kb downstream of rs1413299, which is a central propagator of the TGF-β signalling pathway that is an important regulator of several important biological processes, including cell proliferation, differentiation, migration, apoptosis and matrix accumulation[25]. Participating in the activity of the TGF-β signalling pathway[26,27], TGFBR1 binds to TGF-β and forms a heterodimeric complex with TGFBR2, leading to phosphorylation and activation of SMAD2 and SMAD3. Furthermore, TGFBR1 expression was found to be markedly reduced in recurrent ovarian tumours[28]. In addition, the deletion of TGFBR1* 6A/9A, which was known to be associated with risk of multiple cancers, was also reported to be associated with ovarian cancer risk[29,30]. Therefore, additional studies are warranted to explore a possible role for TGFBR1 in the aetiology of EOC and a potential mechanistic link between rs1413299 and the functions of TGFBR1. Thus, we hypothesized that rs1413299 variants may have a different expression pattern of downstream genes through a long-range chromatin interaction and eventually influence the risk of ovarian cancer.

The second identified SNP, rs1192691, is located 245 kb upstream of exon 1 of ANKRD30A that encodes the ankyrin repeat domain 30A, and this gene is also known as NY-BR-1. NY-BR-1 is a breast cancer differentiation antigen and a potential target for cancer immunotherapy[31]. No associations with EOC risk and functional implications of NY-BR-1 in EOC have been reported in previous ovarian cancer studies; however, the oestrogen-response element-like sequences nearby the NY-BR-1 promoter region suggests that the NY-BR-1 expression may be
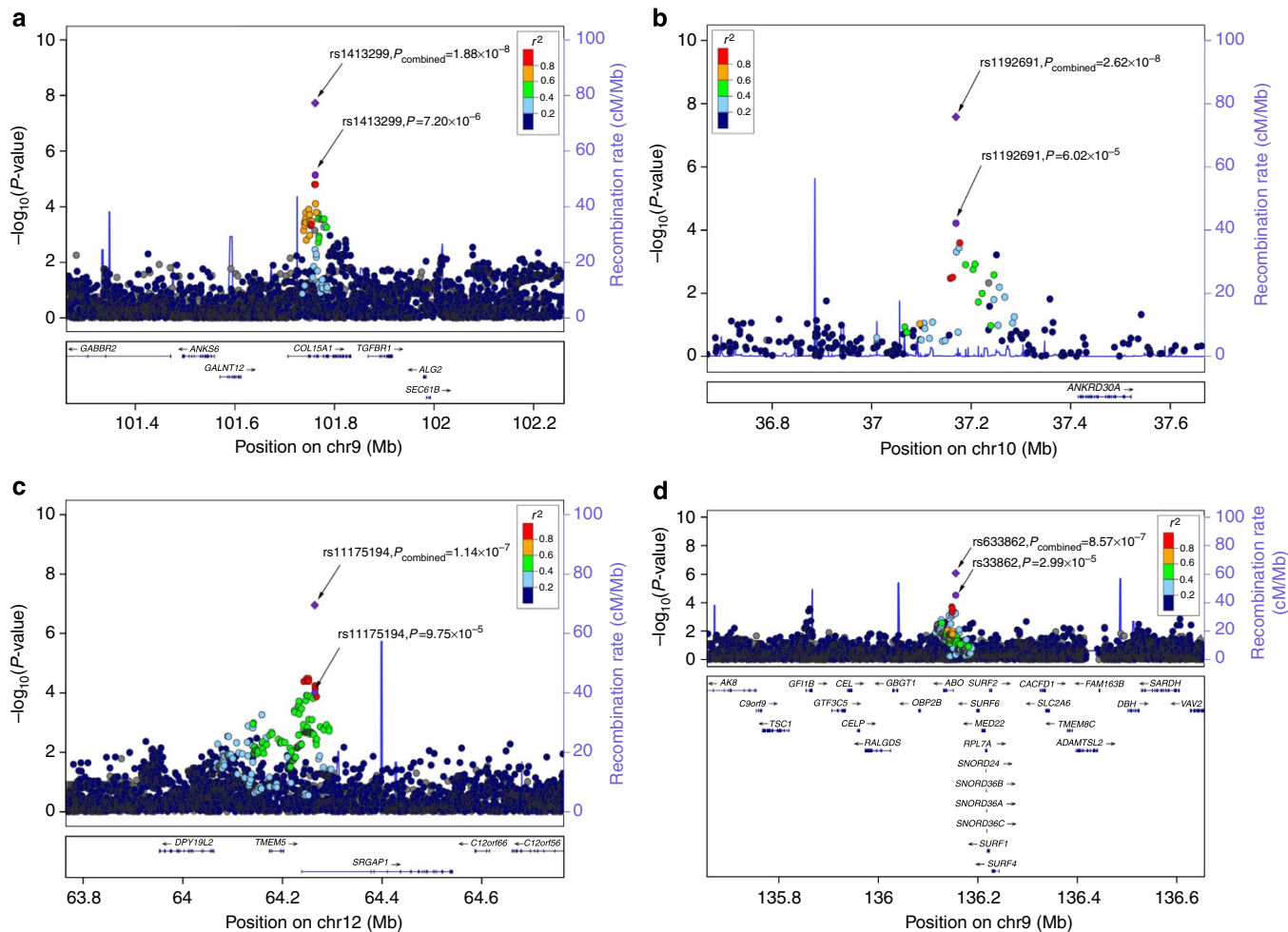
Figure 2 | Regional plots of the four ovarian cancer susceptibility loci. Regional plots of four newly discovered loci associated with the risk of ovarian cancer in Chinese women in the GWAS discovery stage. The results (− log10 P, P-value of the meta-analysis for all populations) are shown for SNPs in the region 500 kb up- or downstream of the marker SNP. The marker SNPs are shown as purple circle in stage I and as purple diamonds in combined stages; the LD values (r^2) between that SNPs and the most strongly associated SNP (diamond) are indicated by the heat scale. The genes within the region of interest are annotated and the directions of transcripts are shown in arrows.

