An Immune-Related Gene-Based Signature as Prognostic Tool in Ovarian Serous Cystadenocarcinoma

Fengjuan Shi, Tongning Deng, Jian Mo, Huiru Wang, Jianxun Lu

Background: Ovarian serous cystadenocarcinoma (OSCC) is a life-threatening malignancy with poor prognosis. Therefore, the identification of immune-related genes associated with OSCC prognosis may reveal new targets of immunotherapy for OSCC.

Patients and Methods: The gene expression profiles of overlapped genes were extracted by weighted gene co-expression network analysis (WGCNA) to identify immune-related modules. Significant genes were identified by univariate Cox regression analysis of model genes. Model characteristic genes were obtained by least absolute shrinkage and selection operator (LASSO) analysis and used to calculate a “signature index”. The model’s ability to predict prognosis in OSCC patients was assessed using time-dependent receiver operator characteristic curves. Differences in the biological processes and Kyoto Encyclopedia of Genes and Genomes pathways between groups with high or low signature index were assessed using gene set enrichment analysis (GSEA). The types of immune cells and their abundance in the two index groups were explored by single-sample GSEA.

Results: The expression profiles of 3517 overlapped genes were extracted by WGCNA, and nine modules related to the immune system of OSCC were obtained. The expression profiles of 114 hub genes were then subjected to LASSO analysis. Among them, 10 immune-related genes were significant, of which six were identified as model characteristic genes and were used to calculate the signature index. Moreover, 24 types of immune cells were identified in the tumor microenvironment, and their abundance was explored in high- and low-signature index groups of two datasets.

Conclusion: ARHGEF18, PLEKHA7, MTOR, VPS45, BRCA1, and HINT2 were identified as characteristic genes and used to develop a new immune-related gene-based signature as a promising prognostic biomarker for OSCC.

Keywords: ovarian serous cystadenocarcinoma, prognostic biomarker, signature index, immune-related genes

Introduction

Ovarian cancer, a malignancy that arises in ovarian tissue, is classified by the World Health Organization as serous carcinoma, mucinous carcinoma, endometrioid carcinoma, or clear-cell carcinoma, or as Brenner, mixed, or undifferentiated types. Ovarian serous cystadenocarcinoma (OSCC) is an aggressive type of ovarian cancer with high recurrence rate, metastatic ability, mortality, and resistance to conventional therapy. Although treatment methods for ovarian cancer are gradually improving, the prognosis of advanced serous ovarian neoplasms is still poor, indicating the need for new biomarkers.
Tumor-associated macrophages (TAMS) are the major lineage cells in the ovarian tumor microenvironment (TME) and have been shown to play an important role in immunosuppression through tumor cell invasion, angiogenesis, metastasis, and early relapse. Immune cells, mesenchymal cells, endothelial cells, inflammatory mediators, and extracellular matrix (ECM) molecules have also been found in the TME of serous ovarian neoplasms, with immune cells playing a defensive role through chemotaxis. Based on these findings and given that immune escape is a major cause of tumorigenesis, we hypothesized that classifying immune cells and estimating their abundance in the ovarian TME would help clarify their involvement in cancer. Therefore, in this study, we used weighted gene co-expression network analysis (WGCNA), univariate Cox regression analysis, and least absolute shrinkage and selection operator (LASSO) analysis to identify immune-related modules and calculate a “signature index”. Time-dependent receiver operator characteristics curve (tROC) analysis was also performed, suggesting that the signature index is a potential immune-related prognostic biomarker for serous ovarian neoplasms.

Materials and Methods

Data Processing

The gene expression profiles of the GSE53963 dataset were extracted from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/). GSE53963 is based on GPL6480 and contains 174 serous ovarian cancer samples, which were used as the training set to create an immune-related gene-based signature. The extracted gene expression profiles were normalized using the normalizeBetweenArrays function in the limma R package. RNA sequencing (displayed as read count) and clinical information from patients with ovarian serous cancer were acquired from The Cancer Genome Atlas (TCGA; https://www.cancer.gov/) to validate the prognostic ability of the immune-related gene-based signature. RNA-sequencing data were normalized using the VOOM function in the limma R package. If a gene was measured using multiple probes, its expression was considered as the mean expression across all the probes.

