Multivariate Analysis of Potential Pleiotropic Genes For Breast, Ovarian And Cervical Cancers Using Gene-Based Association Analysis

XiaoCan Jia

Zhengzhou University

Nian Shi

Zhengzhou University First Affiliated Hospital

Yu Feng

Zhengzhou University

Huili Zhu

Zhengzhou University

Yuping Wang

Zhengzhou University

Jie Lu

Zhengzhou College of Animal Husbandry Engineering: Henan University of Animal Husbandry and Economy

YongLi Yang

Zhengzhou University

Xuezhong shi (✉ xzshi@zzu.edu.cn)

Zhengzhou University <https://orcid.org/0000-0002-1694-1011>

Research Article

Keywords: Breast cancer, Ovarian cancer, Cervical cancer, metaCCA, GWAS

Posted Date: July 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-695526/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Although genome-wide association studies (GWAS) have a dramatic impact on susceptibility locus discovery in gynecological malignancies, the single nucleotide polymorphisms (SNPs) identified by this prevailing univariate approach only explain a small percentage of heredity. The extensive previous studies have repeatedly shown breast, ovarian and cervical cancers have common genetic mechanisms and the overlapping pathophysiological pathways. Novel multivariate analytical methods are necessary to identify shared pleiotropic genes. In this study, a total of 40,859 SNPs mapped in 11,597 gene regions were performed to identify potential common variants by using metaCCA and VEGAS2 analysis. Gene enrichment and protein-protein interaction (PPI) network analysis were used to explore potential biological pathways and connectivity. After metaCCA analysis, 4,203 SNPs ($P < 1.22 \times 10^{-6}$) and 1,886 pleiotropic gene ($P < 4.31 \times 10^{-6}$) were identified. By screening the results of gene-based P-values, the existence of 3 confirmed pleiotropic genes and 16 novel genes that achieved statistical significance in the metaCCA analysis and were also associated with at least one cancer in the VEGAS2 analysis were identified. The enrichment analysis showed the biological pathways of these genes were mainly enriched in 4 signaling pathways and 11 differentially expressed genes were found to encode interacting proteins in PPI network analysis. Altogether, we identified novel genetic variants of breast, ovarian and cervical cancers and provided evidence of biological functions which developed new insights for the diagnosis and treatment of these cancers.

Introduction

Breast, ovarian and cervical cancers are the most common and lethal gynecological malignancies. According to the global cancer statistics, about 3 million new cases are diagnosed and their incidence ranked first in the global burden of cancer (16.4% of the total cases) in 2018 [1]. Also, the epidemiological researches have showed that all three cancers are linked to hereditary and exhibits family clustering [2, 3]. Approximately, 180 genetic loci for breast cancer, 40 susceptibility loci for cervical cancer and 40 association loci for ovarian cancer have been identified in case-control genome-wide association studies (GWAS) and expression quantitative trait loci studies, which indicated that associations between genes and phenotypes are significant risk factors in the development of this three cancers.

Common molecular mechanisms and genetic polymorphism existed widely in the gynecology-related cancers. A molecular study revealed the involvement of autophagy in the development of cervical, endometrial and ovarian cancer [4]. Cross-cancer research showed that variation at the genetic level influenced the formation of multiple gynecology-related cancers. For example, analysis of gene expression has confirmed the association of variants at the 19p13 locus with estrogen receptor negative breast cancer and ovarian cancer [5]. Seven new cross-cancer loci were identified in a comprehensive GWAS meta-analysis of breast, ovarian and prostate cancers and further pathway analysis revealed apoptosis as the common susceptibility mechanism of these three cancers [6]. This suggested the existence of complex and low-penetrance polymorphic genes with shared effects among gynecology-related cancers.

GWAS is a commonly used approach for verifying individual single-nucleotide polymorphisms (SNPs) sequentially against a quantitative phenotype measure. However, a comparative study of GWAS and meta-analysis showed the univariate method were limited in detecting the heritability of intricate phenotypes because the widely existed internal correlation of genotype-genotype and phenotype-phenotype was ignored [7]. Recent research suggested simultaneously performing multiple related traits analysis could increase statistical power and identify complex genotype-phenotype correlations compared to the univariate GWAS analysis [8]. In addition, with the availability of GWAS summary statistical data, the researchers focus on performing multivariate statistical analysis to identify common genetic variant based on public data.