regulated through oestrogen receptors[32]. Additional studies are warranted to further understand the implication of this gene in ovarian cancer.

The third SNP, rs11175194, is located in the first intron of *SRGAP1* that encodes the SLIT-ROBO Rho GTPase-activating protein 1. The Rho family small GTPases serve as molecular switches involved in several cellular progresses[33]. It is reported that overexpression of Rho GTPases in cancers was correlated with poor prognosis[34] and that SNPs in the *SRGAP1* gene were associated with papillary thyroid carcinoma[35]. As a subfamily of Rho GTPase-activating proteins, SRGAP1 is involved in regulation of the RhoA activity through interacting with CDC42 that has been proven to exert an important role in cancer development[36]. Interestingly, CDC42-positive macrophages may prevent malignant transformation of ovarian endometriosis[37].

The fourth SNP, rs633862, is located 5 kb upstream of *ABO* (ABO blood group). ABO determines the ABO blood group of an individual by modifying the oligosaccharides on cell surface glycoproteins. This SNP has a moderate linkage disequilibrium (LD) with rs8176719 ($r^2 = 0.57$, CHB (Han Chinese in Beijing) population in the 1,000 genome project), an SNP also named the O allele frame-shift mutation affecting amino acid 176. Furthermore, the minor allele of rs633862, suggested the O subgroup, has been found to exert a protective role in ovarian

cancer risk in the present study, which is consistent with the findings of previously published studies: A and B subgroups were associated with modestly increased risk of ovarian cancer[38–40].

In summary, in this GWAS of Han Chinese women on EOC susceptibility, we definitively (at $P < 5 \times 10^{-8}$) identified two new susceptibility loci at 9q22.33 and 10p11.21, and two other consistently replicated (at $P < 5 \times 10^{-7}$) loci at 12q14.2 and 9q34.2. In addition, we also confirmed 1 locus at chromosome 17q21 from those 11 SNPs in previously reported GWASs of the Europeans. Further studies with larger sample sizes are warranted to replicate our findings. Fine mapping around these new loci and related functional studies should be carried out to elucidate the underlying mechanism for the observed associations.

