

Fine-scale mapping of chromosome 9q22.33 identifies candidate causal variant in ovarian cancer

Tongyu Xing^{Equal first author, 1}, Yanrui Zhao^{Equal first author, 1}, Lili Wang¹, Wei Geng¹, Wei Liu¹, Jingjing Zhou¹, Caiyun Huang¹, Wei Wang¹, Xinlei Chu¹, Ben Liu¹, Kexin Chen¹, Hong Zheng^{Corresp., 1}, Lian Li^{Corresp. 1}

¹ Department of Epidemiology and Biostatistics, Key Laboratory of Molecular Cancer Epidemiology, Key Laboratory of Prevention and Control of Human Major Diseases, Ministry of Education, Tianjin's Clinical Research Center for Cancer, National Clinical Research Center for Cancer, Tianjin, China

Corresponding Authors: Hong Zheng, Lian Li

Email address: zhengh64@aliyun.com, lilian@tmu.edu.cn

Ovarian cancer is a complex polygenic disease in which genetic factors play a significant role in disease etiology. A genome-wide association study (GWAS) identified a novel variant on chromosome 9q22.33 as a susceptibility locus for epithelial ovarian cancer (EOC) in the Han Chinese population. However, the underlying mechanism of this genomic region remained unknown. In this study, we conducted a fine-mapping analysis of 130kb regions, including 1,039 variants in 200 healthy women. Ten variants were selected to evaluate the association with EOC risk in 1,099 EOC cases and 1,591 controls. We identified two variants that were significantly associated with ovarian cancer risk (rs7027650, $P=1.91 \times 10^{-7}$; rs1889268, $P=3.71 \times 10^{-2}$). Expression quantitative trait locus (eQTL) analysis found that rs7027650 was significantly correlated with *COL15A1* gene expression ($P=0.009$). The Luciferase reporter gene assay confirmed that rs7027650 could interact with the promoter region of *COL15A1*, reducing its activity. Electrophoretic Mobility Shift Assay (EMSA) showed the allele-specific binding capacity of rs7027650. These findings revealed that rs7027650 could be a potential causal variant at 9q22.33 region and may regulate the expression level of *COL15A1*. This study offered insight into the molecular mechanism behind a potential causal variant that affects the risk of ovarian cancer.

Fine-scale mapping of chromosome 9q22.33 identifies candidate causal variant in ovarian cancer

Tongyu Xing^{1,†}, Yanrui Zhao^{1,†}, Lili Wang¹, Wei Geng¹, Wei Liu¹, Jingjing Zhou¹, Caiyun Huang¹, Wei Wang¹, Xinlei Chu¹, Ben Liu¹, Kexin Chen¹, Hong Zheng¹, Lian Li¹

¹ Department of Epidemiology and Biostatistics, Key Laboratory of Molecular Cancer Epidemiology, Key Laboratory of Prevention and Control of Human Major Diseases, Ministry of Education, Tianjin's Clinical Research Center for Cancer, National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin Medical University, Tianjin, China.

[†]These authors have contributed equally to this work.

Corresponding author:

Lian Li & Hong Zheng

Department of Epidemiology and Biostatistics, Key Laboratory of Molecular Cancer Epidemiology, Key Laboratory of Prevention and Control of Human Major Diseases, Ministry of Education, National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin Medical University, Huanhu Xi Road, TiyanBei, Hexi District, Tianjin, P.R. China 300060. Tel.: ++86 (0)2223372231; Fax: ++86 (0)2223372231

Email address: lilian@tmu.edu.cn, or zhengh64@aliyun.com

Abstract

Ovarian cancer is a complex polygenic disease in which genetic factors play a significant role in disease etiology. A genome-wide association study (GWAS) identified a novel variant on chromosome 9q22.33 as a susceptibility locus for epithelial ovarian cancer (EOC) in the Han Chinese population. However, the underlying mechanism of this genomic region remained unknown. In this study, we conducted a fine-mapping analysis of 130kb regions, including 1,039 variants in 200 healthy women. Ten variants were selected to evaluate the association with EOC risk in 1,099 EOC cases and 1,591 controls. We identified two variants that were significantly associated with ovarian cancer risk (rs7027650, $P=1.91 \times 10^{-7}$; rs1889268, $P=3.71 \times 10^{-2}$). Expression quantitative trait locus (eQTL) analysis found that rs7027650 was significantly correlated with *COL15A1* gene expression ($P=0.009$). The Luciferase reporter gene assay confirmed that rs7027650 could interact with the promoter region of *COL15A1*, reducing its activity. Electrophoretic Mobility Shift Assay (EMSA) showed the allele-specific binding capacity of rs7027650. These findings revealed that rs7027650 could be a potential causal variant at 9q22.33 region and may regulate the expression level of *COL15A1*. This study offered insight into the molecular mechanism behind a potential causal variant that affects the risk of ovarian cancer.

Introduction

Ovarian cancer is one of the most common gynecological cancers and is the second most fatal gynecological malignancy in Chinese women (Sung et al., 2021). As a complex polygenic disease (Dareng et al., 2022; Stewart et al., 2019), genetic factors play a significant role in