MetaCCA, a novel systematic statistical analysis method using canonical correlation analysis devised by Cichonska et al [8], was adopted to solve the limitation of GWAS. It aims to identify complex correlations between multiple genotypes and multiple phenotypes through combining several univariate summary statistics. It has been applied to identify novel shared genetic factors involved in the development of psychiatric disorders [9]. In this study, metaCCA was applied to detect potential overlapping pleiotropic genes shared by breast, ovarian and cervical carcinogenesis. In addition, gene enrichment and protein-protein interaction (PPI) network analysis were performed to explore potential biological function and connectivity among them.

Methods

GWAS Datasets

Primary SNP-level summary data for the three cancers were obtained from the Breast Cancer Association Consortium (BCAC), the Ovarian Cancer Association Consortium (OCAC) and the GWAS Catalog. The study included 89,677 breast cancer samples (46,785 cases and 42,892 controls), 66,450 ovarian cancer samples (25,509 cases and 40,941 controls), and 9,628 cervical cancer samples (2,866 cases and 6,762 controls) where all individuals were of European ancestry [10–12]. The summary statistics, including P-values, regression coefficients and standard errors calculated in the meta-analysis are available. Finally, overlapping SNPs of breast, ovarian and cervical cancers were selected for further analysis.

Data Preparation

A total of 294,879 common SNPs were acquired after combining summary statistics for all the three cancers. To select SNPs with small pairwise associations, linkage disequilibrium (LD) based SNP pruning method was used to select SNPs with imputation $r^2 \leq 0.2$ in European population. Following this primary selection of SNPs with lower LD, the process was repeated using a sliding window of five SNPs to confirm complete removal of pairs of SNPs with high LD. The whole process of SNP selection used the HapMap version 3 CEU genotypes panel (Utah residents with Northern and Western European ancestry from the CEPH collection) as a reference. After deleting the SNPs with large LD, 40,967 SNPs was remained. Then, according to the 1,000 Genome datasets, we accomplished the gene annotation for the

pruned SNPs using PLINK1.9 which based on the hg19 human genome build (<https://www.cog-genomics.org/static/bin/plink/glist-hg19>). Finally, 40,859 SNPs were mapped 11,597 gene regions were included in the metaCCA analysis. It is worth noting that the regression coefficient beta of each annotated SNP has to be normalized prior to metaCCA analysis [8].

Multivariate MetaCCA Analysis

The purpose of metaCCA analysis was to identify the potential pleiotropic genes for multiple diseases, which required a full covariance matrix (Σ), consisting of a cross-covariance matrix between all genotypic and phenotypic variables (ΣXY), a genotypic correlation structure between SNPs (ΣXX), and a phenotypic correlation structure between traits (ΣYY). ΣXY was imputed with the normalized regression coefficient β_{gp} (g and p were the number of genotypic and phenotypic variables, respectively). ΣXX was estimated using the reference SNP dataset of the HapMap 3 CEU. Each ΣYY entry corresponded to a Pearson correlation coefficient between vectors of β estimates from p phenotypic variables across g genetic variants [13]. It has been proved that the more the number of genotypic variables there are, the less the error of the estimate there is [14]. Therefore, 294,879 overlapping SNPs were used as the estimation of ΣYY , even if only some of the SNPs were included in further analysis. The canonical correlation coefficient r was used to describe the association between genotypes and phenotypes.

In this study, we performed association analysis to identify canonical correlations in the SNP and gene levels, respectively. In the SNP level, we selected susceptibility loci significantly associated with the three cancers using univariate SNP-multivariate phenotype association analysis. In addition, the GWAS summary statistics were examined to confirm random selection of the sample of 40,859 SNPs. The results revealed that the standardized mean of β approached zero and the corresponding median P -value approached 0.5 for all the three datasets, thus confirming that the sample of SNPs was from a random sample of the whole genome. The significant level α (two-sided) was corrected with Bonferroni method (adjusted $\alpha = 0.05/40,859$). Canonical correlation r of any SNP was significant when the P -value was smaller than 1.22×10^{-6} ($= 0.05/40,859$). In the gene level, we conducted multivariate SNP-multivariate phenotype association analysis to obtain a canonical correlation coefficient between a gene and the three cancers. Similarly, Canonical correlation r of any gene was significant when the P -value was smaller than 4.31×10^{-6} ($= 0.05/11,597$).