Methods

Study populations. We performed a three-stage GWAS in Han Chinese women for EOC susceptibility. A summary of all cases and controls in the study is provided in Supplementary Table 1. The GWA scan phase included 1,044 EOC patients and 1,172 controls (429 cases and 425 controls from Northern Chinese, and 615 cases and 747 controls from Southeastern Chinese), followed by two stages of validation (stage II: 408 cases and 900 controls of Southern Chinese; 552 cases and 899 controls of Southeastern Chinese, and stage III: 492 cases and 1,004 controls of Northern Chinese). There were six studies that were included in this three-stage study. Northern Chinese population was from TOCS (Tianjin Ovarian Cancer study), CAMSCH (Chinese Academy of Medical Sciences Cancer Hospital) and

BUCT (Beijing University of Chemical Technology). Southeastern Chinese population was from NOCS (Nanjing Ovarian Cancer study) and SOCS (Shanghai Ovarian Cancer study). Southern Chinese population was from GOCS (Guangzhou Ovarian Cancer study). All EOC cases were recruited in local hospitals and had the pathologically proven disease. Cancer-free controls were recruited in local hospitals for individuals receiving routine physical examinations or in the communities for those participating in the screening of non-communicable diseases. The cases and controls were frequency-matched for age in both the GWA scan stage and validations. At recruitment, informed consent was obtained from each subject, and this study was approved by the institutional review boards of each participating institution.

The TOCS enrolled patients with newly diagnosed and histologically confirmed EOC from Tianjin Medical University Cancer Hospital in Tianjin, China. Patients with a previous medical history of cancer, or previous radiotherapy or chemotherapy were excluded. Controls were recruited from cancer-free subjects who underwent regular health check-up during the same time when cases were recruited, and lived in the same neighbourhoods or nearby communities. The controls were frequency-matched to cases on age. All subjects were ethnic Chinese. The study was approved by the Committee on Human Research of Tianjin Medical University Cancer Hospital.

The CAMSCH study recruited cases from Beijing city and surrounding provinces of Beijing. The eligible patients with newly diagnosed, histopathologically confirmed and previously untreated (by radiotherapy or chemotherapy) ovarian cancer were accrued between January 2009 and June 2012 at the Cancer Hospital, Chinese Academy of Medical Sciences. All participants were of unrelated Han Chinese ethnicity. Each patient was interviewed for detailed information on demographic characteristics and lifestyles. Controls were recruited from the TOCS and were frequency-matched to cases on age. The study was approved by the Committee on Human Research of Cancer Hospital, Chinese Academy of Medical Sciences.

The BUCT enrolled cases and controls from the Shandong area of China. Cases were incident patients diagnosed with ovarian cancer at the Shandong Cancer Hospital. All of the cases had histologically confirmed ovarian cancer. Controls were healthy, cancer-free subjects selected from health examination clinics of the same hospital. All subjects were ethnic Chinese. The study was approved by the Committee on Human Research of Beijing University of Chemical Technology.

The NOCS EOC cases were recruited from Jiangsu Provincial Hospital of Traditional Chinese Medicine and Nantong Tumor Hospital, southeastern China. The criteria for the recruitment of EOC cases included: (1) Han Chinese; (2) without previous malignant tumours in any other organs; and (3) histopathologically confirmed diagnosis. Cancer-free controls were randomly selected from the health examination clinics of the same hospitals during the same time period of case recruitment and were frequency-matched to cases by age. Each individual was interviewed face-to-face by trained interviewers to get information on demographic characteristics and lifestyles. The study was approved by the Committee on Human Research of Nanjing Medical University.

The SOCS recruited cases from two hospitals, of which 1,167 ovarian cancer cases were consecutively recruited between March 2009 and August 2012 from Fudan University Shanghai Cancer Center (FUSCC), and the other 159 ovarian cancer cases were consecutively recruited between March 2012 and August 2012 at the Jiangsu Cancer Hospital (JCH). All tumours were histopathologically confirmed independently as primary epithelial ovarian carcinoma by two gynaecologic pathologists as routine diagnosis at FUSCC or at JCH. During an in-person interview, all potential subjects provided information about their demographics and known risk factors. The study was approved by the Committee on Human Research of Fudan University Shanghai Cancer Center.

The GOCS enrolled case and controls from Guangdong area of China. Cases were the EOC patients who underwent tumour resection between 2002 and 2012 in Sun Yat-sen University Cancer Center. The diagnosis was confirmed histologically in all cases. Controls were recruited from subjects who underwent regular health check-up during the same time period of cases in several cities including Guangzhou, Zhongshan and Sihui, and matched by age. The study was approved by the Committee on Human Research of Sun Yat-sen University Cancer Center.