ovarian cancer etiology (Jones et al., 2017). So far, over **then** 40 susceptibility loci of ovarian cancer have been identified by genome wide association **studies** (Glubb et al., 2020; Kar et al., 2018; Lawrenson et al., 2019; Manichaikul et al., 2020). However, the mechanisms through which the most of these loci contribute to **the** ovarian cancer carcinogenesis remain unclear. A functional study was conducted at 9p22.2 locus, which was the first and most significant risk locus for ovarian cancer identified through GWAS in European women. Buckley et al. identified multiple candidate causal variants (CCVs) at the locus and these variants may be mediated by changes in a transcriptional regulatory network of several regulatory elements that act on *BNC2*, potentially influencing ovarian cancer susceptibility (Buckley et al., 2019). In our previous **genome wide association study**, we determined that rs1413299 at 9q22.33 was significantly associated with the risk of epithelial ovarian cancer (EOC) in Chinese women (Chen et al., 2014). SNP rs1413299 is located at intron six of *COL15A1*, which encodes the alpha chain of collagen type XV (COL15). COL15, a member of the FACIT (fibril-associated collagens with interrupted triple helices) **collagen** family, is a proteoglycan closely associated with the basement membrane, which has a broad tissue distribution (Breteaud et al., 2020; Clementz & Harris, 2013). *COL15A1* gene expression levels were significantly different between ovarian cancer tumor tissues and normal tissues in The Cancer Genome Atlas (TCGA) data (Chen et al., 2014). However, the effect of the variants in this region on the occurrence and development of ovarian cancer is still unclear. To identify the potential causal variants and additional novel variants at 9q22.33, we conducted a target sequencing followed by a validation study in a large sample size. We also performed a preliminary functional exploration to investigate the molecular mechanism of the susceptibility region.

Materials and methods

Study subjects were recruited from the Tianjin and Guangdong provinces in China. Cases were newly diagnosed and histologically confirmed to be epithelial ovarian cancer. Controls were selected from healthy and cancer free women who underwent regular health check-ups or participants in the cancer screening program. Informed consent was obtained from each subject when **recruitment**. This study was approved by the Medical Ethics Committee of Tianjin Medical University Cancer Institute and Hospital (Approval number: bc201810). 200 cancer free women were selected for target sequencing and a total of 1,099 epithelial ovarian cases and 1,591 age matched controls were included in the association study.

Target region sequencing and genotyping

We explored linkage disequilibrium (LD) structure around the target SNP rs1413299 using the HapMap Project database (Release 3 version 27, CHB+JPT) by Haploview (Barrett et al., 2005). A total of 1,039 variants (Total Bases: 130kb, Overall coverage: 96.9%) were successfully designed within the LD block interval at 9q22.33 (chr9:101,730,000-101,860,000, hg19). Sequencing was conducted using Ion Torrent PGM Next Generation Sequencing (Thermo Fischer Scientific) and four variants **which** with minor allele frequency (MAF) > 5% and in high LD with the target SNP rs1413299 ($r^2 > 0.8$) were selected. In order to identify additional new variants, we also included **the** SNPs with $P < 1 \times 10^{-3}$ in the 1Mb region surrounding rs1413299 from previous GWAS results and the SNPs with high LD ($r^2 > 0.7$) to rs1413299 from SNAP (<https://data.broadinstitute.org/mpg/snp/snp/>). Finally, ten variants were included for the subsequence validation study.

Genotyping of the ten SNPs (rs10819587, rs10988451, rs1413298, rs1572136, rs1889268,

rs4743305, rs7021675, rs7027650, rs7031588, and rs73503719) was performed using the iPLEX MassARRAY platform (Sequenom, Inc.) according to manufacturer's protocols.

Functional annotation

The functional annotations of the variants were performed using several kinds of sources. Epigenomic data was obtained from the Roadmap Epigenomic project (<https://egg2.wustl.edu/roadmap/data/byFileType/signal/consolidated/macs2signal/foldChange/>), including DNase-Seq data and Chip-seq profiles for H3K27ac and H3K4me1 histone modification of primary ovary tissue E097. The ATAC-seq profiles for three ovarian cancer cell lines (OVCA432, DOV13, and SKOV3) were from our previous study (Dai et al., 2021). The construction method of the ATAC-Seq library and analysis method have been previously described previously (Dai et al., 2021). We also used HaploRegv4.1 (<http://www.broadinstitute.org/mammals/haploreg/hap-loreg.php>) (Ward & Kellis, 2012) and RegulomeDB (<http://regulomedb.org/>) (Boyle et al., 2012) to explore the possible functions of the variants.

eQTL analysis

We obtained 102 ovarian tumor tissues from the cancer biobank of Tianjin Medical University Cancer Hospital and extracted the total RNA from these tissues using the standard Trizol method. Real-time quantitative PCR was used to measure the expression of *COL15A1* gene expression. The eQTL analysis was conducted using linear regression, preformed with the R package Matrix eQTL, and adjusted for age.

Cell lines

The SKOV3, 293T, and OVCA433 cell lines were preserved in the Laboratory of Epidemiology and Biostatistics at Tianjin Medical University Cancer Hospital. SKOV3 was cultured in RPMI 1640 medium with 10% FBS. 293T and OVCA433 were cultured in DMEM medium with 10% FBS, and incubated in a cell incubator at 37°C with 5% CO₂.

Luciferase assay

We downloaded the target DNA fragment containing rs7027650 from the NCBI (National Center for Biotechnology Information (nih.gov)), and designed the upstream and downstream primers of the target fragment. PCR amplification was performed using the whole genome DNA of the ovarian cancer cell line SKOV3 as a template to obtain a 500bp DNA fragment containing the SNP rs7027650. Using the PGL3 Basic-*COL15A1* Promoter as a control, we inserted DNA fragment-A and DNA fragment-T into the PGL3 Basic-*COL15A1* Promoter plasmid vector to obtain PGL3 Basic-A-*COL15A1* Promoter and PGL3 Basic-T-*COL15A1* Promoter plasmid vector. The three plasmids were then transfected into SKOV3 ovarian cancer cell line or 293T cell line, respectively, and luciferase protein activities were detected after 30 hours later. Luciferase activities were measured using the Dual Luciferase Assay Kit (Promega, Madison, WI).

Electrophoretic mobility shift assay (EMSA)

The probes were commercially synthesized and labeled with biotin at the 5' end (Table S1), and nuclear protein was obtained from ovarian cancer cell line OVCA433. EMSA assays were performed using the LightShift Chemiluminescent EMSAs Kit (Thermo Fisher Scientific) according to the manufacturer's instructions and as described previously (Dai et al., 2021).