As a complement to the metaCCA analysis, we performed Versatile Gene-based Association Study-2 (VEGAS-2) analysis to refine the identified genes (<https://vegas2.qimrberghofer.edu.au/>). This method was a gene-based approach considering associations between a trait and all SNPs within a gene rather than each SNP individually [15]. All SNPs in each gene were incorporated into the VEGAS-2 analysis and then P -value was calculated for each trait in the gene level. Ultimately, we selected the pleiotropic genes associating with at least one phenotype using the adjusted threshold ($= 1 \times 10^{-6}$).

Functional Annotation

For identifying biological functions of the significant pleiotropic genes in our study, we explored the potential biological pathways using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Gene enrichment analysis was performed using KOBAS3.0 website developed by Center for Bioinformatics, Peking University (<http://kobas.cbi.pku.edu.cn/index.php>). In the enrichment analysis, $P < 0.01$ indicated statistically significant differences. To evaluate functional connectivity of genes, PPI network analysis were performed through Search Tool for the Retrieval Interacting Genes (STRING11.0) database (<https://string-db.org/>) [16]. We considered the total score above 0.400 that correspond to the combination of the following 7 different scores: text mining, experiments, databases, co-expression, gene neighborhood, gene fusion and gene co-occurrence.

Data availability

This data is publicly available from the BCAC, OCAC and GWAS Catalog.

BCAC: <http://bcac.ccge.medschl.cam.ac.uk/>

OCAC: <http://ocac.ccge.medschl.cam.ac.uk/>

GWAS Catalog: <https://www.ebi.ac.uk/gwas/>

Results

A two-step analysis strategy was performed to identify common SNPs and genes associated with breast, ovarian and cervical cancers. First, associations between multiple genotypes and multiple phenotypes were evaluated using metaCCA. The next step was to validate these pleiotropic genes for their specific associations with single cancer using the VEGAS-2 method. In addition, we performed gene enrichment and PPI network analysis to explore potential biological function and connectivity of these pleiotropic genes.

Pleiotropic SNPs and Genes Identified by MetaCCA Analysis

After the gene annotation and SNP pruning, we performed metaCCA analysis with a total of 40,859 SNPs located in 11,597 gene regions. In the univariate SNP-multivariate phenotype analysis, we obtained 4,203 significant SNPs ($P < 1.22 \times 10^{-6}$), and the canonical correlation coefficient r ranged from 0.018 to 0.083. The associations between SNP and phenotype are represented in the Manhattan plot (Fig. 1). If the $-\log_{10}$ (metaCCA) value of a SNP was more than 5.91 ($= -\log_{10}(1.22 \times 10^{-6})$), this SNP was regarded as a potential pleiotropic SNP for the 3 cancers. In the multivariate SNP-multivariate phenotype analysis, we identified 1,886 significant genes ($P < 4.31 \times 10^{-6}$), and the canonical correlation r of genes ranged from 0.022 to 0.680.

Refining the Pleiotropic Genes by VEGAS-2 Analysis

To evaluate the relationship between genes and cancers identified by metaCCA, we performed VEGAS-2 analysis to identify genes associated with one of the three traits. Using the VEGAS-2 methods, 98 genes were found to be related only with one trait (18 genes for cervical cancer, 61 genes for breast cancer and 20 genes for ovarian cancer).

By combining the results, 19 putative pleiotropic genes were confirmed to have a statistical significance in the metaCCA analysis and were associated with at least one trait in the VEGAS-2 analysis. According to the VEGAS-2 analysis, 3 pleiotropic genes (*EHMT2*, *LST1*, *LTA*) were found to be associated with cervical cancer, 14 pleiotropic genes were found to be associated with breast cancer and 3 pleiotropic genes (*BNC2*, *CRHR1*, *MLLT10*) were found to be associated with ovarian cancer. The pleiotropic gene, *MLLT10*, was associated with ovarian cancer and breast cancer based on VEGAS-2 analysis ($P < 1 \times 10^{-6}$). The findings of the metaCCA and VEGAS-2 analysis were summarized in Table 1.