Genotyping and QC in the GWAS. We genotyped a total of 900,015 SNPs in the GWA scan with 1,057 cases and 1,191 controls by using Illumina HumanOmniZhongHua-8 BeadChip. Before the association analysis, a systematic QC procedure was applied to the raw genotyping data to filter both unqualified SNPs and samples (see flow diagram in Supplementary Fig. 1). We excluded SNPs from further analysis if they (1) did not map to autosomal chromosomes; (2) had a low call rate in GWAS samples (<95%); (3) had MAF<0.05; and (4) were deviated from Hardy–Weinberg equilibrium ($P < 1.0 \times 10^{-5}$). We also removed individuals from further analysis if they (1) had an overall successful genotyping call rate<95%; (2) had sex discrepancies between the records and the genetically inferred data; and (3) were the unexpected duplicates or probable relatives (all PI_HAT>0.25). We detected population outliers by using a method based on PCA. A set of 106,963 common autosomal SNPs (MAF>0.25) with low LD ($r^2 < 0.20$) were used to identify population outliers in the samples that had passed QC, using the founders of the HapMap trios of Yoruba in Ibadan ($N = 90$), Utah residents of Northern and Western European ancestry ($N = 90$), CHB (Han Chinese in Beijing) ($N = 45$) and JPT (Japanese in Tokyo) ($N = 44$) as the

internal controls (Supplementary Fig. 2). The PCA showed that the cases and controls were genetically matched in both the Northern and Southeastern Chinese populations (Supplementary Fig. 2), and the genomic control inflation factor (λ) were 1.028 and 1.015, respectively, after adjustment for the first three principal components. After the QC process, a total of 1,044 cases and 1,172 controls with 710,714 SNPs were included in further analyses.

SNP selection and genotyping in the replication studies. We selected SNPs from the GWA scan phase for the stage II analysis based on the following criteria: (i) $P_{meta} < 1.0 \times 10^{-4}$; (ii) the same association direction between Northern and Southeastern studies; and (iii) only the SNP with the lowest P-value was selected when multiple SNPs were observed but in strong LD ($r^2 \geq 0.8$). Eight SNPs met the criteria for replication, but those that had a strong LD ($r^2 > 0.8$) with the selected SNPs were not included in replication (Supplementary Table 5). The SNPs that were significantly associated with EOC risk in the stage II analysis ($P < 0.05$) were further genotyped in stage III samples using Taqman. Genotyping in stages II and III was performed using the iPLEX MassARRAY platform (Sequenom, Inc.) of the 41 SNPs. The primers and probes are available on request. Laboratory technicians who performed genotyping experiments were blinded to case or control status. Five per cent of the samples were randomly selected for repeated genotyping as blind duplicates, and the reproducibility was 100%.

Statistical analysis. Associations between SNP genotypes and disease status were assessed in an additive model in PLINK v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/) using logistic regression modeling (1 degree of freedom) with adjustment for age and the first three principal components. For the meta-analysis, a fixed-effects model was used when the Cochran's Q statistic showed no heterogeneity (P for $Q > 0.05$); otherwise, a random-effects model (DerSimonian–Laird) was applied. Specifically, for the selection of SNPs used in the stage II analysis, we used the fixed-effects model that allowed us to include more candidate SNPs. Population structure was evaluated by the PCA in the software package EIGENSTRAT 4.2, and Manhattan plot of $-\log_{10} P$ was generated using package ggplot2 in R 2.15.1 (http://www.r-project.org./). The chromosome regions were plotted using an online tool, LocusZoom 1.1 (http://csg.sph.umich.edu/locuszoom/). We used the Shapeit v2 (Phasing step, http://www.shapeit.fr/) and IMPUTE2 (Imputation step) software (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) to impute untyped SNPs using the LD information from the 1000 Genomes Project (Phase I integrated variant set release, v3, across all 1,092 individuals, http://www.1000genomes.org/) and imputed SNPs were with info score >0.8. All other analyses were performed using R 2.15.1.