Statistical analysis

The association between the variants and ovarian cancer risk was estimated using logistic regression adjusted for age. Linkage disequilibrium between the variants and lead SNP rs1413299 was evaluated by PLINK 1.9. The figure was made using R software version 4.0.4

and the difference was estimated using a Student's t-test. The P values were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Targeted sequencing of the 9q22.33 region

Variants for fine-mapping were selected by Haploview software based on the Asian population (CHB+JPT) data from Hapmap Release3 version27. We obtained 1,039 variants at 130 kb (Chr9: 101,730,000-101,860,000) (Fig. S1) of the 9q22.33 region and performed targeted sequencing in 200 healthy women. We identified four variants that have high LD ($r^2 > 0.70$) with rs1413299 and MAF > 0.05 in target sequencing. We also selected six variants according to the results of the previous GWAS (Chen et al., 2014) and analyzed these variants with by SNAP (details in method). In total, ten variants were available for genotyping using Sequenom MassArray and were used for subsequent association analysis (Table S2-S4 and Materials and methods).

Association analysis of the variants and ovarian cancer risk

Ten variants were genotyped in 1,099 EOC cases and 1,591 healthy controls (Table S5) using the Sequenom iPLEX platforms. We identified that rs7027650 ($P = 1.91 \times 10^{-7}$, OR=1.31, 95%CI=1.21-1.41) and rs1889268 ($P = 3.71 \times 10^{-2}$, OR=1.14, 95%CI=1.02-1.26) were significantly associated with ovarian cancer risk in additive model (Table 1 and Table S6). The association between rs7027650 and ovarian cancer risk was also significant in both the dominant ($P = 5.90 \times 10^{-10}$, OR=1.80, 95%CI=1.62-1.98) and recessive models ($P = 4.88 \times 10^{-3}$, OR=1.26, 95%CI=1.10-1.42). For rs1889268, the significant association was also observed in recessive model ($P = 2.32 \times 10^{-2}$, OR=1.20, 95%CI=1.04-1.35). We then performed stratification analysis by histological subtype, finding that rs7027650 was significantly associated with three major histotypes of epithelial ovarian cancer, (Serous OC: $P = 1.33 \times 10^{-3}$; Endometrioid OC: $P = 5.57 \times 10^{-6}$; Mucinous OC: $P = 1.27 \times 10^{-3}$). Additionally, we found that rs1889268 was significantly associated with Serous OC ($P = 4.99 \times 10^{-2}$), and rs1413298 was associated with Endometrioid OC ($P = 8.72 \times 10^{-3}$) (Fig. 1 and Table S7)

Functional annotation and eQTL analysis of candidate causal variants

We annotated these SNPs using publicly available data from the Roadmap project (DNase-Seq data and Chip-seq profiles for H3K27ac and H3K4me1 histone modification of primary ovary tissue) and in-house ATAC-seq profiles in three ovarian cancer cell lines (OVCA432, DOV13, and SKOV3). We found that rs7027650, the most significant variant, was overlapped with DNase hypersensitivity peaks in normal primary ovarian tissue which represent the open chromosome region. And the ATAC-seq analyses showed that rs1889268 and rs1413298 resided in open chromatin regions in ovarian cancer cells (Fig. 2A).

To identify candidate functional variants, we then performed expression quantitative trait locus (eQTL) analysis in 102 ovarian tumor tissues from Chinese women. We found that the allele T of rs7027650 was significantly associated with the decreased mRNA expression of *COL15A1* ($P = 0.009$) (Fig. 2B and Fig. S2). However, there was no significant association between the other two variants and expression of the *COL15A1* gene (rs1889268: $P = 0.072$; rs1413298: $P = 0.193$). We have also explored the GTEx project data, we found that there were several variants associated with gene expression. The most significant variant, 7027650, was associated with *COL15A1* expression ($P < 0.05$) in Thyroid and Nerve-Tibial tissues. However, the significant correlation was not observed in normal ovary tissues (Table S8).

Validation of the candidate causal variants

We performed the luciferase assay and EMSA to explore the possible function of potential

variants. In order to explore whether the variant interacts with the promoter of *COL15A1*, we constructed the luciferase reporter vectors containing the rs7027650-A allele or rs7027650-T allele and transfected these plasmids into the SKOV3 and 293T cell lines, respectively. Compared to the fragment containing allele-A, the transcription activity of *COL15A1* in the fragment with allele-T was lower in both SKOV3 and 293T cells (Fig. 3A). The results suggested that the gene fragment containing rs7027650 could interact with the promoter region of *COL15A1* and may reduce its activity. Next, we performed EMSAs using the probes containing allele A, allele T, and controls. The results showed differences in the allele-specific protein binding capacity of rs7027650 (Fig. 3B). These results indicate that rs7027650 may be a potential functional SNP.

Discussion

In this study, we conducted targeted sequencing of 130kb on the *COL15A1* gene at the 9q22.33 region and subsequently validated the results in 1,099 EOC cases and 1,591 controls. We identified an unreported genetic variant, rs7027650, that was significantly associated with ovarian cancer risk and significantly correlated with the expression of *COL15A1* in eQTL analysis. Preliminary molecular analysis revealed that rs7027650 may regulate the expression level of *COL15A1*.

In the previous three-stage GWAS study of ovarian cancer in Han Chinese women, we discovered two new susceptibility loci associated with the risk of epithelial ovarian cancer (Chen et al., 2014). SNP rs1413299 at the 9q22.33 region was the most significant SNP in this GWAS study. However, the functional variant at this locus and the molecular mechanism accounting for the risk of ovarian cancer were still unknown. Therefore, we performed fine-scale mapping, validation in a large case control cohort, and eQTL analysis to explore the candidate functional variants at the susceptibility risk loci 9q22.33. We determined that rs7027650 located at intron two of the *COL15A1* gene was the potential candidate causal variant in this region.