Table 1
The 19 pleiotropic genes identified by the metaCCA and VEGAS-2 analysis

Number	Gene	MetaCCA p-value	VEGAS-2 p-value		
			breast cancer	ovarian cancer	cervical cancer
1	<i>BNC2</i>	3.71E-55	2.30E-04	1.00E-06	1.30E-03
2	<i>CASC21</i>	6.49E-40	1.00E-06	0.06	8.90E-04
3	<i>CASC8</i>	2.16E-52	1.00E-06	0.06	8.90E-04
4	<i>CCDC170</i>	2.58E-17	1.00E-06	5.10E-04	0.03
5	<i>CRHR1</i>	1.91E-07	0.03	1.00E-06	0.75
6	<i>EHMT2</i>	1.90E-12	0.03	0.05	1.00E-06
7	<i>ESR1</i>	6.11E-23	1.00E-06	1.10E-02	4.10E-03
8	<i>FAM227A</i>	1.17E-11	1.00E-06	0.17	0.08
9	<i>FGFR2</i>	2.92E-78	1.00E-06	4.80E-02	0.08
10	<i>ITPR1</i>	7.44E-65	1.00E-06	2.90E-02	9.30E-03
11	<i>LOC100506674</i>	9.53E-11	1.00E-06	9.10E-02	0.85
12	<i>LST1</i>	5.15E-30	0.02	0.66	1.00E-06
13	<i>LTA</i>	5.01E-23	0.02	0.66	1.00E-06
14	<i>MAP3K1</i>	8.59E-13	1.00E-06	0.11	0.38
15	<i>MLLT10</i>	4.55E-10	1.00E-06	6.94E-08	0.18
16	<i>NEK10</i>	3.90E-16	1.00E-06	0.01	0.20
17	<i>PEX14</i>	6.03E-08	1.00E-06	0.01	0.01
18	<i>SYT8</i>	1.17E-07	1.00E-06	0.20	0.02
19	<i>TNNT3/LSP1</i>	2.26E-08	1.00E-06	0.22	0.58

Particularly, among the 19 putative pleiotropic genes identified in this study, 3 (*CRHR1*, *LTA* and *MLLT10*) were previously reported to have an association with more than one trait. While the other 16 were novel pleiotropic genes, 10 genes (*BNC2*, *CASC8*, *CCDC170*, *ESR1*, *FGFR2*, *ITPR1*, *MAP3K1*, *NEK10*, *PEX14*, and *TNNT3/LSP1*) have been previously confirmed to be associated with one of the three cancers but was identified to be associated with all three cancers in our study. The remaining 6 genes (*CASC21*, *EHMT2*, *FAM227A*, *LOC100506674*, *LST1*, and *SYT8*) might represent novel pleiotropic candidate genes for breast, ovarian and cervical cancers. The detailed features of the 19 significant pleiotropic genes were shown in Table 2.