References

1. Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127,** 2893–2917 (2010).
2. Sueblinvong, T. & Carney, M. E. Current understanding of risk factors for ovarian cancer. *Curr. Treat. Options Oncol.* **10,** 67–81 (2009).
3. Lancaster, J. M. *et al.* Society of gynecologic oncologists education committee statement on risk assessment for inherited gynecologic cancer predispositions. *Gynecol. Oncol.* **107,** 159–162 (2007).
4. Doherty, J. A. *et al.* ESR1/SYNE1 polymorphism and invasive epithelial ovarian cancer risk: an Ovarian Cancer Association Consortium study. *Cancer Epidemiol. Biomarkers Prev.* **19,** 245–250 (2010).
5. Ratner, E. *et al.* A KRAS-variant in ovarian cancer acts as a genetic marker of cancer risk. *Cancer Res.* **70,** 6509–6515 (2010).
6. Braem, M. G., Schouten, L. J., Peeters, P. H., van den Brandt, P. A. & Onland-Moret, N. C. Genetic susceptibility to sporadic ovarian cancer: a systematic review. *Biochim. Biophys. Acta* **1816,** 132–146 (2011).
7. Notaridou, M. *et al.* Common alleles in candidate susceptibility genes associated with risk and development of epithelial ovarian cancer. *Int. J. Cancer* **128,** 2063–2074 (2011).
8. Cunningham, J. M. *et al.* Cell cycle genes and ovarian cancer susceptibility: a tagSNP analysis. *Br. J. Cancer* **101,** 1461–1468 (2009).
9. Lurie, G. *et al.* Vitamin D receptor rs2228570 polymorphism and invasive ovarian carcinoma risk: pooled analysis in five studies within the Ovarian Cancer Association Consortium. *J. Int. Cancer* **128,** 936–943 (2011).
10. Phelan, C. M. *et al.* Polymorphism in the GALNT1 gene and epithelial ovarian cancer in non-Hispanic white women: the Ovarian Cancer Association Consortium. *Cancer Epidemiol. Biomarkers Prev.* **19,** 600–604 (2010).
11. Schildkraut, J. M. *et al.* Association between DNA damage response and repair genes and risk of invasive serous ovarian cancer. *PLoS ONE* **5,** e10061 (2010).
12. Bolton, K. L. *et al.* Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nat. Genet.* **42,** 880–884 (2010).
13. Song, H. *et al.* A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. *Nat. Genet.* **41,** 996–1000 (2009).
14. Goode, E. L. *et al.* A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. *Nat. Genet.* **42,** 874–879 (2010).
15. Pharoah, P. D. *et al.* GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. *Nat. Genet.* **45,** 362–370 (2013).

16. Permuth-Wey, J. et al. Identification and molecular characterization of a new ovarian cancer susceptibility locus at 17q21.31. Nat. Commun. 4, 1627 (2013).

17. Bojesen, S. E. et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat. Genet. 45, 371–384 (2013).

18. Henderson, B. E., Lee, N. H., Seewaldt, V. & Shen, H. The influence of race and ethnicity on the biology of cancer. Nat. Rev. Cancer 12, 648–653 (2012).

19. Amenta, P. S. et al. Type XV collagen in human colonic adenocarcinomas has a different distribution than other basement membrane zone proteins. Hum. Pathol. 31, 359–366 (2000).

20. Amenta, P. S. et al. Loss of types XV and XIX collagen precedes basement membrane invasion in ductal carcinoma of the female breast. J. Pathol. 199, 298–308 (2003).

21. Harris, A., Harris, H. & Hollingsworth, M. A. Complete suppression of tumor formation by high levels of basement membrane collagen. Mol. Cancer Res. 5, 1241–1245 (2007).

22. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74 (2012).

23. Consortium, E. P. et al. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74 (2012).

24. Li, G. et al. ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. Genome Biol. 11, R22 (2010).

25. Gordon, K. J. & Blobe, G. C. Role of transforming growth factor-β superfamily signaling pathways in human disease. Biochim. Biophys. Acta 1782, 197–228 (2008).

26. Franzen, P. et al. Cloning of a TGF beta type I receptor that forms a heteromeric complex with the TGF beta type II receptor. Cell 75, 681–692 (1993).

27. Johnson, D. W., Qumsiyeh, M., Benkhalifa, M. & Marchuk, D. A. Assignment of human transforming growth factor-beta type I and type III receptor genes (TGFBR1 and TGFBR3) to 9q33-q34 and 1p32-p33, respectively. Genomics 28, 356–357 (1995).

28. Bristow, R. E., Baldwin, R. L., Yamada, S. D., Korc, M. & Karlan, B. Y. Altered expression of transforming growth factor-beta ligands and receptors in primary and recurrent ovarian carcinoma. Cancer 85, 658–668 (1999).