Functional annotation revealed that rs7027650 overlaps with DNase I hypersensitivity sites markers in normal ovarian tissue and ovarian cancer cell lines, suggesting that this variant resides in open chromatin region and may serve a transcriptional regulation function. The Luciferase assay revealed that the gene fragments with rs7027650 have lower activity in SKOV3 and 293T cells indicating that the SNP could interact with the promoter region of *COL15A1*. The EMSA showed that rs7027650 exhibited different protein binding capacities between allele-A and allele-T. Collectively, these molecular analyses and experiments suggested that SNP rs7027650 may be a functional variant, potentially affecting in *COL15A1* transcription regulation and may affect ovarian cancer carcinogenesis.

The *COL15A1* gene encodes the alpha chain of type COL15 which is a member of the FACIT collagen family. COL15 is a structurally complex macromolecule with many unique features and a range of biological properties. Light microscopy experiments have reported that COL15 is localized to endothelial, muscle, nerve, adipose, and most epithelial basement membrane regions of human tissues. This wide distribution suggests COL15 has an adhesive role, connecting the basement membrane to connective tissue (Amenta et al., 2005; Clementz & Harris, 2013). In breast cancer, collagen types XV and XIX were lost early in the development of invasive tumors. The disappearance of these proteins may signal remodeling of the extracellular matrix to promote tumor cell infiltration (Amenta et al., 2003). In a transgenic *MMTV-PyMT* mouse mammary carcinoma model, inactivation of *COL15A1* modulated the tumor extracellular matrix and to increased mammary tumor growth (Martinez-Nieto et al., 2021). Type XV collagen

played a role in the adherence of the basement membrane to surrounding connective tissue and it may be associated with the tumorigenesis of keratinocytes and melanocytes (Fukushige et al., 2005). In cervical carcinoma cells, collagen XV functioned as a dose-dependent suppressor of tumorigenicity (Harris et al., 2007), and collagen XV was found to inhibit the adhesion and migration of fibrosarcoma cells when present in fibronectin-containing matrices (Hurskainen et al., 2010). Kouji Kimura et.al. reported that immunohistochemical examination of collagen XV and *COL15A1* mRNA both showed increased expression in tumoral regions than in non-tumoral regions in human hepatocellular carcinoma (HCC), and collagen XV was considered to be the factor that contributed to the capillarization of HCC (Kimura et al., 2016). Collagen XV was also reported to function as a metastasis inhibitor in HCC by regulating the discoidin domain receptor 1 (DDR1)-Snail/Slug axis, thus regulating epithelial mesenchymal transition (EMT) (Yao et al., 2022). The *COL15A1* gene has been reported to be upregulated by the interaction between Clusterin (CLU) and myocyte enhancer factor 2A (MEF2A). This upregulation is associated with blocking the EMT process, inhibiting the invasion and metastasis of testicular seminoma (Cui et al., 2021). In ovarian cancer, *COL15A1* gene expression was significantly lower in tumor tissues than normal tissues in TCGA data (Chen et al., 2014), suggesting *COL15A1* had a potential tumor suppressor role in ovarian cancer. Further experimental investigation is warranted to reveal the underlying molecular mechanism of *COL15A1* in ovarian cancer carcinogenesis. There are still some limitations to this study. Due to insufficient samples, we did not validate our results in another larger sample. We also did not find any significant association between rs7027650 and ovarian cancer risk in the European population from the Ovarian Cancer Association Consortium (OCAC) data. The inconsistent result with European population may be due to the race- and ethnicity heterogeneity of cancer susceptibility (Henderson et al., 2012). The ethnicity specific cancer susceptibility was also observed at several other kinds of cancers, such as colorectal cancer and prostate cancer etc. (Lu et al., 2019; Rebbeck, 2017). Additionally, this study only conducted a preliminary exploration of the potential function of the causal variant, and more experiments are needed to provide additional evidence to support the regulatory function of this variant.

Conclusions

In conclusion, we identified a potential causal variant within the chromosome 9q22.33 region, rs7027650, which was associated with ovarian cancer susceptibility. We also demonstrated that the potential causal rs7027650 may regulate *COL15A1* gene expression, offering suggestions for further functional studies to clarify the underlying biological mechanism in ovarian cancer etiology. Nonetheless, the transcriptional regulation mechanism of rs7027650 on *COL15A1* remains unclear and needs to be clarified through further functional experiments.

References

- Amenta PS, Hadad S, Lee MT, Barnard N, Li D, and Myers JC. 2003. Loss of types XV and XIX collagen precedes basement membrane invasion in ductal carcinoma of the female breast. *J Pathol* 199:298-308. 10.1002/path.1303
- Amenta PS, Scivoletti NA, Newman MD, Sciancalepore JP, Li D, and Myers JC. 2005. Proteoglycan-collagen XV in human tissues is seen linking banded collagen fibers subjacent to the basement membrane. *Journal of Histochemistry* 53:165-176.
- Barrett JC, Fry B, Maller J, and Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263-265. 10.1093/bioinformatics/bth457
- Boyle AP, Hong EL, Hariharan M, Yong C, and Snyder M. 2012. Annotation of functional variation in personal