Table 2
The features of 19 significant pleiotropic genes

Chr	Gene	most significant SNP	Gene type	r-value	p-value
9	<i>BNC2</i>	rs10756819	Novel*[17]	0.07	3.71E-55
8	<i>CASC21</i>	rs10098985	Novel	0.06	6.49E-40
8	<i>CASC8</i>	rs1562430	Novel*[18]	0.06	2.16E-52
6	<i>CCDC170</i>	rs3734805	Novel*[19]	0.04	2.58E-17
17	<i>CRHR1</i>	rs7209556	Confirmed[20, 21]	0.03	1.91E-07
6	<i>EHMT2</i>	rs2227956	Novel	0.03	1.90E-12
6	<i>ESR1</i>	rs3734805	Novel*[22]	0.05	6.11E-23
22	<i>FAM227A</i>	rs2280790	Novel	0.03	1.17E-11
10	<i>FGFR2</i>	rs2981579	Novel*[23]	0.09	2.92E-78
3	<i>ITPR1</i>	rs9867580	Novel*[24]	0.07	7.44E-65
5	<i>LOC100506674</i>	rs11746980	Novel	0.03	9.53E-11
6	<i>LST1</i>	rs3134899	Novel	0.05	5.15E-30
6	<i>LTA</i>	rs3134899	Confirmed[25, 26]	0.05	5.01E-23
5	<i>MAP3K1</i>	rs832577	Novel*[23]	0.04	8.59E-13
10	<i>MLLT10</i>	rs2183271	Confirmed[27, 24]	0.03	4.55E-10
3	<i>NEK10</i>	rs2100006	Novel*[28]	0.04	3.90E-16
1	<i>PEX14</i>	rs616488	Novel*[24]	0.03	6.03E-08
11	<i>SYT8</i>	rs4980379	Novel	0.03	1.17E-07
11	<i>TNNT3/LSP1</i>	rs909116	Novel*[19]	0.03	2.26E-08
Note:					
Confirmed: Genes were associated with more than one cancer identified by previously studies.					
Novel*: Genes were associated with only one cancer identified by previously studies.					
Novel: New genes identified by this study.					

Gene Enrichment and PPI Analysis

When significant genes associated with the three cancers were analyzed with KOBAS3.0 online analysis, the genes were mainly enriched in the gonadotropin-releasing hormone (GnRH) signaling pathway, estrogen signaling pathway, proteoglycans in cancer and mitogen-activated protein kinase (MAPK)

signaling pathway. Details of the functional enrichment analysis results were shown in Table 3. PPI network analysis showed that 11 differentially expressed genes were found to encode interacting proteins (Fig. 2).

Table 3
Significant KEGG enrichment pathways of the pleiotropic genes (*p*-value < 0.01)

Biology pathways (KEGG Database)	KEGG ID	<i>p</i> -value	Genes
GnRH signaling pathway	hsa04912	8.1E-04	<i>ITPR1</i> ; <i>MAP3K1</i>
Estrogen signaling pathway	hsa04915	9.5E-04	<i>ESR1</i> ; <i>ITPR1</i>
Proteoglycans in cancer	hsa05205	3.9E-03	<i>ESR1</i> ; <i>ITPR1</i>
MAPK signaling pathway	hsa04010	5.9E-03	<i>FGFR2</i> ; <i>MAP3K1</i>

Discussion

In this study, 19 pleiotropic genes were identified to associate with breast, ovarian and cervical cancers, including 3 confirmed pleiotropic genes and 16 novel pleiotropic genes. Gene enrichment and PPI network analysis showed shared biological function and connectivity of the three cancers which developed new insights for the diagnosis and treatment of gynecology-related cancers.

Three confirmed pleiotropic genes (*CRHR1*, *LTA*, and *MLLT10*) have also been identified in this study, implying that these genes exerted a crucial influence on gynecology-related cancers. *CRHR1* increased the expression of Fas ligand through corticotropin-releasing hormone and positive expression of Fas ligand was associated with higher ovarian tumor stage [29]. However, the activation of *CRHR1* diminished breast cancer progression through participating in Urocortin-repressed transforming growth factor β 1 signaling [30]. *LTA* was a member of the tumor necrosis factor family, which encoded a cytokine produced by lymphocytes [2]. *LTA* locus variants was likely to increase the risk of gynecology-related cancers including breast and cervical cancers [31]. In addition, *LTA* gene was significantly associated with increased risks of cervical and vulvar cancers according to gene-based analysis [25]. *MLLT10* has been reported to associate with breast and ovarian cancers [27]. Interestingly, the genes were also identified to associate with breast and ovarian cancers in VEGAS-2 analysis and with all three cancers in metaCCA analysis. Potential biological mechanisms of the pleiotropic gene may include inhibition of cell cycle arrest and apoptosis, impaired Deoxyribonucleic acid repair and myelopoiesis regulation [32].