29. Liao, R. Y. et al. TGFBR1*6A/9A polymorphism and cancer risk: a meta-analysis of 13,662 cases and 14,147 controls. Mol. Biol. Rep. 37, 3227–3232 (2010).

30. Wang, Y. Q., Qi, X. W., Wang, F., Jiang, J. & Guo, Q. N. Association between TGFBR1 polymorphisms and cancer risk: a meta-analysis of 35 case-control studies. PLoS ONE 7, e42899 (2012).

31. Seil, I. et al. The differentiation antigen NY-BR-1 is a potential target for antibody-based therapies in breast cancer. Int. J. Cancer 120, 2635–2642 (2007).

32. Theurillat, J. P. et al. Distinct expression patterns of the immunogenic differentiation antigen NY-BR-1 in normal breast, testis and their malignant counterparts. Int. J. Cancer 122, 1585–1591 (2008).

33. Moon, S. Y. & Zheng, Y. Rho GTPase-activating proteins in cell regulation. Trends Cell Biol. 13, 13–22 (2003).

34. Fritz, G., Just, I. & Kaina, B. Rho GTPases are over-expressed in human tumors. Int. J. Cancer 81, 682–687 (1999).

35. He, H. et al. SRGAP1 is a candidate gene for papillary thyroid carcinoma susceptibility. J. Clin. Endocrinol. Metab. 98, E973–E980 (2013).

36. Wong, K. et al. Signal transduction in neuronal migration: roles of GTPase activating proteins and the small GTPase Cdc42 in the Slit-Robo pathway. Cell 107, 209–221 (2001).

37. Canet, B., Pons, C., Espinosa, I. & Prat, J. CDC42-positive macrophages may prevent malignant transformation of ovarian endometriosis. Hum. Pathol. 43, 720–725 (2012).

38. Yuzhalin, A. E. & Kutikhin, A. G. ABO and Rh blood groups in relation to ovarian, endometrial and cervical cancer risk among the population of South-East Siberia. Asian Pac. J. Cancer Prev. 13, 5091–5096 (2012).

39. Gates, M. A., Wolpin, B. M., Cramer, D. W., Hankinson, S. E. & Tworoger, S. S. ABO blood group and incidence of epithelial ovarian cancer. Int. J. Cancer 128, 482–486 (2011).

40. Poole, E. M. et al. ABO blood group and risk of epithelial ovarian cancer within the Ovarian Cancer Association Consortium. Cancer Causes Control 23, 1805–1810 (2012).

Acknowledgements

Author contributions

K.C., Q.W., H.S., W.J. and D.M. designed and supervised the overall study, obtained financial support and finalized the manuscript. H.M., L.L. and R.Z. coordinated the studies, performed statistical analyses, summarized results and drafted the manuscript. T.S., D.Y., M.Y., W.X., F.S., S.L., C.W., Y.Z., X.W., D.L., F.X., H.L., Z.J., J.L., W.T., J.H., J.J., Q.H. and D.L. recruited study subjects and prepared samples. J.D. conducted stage I analysis and SNP selection. Z.H. and C.W. conducted stage II and III analysis, and drafted the manuscript. H.Z. and P.L. conducted stage III genotyping. Y.L. conducted the expression analysis using The Cancer Genome Atlas data. All authors contributed to the final manuscript.

Additional information

Corrigendum: Genome-wide association study identifies new susceptibility loci for epithelial ovarian cancer in Han Chinese women

Kexin Chen, Hongxia Ma, Lian Li, Rongyu Zang, Cheng Wang, Fengju Song, Tingyan Shi, Dianke Yu, Ming Yang, Wenqiong Xue, Juncheng Dai, Shuang Li, Hong Zheng, Chen Wu, Ying Zhang, Xiaohua Wu, Dake Li, Fengxia Xue, Haixin Li, Zhi Jiang, Jibin Liu, Yuexin Liu, Pei Li, Wen Tan, Jing Han, Jiang Jie, Quan Hao, Zhibin Hu, Dongxin Lin, Ding Ma, Weihua Jia, Hongbing Shen & Qingyi Wei

The original version of the Supplementary Information attached to this Article contained an error in the numbering of the Supplementary Tables. The HTML has now been updated to include a corrected version of the Supplementary Information.