- genomes using RegulomeDB. *Genome Research* 22:1790-1797.
- Breitaud S, Guillon E, Karppinen SM, Pihlajaniemi T, and Ruggiero F. 2020. Collagen XV, a multifaceted multiplexin present across tissues and species. *Matrix Biol Plus* 6-7:100023. 10.1016/j.mbplus.2020.100023
- Buckley MA, Woods NT, Tyrer JP, Mendoza-Fandino G, Lawrenson K, Hazelett DJ, Najafabadi HS, Gjyshi A, Carvalho RS, Lyra PC, Jr., Coetzee SG, Shen HC, Yang AW, Earp MA, Yoder SJ, Risch H, Chenevix-Trench G, Ramus SJ, Phelan CM, Coetzee GA, Noushmehr H, Hughes TR, Sellers TA, Goode EL, Pharoah PD, Gayther SA, Monteiro ANA, and Ovarian Cancer Association C. 2019. Functional Analysis and Fine Mapping of the 9p22.2 Ovarian Cancer Susceptibility Locus. *Cancer Res* 79:467-481. 10.1158/0008-5472.CAN-17-3864
- Chen K, Ma H, Li L, Zang R, Wang C, Song F, Shi T, Yu D, Yang M, Xue W, Dai J, Li S, Zheng H, Wu C, Zhang Y, Wu X, Li D, Xue F, Li H, Jiang Z, Liu J, Liu Y, Li P, Tan W, Han J, Jie J, Hao Q, Hu Z, Lin D, Ma D, Jia W, Shen H, and Wei Q. 2014. Genome-wide association study identifies new susceptibility loci for epithelial ovarian cancer in Han Chinese women. *Nat Commun* 5:4682. 10.1038/ncomms5682
- Clementz AG, and Harris A. 2013. Collagen XV: exploring its structure and role within the tumor microenvironment. *Mol Cancer Res* 11:1481-1486. 10.1158/1541-7786.MCR-12-0662
- Cui Y, Miao C, Liu S, Tang J, Zhang J, Bu H, Wang Y, Liang C, Bao M, Hou C, Wu J, Chen X, Zhang X, Wang Z, and Liu B. 2021. Clusterin suppresses invasion and metastasis of testicular seminoma by upregulating COL15a1. *Mol Ther Nucleic Acids* 26:1336-1350. 10.1016/j.omtn.2021.11.004
- Dai H, Chu X, Liang Q, Wang M, Li L, Zhou Y, Zheng Z, Wang W, Wang Z, Li H, Wang J, Zheng H, Zhao Y, Liu L, Yao H, Luo M, Wang Q, Kang S, Li Y, Wang K, Song F, Zhang R, Wu X, Cheng X, Zhang W, Wei Q, Li MJ, and Chen K. 2021. Genome-wide association and functional interrogation identified a variant at 3p26.1 modulating ovarian cancer survival among Chinese women. *Cell Discovery* 7. 10.1038/s41421-021-00342-6
- Dareng EO, Tyrer JP, Barnes DR, Jones MR, Yang X, Aben KKH, Adank MA, Agata S, Andrulis IL, Anton-Culver H, Antonenkova NN, Aravantinos G, Arun BK, Augustinsson A, Balmana J, Bandera EV, Barkardottir RB, Barrowdale D, Beckmann MW, Beeghly-Fadiel A, Benitez J, Bermisheva M, Bernardini MQ, Bjorge L, Black A, Bogdanova NV, Bonanni B, Borg A, Brenton JD, Budzylowska A, Butzow R, Buys SS, Cai H, Caligo MA, Campbell I, Cannioto R, Cassingham H, Chang-Claude J, Chanock SJ, Chen K, Chiew YE, Chung WK, Claes KBM, Colonna S, Collaborators GS, Collaborators G-HS, Collaborators E, Cook LS, Couch FJ, Daly MB, Dao F, Davies E, de la Hoya M, de Putter R, Dennis J, DePersia A, Devilee P, Diez O, Ding YC, Doherty JA, Domchek SM, Dork T, du Bois A, Durst M, Eccles DM, Eliassen HA, Engel C, Evans GD, Fasching PA, Flanagan JM, Fortner RT, Machackova E, Friedman E, Ganz PA, Garber J, Gensini F, Giles GG, Glendon G, Godwin AK, Goodman MT, Greene MH, Gronwald J, Group OS, Group A, Hahnen E, Haiman CA, Hakansson N, Hamann U, Hansen TVO, Harris HR, Hartman M, Heitz F, Hildebrandt MAT, Hogdall E, Hogdall CK, Hopper JL, Huang RY, Huff C, Hulick PJ, Huntsman DG, Imyanitov EN, Investigators KC, Investigators H, Isaacs C, Jakubowska A, James PA, Janavicius R, Jensen A, Johannsson OT, John EM, Jones ME, Kang D, Karlan BY, Karnezis A, Kelemen LE, Khusnutdinova E, Kiemeny LA, Kim BG, Kjaer SK, Komenaka I, Kupryjanczyk J, Kurian AW, Kwong A, Lambrechts D, Larson MC, Lazaro C, Le ND, Leslie G, Lester J, Lesueur F, Levine DA, Li L, Li J, Loud JT, Lu KH, Lubinski J, Mai PL, Manoukian S, Marks JR, Matsuno RK, Matsuo K, May T, McGuffog L, McLaughlin JR, McNeish IA, Mebirouk N, Menon U, Miller A, Milne RL, Minlikeeva A, Modugno F, Montagna M, Moysich KB, Munro E, Nathanson KL, Neuhausen SL, Nevanlinna H, Yie JNY, Nielsen HR, Nielsen FC, Nikitina-Zake L, Odunsi K, Offit K, Olah E, Olbrecht S, Olopade OI, Olson SH, Olsson H, Osorio A, Papi L, Park SK, Parsons MT, Pathak H, Pedersen IS, Peixoto A, Pejovic T, Perez-Segura P, Permuth JB, Peshkin B, Peterlongo P, Piskorz A, Prokofyeva D, Radice P, Rantala J, Riggan MJ, Risch HA, Rodriguez-Antona C, Ross E, Rossing MA, Runnebaum I, Sandler DP, Santamarina M, Soucy P, Schmutzler RK, Setiawan VW, Shan K, Sieh W, Simard J, Singer CF, Sokolenko AP, Song H, Southey MC, Steed H, Stoppa-Lyonnet D, Sutphen R, Swerdlow AJ, Tan YY, Teixeira MR, Teo SH, Terry KL, Terry MB, Consortium O, Consortium C, Thomassen M, Thompson PJ, Thomsen LCV, Thull DL, Tischkowitz M, Titus L, Toland AE, Torres D, Trabert B, Travis R, Tung N, Tworoger SS, Valen E, van Altena AM, van der Hout AH, Van Nieuwenhuysen E, van Rensburg EJ, Vega A, Edwards DV, Vierkant RA, Wang F, Wappenschmidt B, Webb PM, Weinberg CR, Weitzel JN, Wentzensen N, White E, Whittemore AS, Winham SJ, Wolk A, Woo YL, Wu AH, Yan L, Yannoukakos D, Zavaglia KM, Zheng W, Ziogas A, Zorn KK, Kleibl Z, Easton D, Lawrenson K, DeFazio A, Sellers TA, Ramus SJ, Pearce CL, Monteiro AN, Cunningham J, Goode EL, Schildkraut JM, Berchuck A, Chenevix-Trench G, Gayther SA, Antoniou AC, and Pharoah PDP. 2022. Polygenic risk modeling for prediction of epithelial ovarian cancer risk. *Eur J Hum Genet* 30:349-362. 10.1038/s41431-021-00987-7