The other 10 novel pleiotropic genes associated with one of the three cancers in previously studies, some of them were showed in the PPI network co-expressed with each other suggesting important roles of these genes in gynecology-related cancers. *BNC2* was highly expressed in endometrium and ovary, which indicated an important role of *BNC2* in the development of other gynecology-related cancers. *ESR1* played a crucial role in the etiology of breast cancer by regulating estrogen signal transduction. *ESR1* expression controlled a feed-forward that sustained activation of the CXCR7/CXCL11 chemokine axis to induce the metastatic behavior of ovarian cancer cells [33]. Besides, loss of estrogen receptor- α had a major role in

mediating cervical cancer invasion and progression [34]. *FGFR2* was a member of the fibroblast growth factor receptor family, which encoded the fibroblast growth factor receptor. *FGFR2* aberrations might affect multiple gynecology-related cancers, such as breast, ovarian, cervical and endometrial cancers [35]. *ITPR1* was identified as an autophagy gene and significantly associated with overall survival in breast cancer [36]. Pathway analysis also showed that *ITPR1* and *ESR1* were jointly involved in the estrogen signaling pathway that promoted breast cancer progression in cellular.

Compared with genes identified in previous GWAS studies of these three cancers, the remaining six genes (*CASC21*, *EHMT2*, *FAM227A*, *LOC100506674*, *LST1*, and *SYT8*) have never been reported. Research showed that *CASC8* and *CASC21* were a hotspot gene integrated by human papilloma virus and promoted the development of cervical cancer [37]. *LST1* was confirmed to correlate with relapse-free survival and distal metastasis-free survival in triple-negative breast cancer and general breast cancer [38]. Another important gene, *EHMT2* encoded a methyltransferase and higher *EHMT2* expression was found to be associated with an adverse prognosis with breast cancer patients and predicted a shorter overall survival in ovarian cancer patients. Therefore, *CASC21*, *EHMT2* and *LST1* might have important effect in the development and therapeutics of other gynecology-related cancers. Detailed biological mechanism of 6 genes remain unclear, and further experiments might need to be conducted to identify the new findings.

Compared with the standard single phenotype GWAS, this cost-effective analytical method provided advantages for the exploration of genetic polymorphisms. First, higher statistical power and larger sample size were obtained by combining the summary statistics of the three large GWAS. Second, present analysis of multiple related traits could contribute to richer findings and increasing the possibility of discovering associations that are more complicated. Inevitably, there were also some limitations in our study. First, the research was unable to determine the proportions of variability explained by the identified genes due to shortfall of detailed primary individual measures. Second, our findings were based on data analysis without experimental approach. Additional experiments might be necessary to confirm the novel genes of gynecology-related cancers in further studies.

Conclusion

In summary, we performed multivariate statistical analysis of breast, ovarian and cervical cancers using metaCCA and VEGAS-2 based on summary statistics, identified 3 confirmed pleiotropic genes in the previous study and 16 significant genes that may be the novel pleiotropic candidate genes associated with all three cancers, which developed new insights for the diagnosis and treatment of gynecology-related cancers.

Declarations

Acknowledgments

All authors thank the MogoEdit (<http://www.mogoedit.com/>) for language modification.

Funding

This study was supported by the National Natural Science Foundation of China (No. 82073670).

Author information

Affiliations

Department of Epidemiology and Biostatistics, College of Public Health, Zhengzhou University, Zhengzhou, P. R. China

XiaoCan Jia, Yu Feng, Huili Zhu, Yuping Wang, Jie Lu, YongLi Yang & XueZhong Shi

Department of Physical Diagnosis, the First Affiliated Hospital of Zhengzhou University, Zhengzhou, P. R. China

Nian Shi

Corresponding author

Correspondence to YongLi Yang & XueZhong Shi.

Ethics declarations

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492.
2. Ding H, Fan G-L, Yi Y-X, et al (2020) Prognostic Implications of Immune-Related Genes' (IRGs) Signature Models in Cervical Cancer and Endometrial Cancer. *Front Genet* 11. doi:10.3389/fgene.2020.00725.
3. Wu A, Zhang S, Liu J, et al (2020) Integrated Analysis of Prognostic and Immune Associated Integrin Family in Ovarian Cancer. *Front Genet* 11. doi:10.3389/fgene.2020.00705.
4. Orfanelli T, Jeong JM, Douglaveris G, et al. Involvement of autophagy in cervical, endometrial and ovarian cancer. *International journal of cancer*. 2014;135(3):519–28. doi:10.1002/ijc.28524.