- Fukushige T, Kanekura T, Ohuchi E, Shinya T, and Kanzaki T. 2005. Immunohistochemical studies comparing the localization of type XV collagen in normal human skin and skin tumors with that of type IV collagen. *Journal of Dermatology* 32:74-83.
- Glubb DM, Thompson DJ, Aben KK, Alsulimani A, O'Mara TAJCEB, and Prevention. 2020. Cross-Cancer Genome-Wide Association Study of Endometrial Cancer and Epithelial Ovarian Cancer Identifies Genetic Risk Regions Associated with Risk of Both Cancers. 30:cebp.0739.2020.
- Harris A, Harris H, and Hollingsworth MA. 2007. Complete suppression of tumor formation by high levels of basement membrane collagen. *Mol Cancer Res* 5:1241-1245. 10.1158/1541-7786.MCR-07-0200
- Henderson BE, Lee NH, Seewaldt V, and Shen H. 2012. The influence of race and ethnicity on the biology of cancer. *Nat Rev Cancer* 12:648-653. 10.1038/nrc3341
- Hurskainen M, Ruggiero F, Hagg P, Pihlajaniemi T, and Huhtala P. 2010. Recombinant human collagen XV regulates cell adhesion and migration. *J Biol Chem* 285:5258-5265. 10.1074/jbc.M109.033787
- Jones MR, Kamara D, Karlan BY, Pharoah PDP, and Gayther SA. 2017. Genetic epidemiology of ovarian cancer and prospects for polygenic risk prediction. *Gynecol Oncol* 147:705-713. 10.1016/j.ygyno.2017.10.001
- Kar SP, Berchuck A, Gayther SA, Goode EL, Moysich KB, Pearce CL, Ramus SJ, Schildkraut JM, Sellers TA, and Pharoah PDP. 2018. Common Genetic Variation and Susceptibility to Ovarian Cancer: Current Insights and Future Directions. *Cancer Epidemiol Biomarkers Prev* 27:395-404. 10.1158/1055-9965.EPI-17-0315
- Kimura K, Nakayama M, Naito I, Komiyama T, Ichimura K, Asano H, Tsukuda K, Ohtsuka A, Oohashi T, Miyoshi S, and Ninomiya Y. 2016. Human collagen XV is a prominent histopathological component of sinusoidal capillarization in hepatocellular carcinogenesis. *Int J Clin Oncol* 21:302-309. 10.1007/s10147-015-0888-2
- Lawrenson K, Song F, Hazelett DJ, Kar SP, Tyrer J, Phelan CM, Corona RI, Rodriguez-Malave NI, Seo JH, Adler E, Coetzee SG, Segato F, Fonseca MAS, Amos CI, Carney ME, Chenevix-Trench G, Choi J, Doherty JA, Jia W, Jin GJ, Kim BG, Le ND, Lee J, Li L, Lim BK, Adenan NA, Mizuno M, Park B, Pearce CL, Shan K, Shi Y, Shu XO, Sieh W, Australian Ovarian Cancer Study G, Thompson PJ, Wilkens LR, Wei Q, Woo YL, Yan L, Karlan BY, Freedman ML, Noushmehr H, Goode EL, Berchuck A, Sellers TA, Teo SH, Zheng W, Matsuo K, Park S, Chen K, Pharoah PDP, Gayther SA, and Goodman MT. 2019. Genome-wide association studies identify susceptibility loci for epithelial ovarian cancer in east Asian women. *Gynecol Oncol* 153:343-355. 10.1016/j.ygyno.2019.02.023
- Lu Y, Kweon SS, Tanikawa C, Jia WH, Xiang YB, Cai Q, Zeng C, Schmit SL, Shin A, Matsuo K, Jee SH, Kim DH, Kim J, Wen W, Shi J, Guo X, Li B, Wang N, Zhang B, Li X, Shin MH, Li HL, Ren Z, Oh JH, Oze I, Ahn YO, Jung KJ, Conti DV, Schumacher FR, Rennert G, Jenkins MA, Campbell PT, Hoffmeister M, Casey G, Gruber SB, Gao J, Gao YT, Pan ZZ, Kamatani Y, Zeng YX, Shu XO, Long J, Matsuda K, and Zheng W. 2019. Large-Scale Genome-Wide Association Study of East Asians Identifies Loci Associated With Risk for Colorectal Cancer. *Gastroenterology* 156:1455-1466. 10.1053/j.gastro.2018.11.066
- Manichaikul A, Peres LC, Wang XQ, Barnard ME, Chyn D, Sheng X, Du Z, Tyrer J, Dennis J, Schwartz AG, Cote ML, Peters E, Moorman PG, Bondy M, Barnholtz-Sloan JS, Terry P, Alberg AJ, Bandera EV, Funkhouser E, Wu AH, Pearce CL, Pike M, Setiawan VW, Haiman CA, African American Breast Cancer C, African Ancestry Prostate Cancer C, Palmer JR, LeMarchand L, Wilkens LR, Berchuck A, Doherty JA, Modugno F, Ness R, Moysich K, Karlan BY, Whittemore AS, McGuire V, Sieh W, Lawrenson K, Gayther S, Sellers TA, Pharoah P, Schildkraut JM, African American Cancer Epidemiology S, and the Ovarian Cancer Association C. 2020. Identification of novel epithelial ovarian cancer loci in women of African ancestry. *Int J Cancer* 146:2987-2998. 10.1002/ijc.32653
- Martinez-Nieto G, Heljasvaara R, Heikkinen A, Kaski HK, Devarajan R, Rinne O, Henriksson C, Thomson E, von Hertzen C, Miinalainen I, Ruotsalainen H, Pihlajaniemi T, and Karppinen SM. 2021. Deletion of Col15a1 Modulates the Tumour Extracellular Matrix and Leads to Increased Tumour Growth in the MMTV-PyMT Mouse Mammary Carcinoma Model. *Int J Mol Sci* 22. 10.3390/ijms22189978
- Rebbeck TR. 2017. Prostate Cancer Genetics: Variation by Race, Ethnicity, and Geography. *Semin Radiat Oncol* 27:3-10. 10.1016/j.semradonc.2016.08.002
- Stewart C, Ralyea C, and Lockwood S. 2019. Ovarian Cancer: An Integrated Review. *Semin Oncol Nurs* 35:151-156. 10.1016/j.soncn.2019.02.001
- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, and Bray F. 2021. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 71:209-249. 10.3322/caac.21660
- Ward LD, and Kellis M. 2012. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic acids research* 40:D930-D934.
- Yao T, Hu W, Chen J, Shen L, Yu Y, Tang Z, Zang G, Zhang Y, and Chen X. 2022. Collagen XV mediated the