5. Bolton KL, Tyrer J, Song H, et al. Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nat Genet.* 2010;42(10):880–4. doi:10.1038/ng.666.
6. Kar SP, Beesley J, Amin Al Olama A, et al. Genome-Wide Meta-Analyses of Breast, Ovarian, and Prostate Cancer Association Studies Identify Multiple New Susceptibility Loci Shared by at Least Two Cancer Types. *Cancer Discov.* 2016;6(9):1052–67. doi:10.1158/2159-8290.CD-15-1227.
7. Pei YF, Zhang L, Papasian CJ, Wang YP, Deng HW. On individual genome-wide association studies and their meta-analysis. *Hum Genet.* 2014;133(3):265–79. doi:10.1007/s00439-013-1366-4.
8. Cichonska A, Rousu J, Marttinen P, et al. metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. *Bioinformatics.* 2016;32(13):1981–9. doi:10.1093/bioinformatics/btw052.
9. Jia X, Yang Y, Chen Y, et al. Multivariate analysis of genome-wide data to identify potential pleiotropic genes for five major psychiatric disorders using MetaCCA. *J Affect Disord.* 2019;242:234–43. doi:10.1016/j.jad.2018.07.046.
10. Leo PJ, Madeleine MM, Wang S, et al (2017) Defining the genetic susceptibility to cervical neoplasia-A genome-wide association study. *Plos Genetics* 13 (8). doi:10.1371/journal.pgen.1006866.
11. Phelan CM, Kuchenbaecker KB, Tyrer JP, et al. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet.* 2017;49(5):680–91. doi:10.1038/ng.3826.
12. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet.* 2013;45(4):353–61. doi:10.1038/ng.2563.
13. Jia X, Yang Y, Chen Y, et al. Multivariate analysis of genome-wide data to identify potential pleiotropic genes for type 2 diabetes, obesity and coronary artery disease using MetaCCA. *Int J Cardiol.* 2019;283:144–50. doi:10.1016/j.ijcard.2018.10.102.
14. Jia X, Yang Y, Chen Y, et al. Multivariate analysis of genome-wide data to identify potential pleiotropic genes for five major psychiatric disorders using MetaCCA. *J Affect Disord.* 2019;242:234–43. doi:10.1016/j.jad.2018.07.046.
15. Wojcik GL, Kao WHL, Duggal P. (2015) Relative performance of gene- and pathway-level methods as secondary analyses for genome-wide association studies. *Bmc Genetics* 16. doi:10.1186/s12863-015-0191-2.
16. Chen EY, Tan CM, Kou Y, et al (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *Bmc Bioinformatics* 14. doi:10.1186/1471-2105-14-128.
17. Song H, Ramus SJ, Tyrer J, et al. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. *Nat Genet.* 2009;41(9):996–1000. doi:10.1038/ng.424.
18. Turnbull C, Ahmed S, Morrison J, et al. Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat Genet.* 2010;42(6):504–7. doi:10.1038/ng.586.
19. Lindstrom S, Thompson DJ, Paterson AD, et al. Genome-wide association study identifies multiple loci associated with both mammographic density and breast cancer risk. *Nature communications.* 2014;5:5303. doi:10.1038/ncomms6303.

20. Wu L, Shi W, Long J, Guo X. (2018) A transcriptome-wide association study of 229,000 women identifies new candidate susceptibility genes for breast cancer. 50 (7):968–978. doi:10.1038/s41588-018-0132-x.
21. Lu Y, Beeghly-Fadiel A, Wu L, et al (2018) A Transcriptome-Wide Association Study Among 97,898 Women to Identify Candidate Susceptibility Genes for Epithelial Ovarian Cancer Risk. 78 (18):5419–5430. doi:10.1158/0008-5472.can-18-0951.
22. Fletcher O, Johnson N, Orr N, et al. Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. J Natl Cancer Inst. 2011;103(5):425–35. doi:10.1093/jnci/djq563.
23. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007;447(7148):1087–93. doi:10.1038/nature05887.
24. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet. 2013;45(4):353–61. doi:10.1038/ng.2563. 361e351-352.
25. Bodelon C, Madeleine MM, Johnson LG, et al. Genetic variation in the TLR and NF-kappaB pathways and cervical and vulvar cancer risk: a population-based case-control study. International journal of cancer. 2014;134(2):437–44. doi:10.1002/ijc.28364.
26. Quan L, Gong Z, Yao S, et al. Cytokine and cytokine receptor genes of the adaptive immune response are differentially associated with breast cancer risk in American women of African and European ancestry. International journal of cancer. 2014;134(6):1408–21. doi:10.1002/ijc.28458.
27. Pharoah PD, Tsai YY, Ramus SJ, et al. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. Nat Genet. 2013;45(4):362–70. doi:10.1038/ng.2564. 370e361-362.
28. Ahmed S, Thomas G, Ghoussaini M, et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet. 2009;41(5):585–90. doi:10.1038/ng.354.
29. Minas V, Rolaki A, Kalantaridou SN, et al. Intratumoral CRH modulates immuno-escape of ovarian cancer cells through FasL regulation. British journal of cancer. 2007;97(5):637–45. doi:10.1038/sj.bjc.6603918.
30. Jin L, Zhu C, Wang X, et al. Urocortin attenuates TGFbeta1-induced Snail1 and slug expressions: inhibitory role of Smad7 in Smad2/3 signaling in breast cancer cells. Journal of cellular biochemistry. 2015;116(11):2494–503. doi:10.1002/jcb.25194.
31. Huang Y, Yu X, Wang L, et al. Four genetic polymorphisms of lymphotoxin-alpha gene and cancer risk: a systematic review and meta-analysis. PloS one. 2013;8(12):e82519. doi:10.1371/journal.pone.0082519.
32. Sakoda LC, Jorgenson E, Witte JS. Turning of COGS moves forward findings for hormonally mediated cancers. Nat Genet. 2013;45(4):345–8. doi:10.1038/ng.2587.
33. Benhadjeba S, Edjekouane L, Sauve K, et al (2018) Feedback control of the CXCR7/CXCL11 chemokine axis by estrogen receptor alpha in ovarian cancer. 12 (10):1689–1705. doi:10.1002/1878-0261.12362.

34. Zhai Y, Bommer GT, Feng Y, et al. Loss of estrogen receptor 1 enhances cervical cancer invasion. *Am J Pathol.* 2010;177(2):884–95. doi:10.2353/ajpath.2010.091166.
35. Helsten T, Elkin S, Arthur E, et al. The FGFR Landscape in Cancer: Analysis of 4,853 Tumors by Next-Generation Sequencing. *Clinical cancer research: an official journal of the American Association for Cancer Research.* 2016;22(1):259–67. doi:10.1158/1078-0432.ccr-14-3212.
36. Gu Y, Li P, Peng F, et al. Autophagy-related prognostic signature for breast cancer. *Molecular carcinogenesis.* 2016;55(3):292–9. doi:10.1002/mc.22278.
37. Li W, Qi Y, Cui X. (2018) Characteristic of HPV Integration in the Genome and Transcriptome of Cervical Cancer Tissues. 2018:6242173. doi:10.1155/2018/6242173.
38. Hsu HM, Chu CM, Chang YJ, et al. Six novel immunoglobulin genes as biomarkers for better prognosis in triple-negative breast cancer by gene co-expression network analysis. *Scientific reports.* 2019;9(1):4484. doi:10.1038/s41598-019-40826-w.

Figures

Figure 1

Manhattan plot of $-\log_{10}(\text{metaCCA})$ values for univariate SNP-three cancers analysis. The red line labels the $-\log_{10}(\text{metaCCA})$ value of 5.91. If the $-\log_{10}(\text{metaCCA})$ value of a certain SNP was greater than 5.91, this SNP was identified as a pleiotropic SNP for breast, ovarian and cervical cancers.

Figure 2

The PPI network generated from the 11 susceptibility genes with total score above 0.400. The light blue and purple lines inside the nodes represent known interaction from curated databases and experimentally determined, respectively. The green, red and blue lines inside the nodes represent three predicted interactions: gene neighborhood, gene fusions and gene co-occurrence. The black, gray and yellow lines represent other interaction: text mining, co-expression and protein homology.