390 epithelial-mesenchymal transition to inhibit hepatocellular carcinoma metastasis. *J Gastrointest Oncol*
 391 13:2472-2484. 10.21037/jgo-22-299

392
393

Table 1(on next page)

Association results of SNPs at 9q22.33 in validation stage

CI, confidence interval; OR, odds ratio; RAF, risk allele frequency; SNP, singlenucleotide polymorphism. ^a Position is given with respect to genome build 37. ^b Risk allele/other allele.

Table 1:
Association results of SNPs at 9q22.33 in validation stage

SNP	Position ^a	Allele ^b	RAF	Genotypes	Case/Control	OR (95%CI)	<i>P</i>
rs7027650	101741969	T/A	0.57	AA	182/422	1.00(reference)	
				TA	450/569	1.82 (1.47-2.26)	3.44E-08
				TT	419/522	1.86 (1.50-2.31)	1.69E-08
				Addictive		1.39(1.24-1.55)	1.91E-07
				Dominant		1.80(1.62-1.98)	5.90E-10
				Recessive		1.26(1.10-1.42)	4.88E-03
rs1889268	101767961	T/C	0.32	CC	468/755	1.00 (reference)	
				TC	488/656	1.20 (1.02-1.41)	3.16E-02
				TT	115/153	1.21 (0.93-1.58)	1.61E-01
				Addictive		1.13(1.02-1.25)	3.71E-02
				Dominant		1.11(0.85-1.37)	4.23E-01
				Recessive		1.20(1.04-1.35)	2.32E-02

CI, confidence interval; OR, odds ratio; RAF, risk allele frequency; SNP, singlenucleotide polymorphism.

^a Position is given with respect to genome build 37.

^b Risk allele/other allele.

Figure 1

The association between three variants and EOC risk stratified by histological subtype.

ORs (95%CI) and *P* values for associations of three variants with each of the histological subtypes of EOC risk which were estimated using logistic regression adjusted for age.

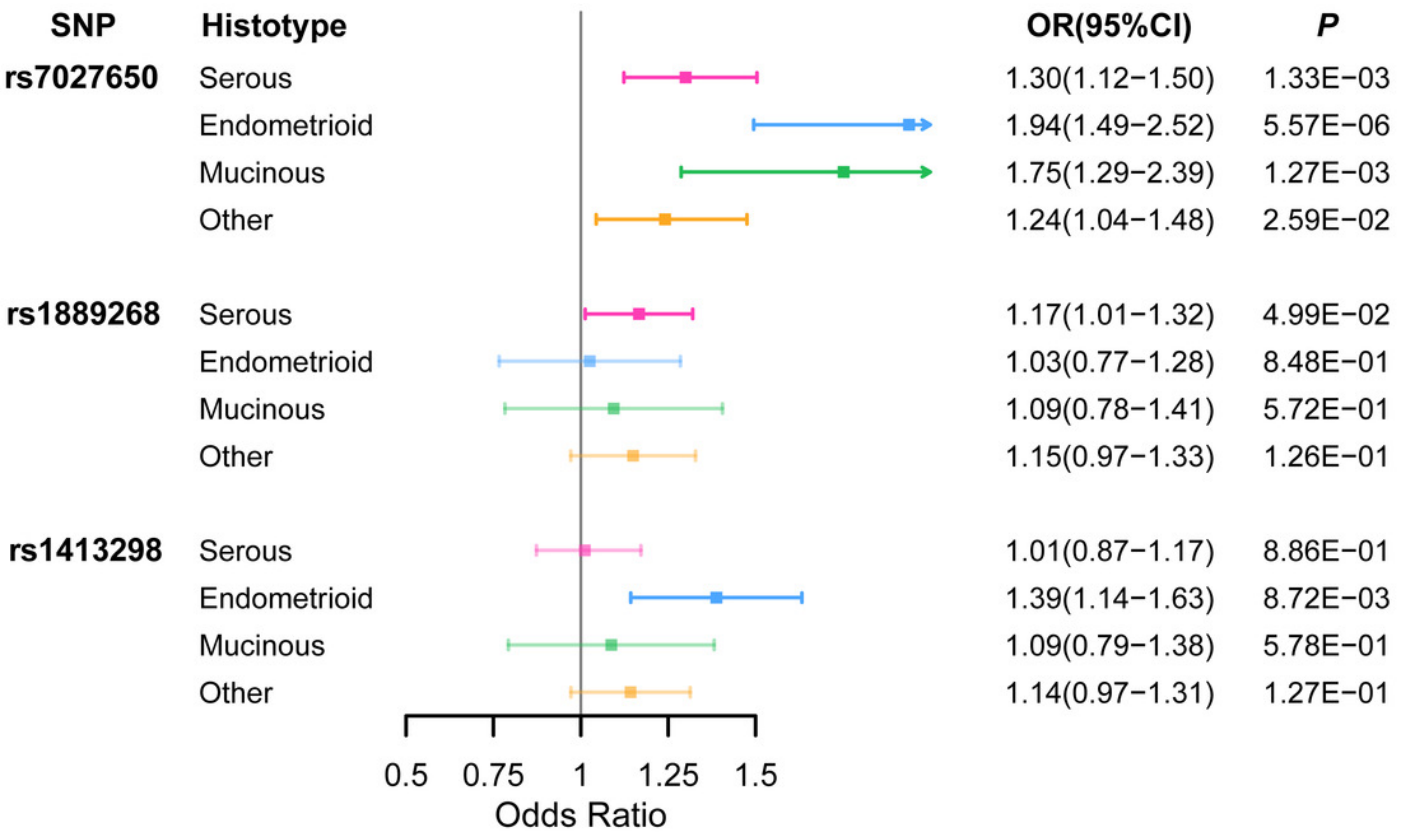


Figure 2

Epigenetic annotations and eQTL analysis for three candidate variants.

(A) Functional annotations with epigenomics data of normal ovary tissue and ovarian cancer cell lines. The lines represent H3K27ac and H3K4me1 histone modification ChIP-seq profiles, DNase-Seq for primary ovary tissue, and ATAC-seq profiles for ovarian cancer cell lines (OVCA432, DOV13, and SKOV3) from top to bottom. The region of rs7027650 overlaps with the peak of the DNase I hypersensitivity site measured by DNase-Seq. (B) Results of eQTL analysis in 102 ovarian cancer tumor tissues showed the significant association between rs7027650 and mRNA expression of *COL15A1*.

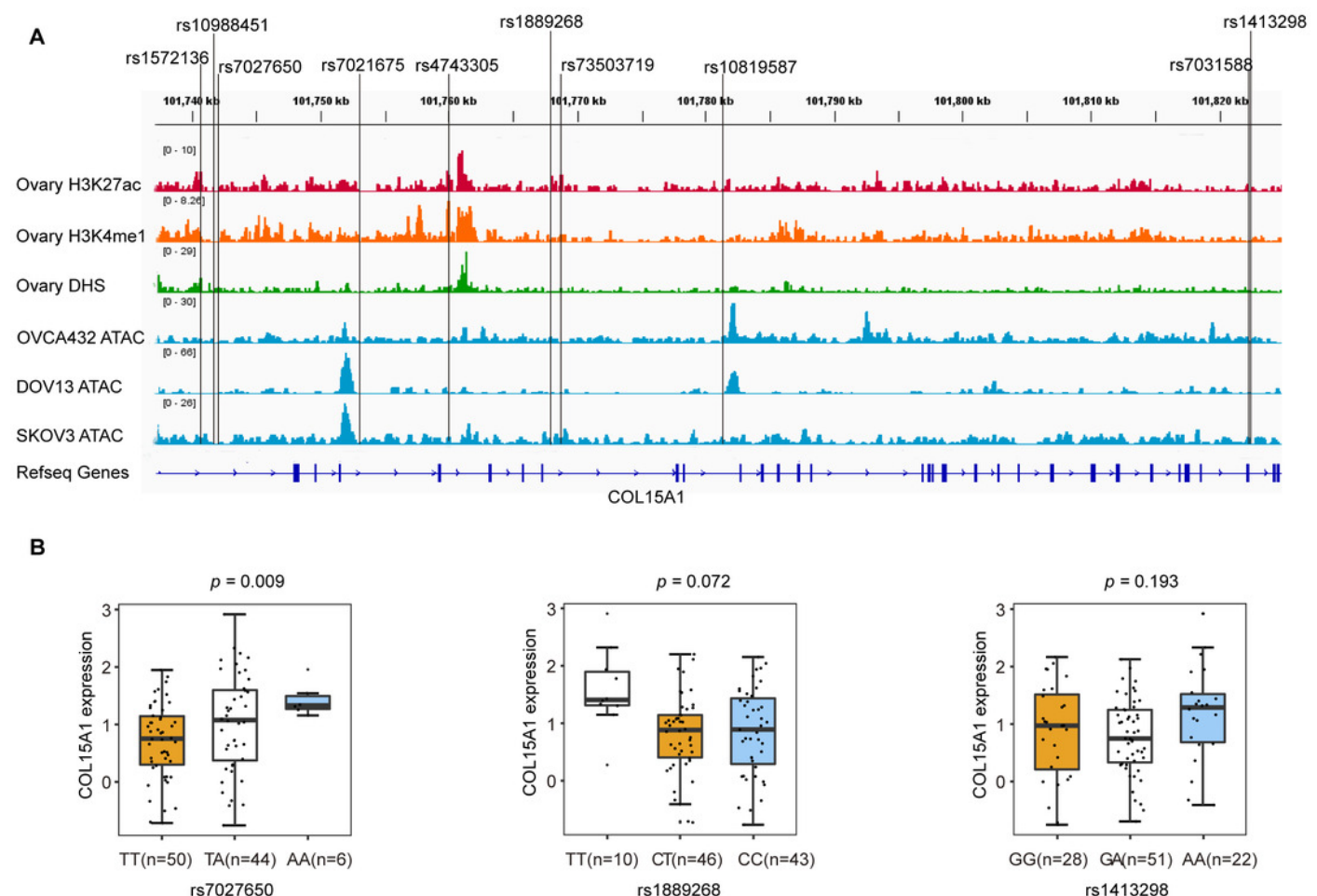


Figure 3

Functional validation of rs7027650.

(A) Luciferase reporter assay using vector containing rs7027650 in SKOV3 and 293T cells. The Luciferase reporter assay showed the different activity of effect allele T fragment and non-effect allele A fragment in SKOV3 and 293T cells. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. (B) The EMSA assay showed differences in the allele-specific protein binding capacity of rs7027650.