WGCNA

Immune-related genes (IRGs) were obtained from the InnateDB database (http://www.innatedb.com) and their expression profiles in the GSE53963 dataset were extracted by WGCNA to identify phenotype-related gene modules. A module eigengene was defined as the first principal component of the module’s expression matrix; module membership (MM), as the association between a gene and its module; and gene significance (GS), as the correlation of a gene with a phenotype. Genes significantly associated with MM ($P < 0.05$) were defined as hub genes. Clinicopathological features including death (event), tumor grade and overall survival (OS) were considered as phenotypes in WGCNA.

Univariate Cox Regression and LASSO Analysis

Univariate Cox regression was performed to assess the potential relation of hub IRGs with prognosis. $P$ values $< 0.05$ were considered statistically significant. The LASSO method was then applied using the glmnet package (https://CRAN.R-project.org/package=glmnet) in R for feature selection to identify IRGs with non-zero coefficients. The IRG-based signature to predict prognosis was generated based on the following formula:

$$\text{Prognostic Index (PI) = Expr_{Gene1} \times Coef_1 + Expr_{Gene2} \times Coef_2 + Expr_{Gene3} \times Coef_3 + \ldots}$$

where Expr represents the expression value of the gene and Coef the coefficient of the corresponding gene provided by LASSO.

The survivalROC package was used to draw the tROC curves and assess the prognostic value of the IRG-based signature. OSCC patients were divided into high- and low-signature index groups according to the corresponding cut-off, and their Kaplan–Meier survival curves were compared using the log-rank method. Differences with $P < 0.05$ were considered statistically significant. Univariate/multivariate Cox proportional hazards regression was also applied to compare the relative prognostic value of the signature index with that of routine clinicopathological features.

Immunohistochemical Analysis

The expression of signature IRGs in normal ovarian and ovarian cancer tissues was explored using the Human Protein Atlas database (http://www.proteinatlas.org), which contains information on a wide variety of human protein-coding genes. The expression of a given gene was investigated in control and cancer tissues using the
same antibody, and expression level was semi-quantitated as “high” (at least 75% of cells staining positive), “medium” (25–75%) or “low” (no more than 25%).

Gene Set Enrichment Analysis (GSEA)
GSEA was performed using normalized mRNA expression profiles of a JAVA-based software (http://software.broad institute.org/gsea/index.jsp) to explore the differences in biological processes (BPs) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways between the high- and low-signature index groups. The gene sets c5.bp.v6.2.symbols.gmt and c2.cp.kegg.v6.2.symbols.gmt from the Molecular Signatures Database (MSigDB) were used as a reference. Nominal P values < 0.05 were considered statistically significant.

Tumor Microenvironment Cell (TMC) Abundance Between High- and Low-Signature Index Groups
A total of 24 types of TMCs were identified for each sample using single-sample GSEA (ssGSEA) based on a previous study. The marker genes of all TMCs were also obtained from previous studies. The ssGSEA analysis was performed using the gsva function of the GSVA package in R and the abundance of TMCs was compared between the high- and low-signature index groups in the GSE53963 and TCGA datasets.

Statistical Analysis
Student’s t-test was used to compare continuous variables. Univariate/multivariate Cox proportional hazards analysis and Kaplan–Meier survival analysis were used to compare survival between low- and high-signature index groups. All analyses were conducted in R version 3.6 (https://www.r-project.org/). All tests were two-sided, and values with P < 0.05 were considered statistically significant.

Results
Identification of Survival-Related Gene Modules
To identify the survival-related gene modules of OSCC, the expression profiles of 3517 overlapped genes were extracted by WGCNA (Figure 1A). The power value was set to 3, which was the lowest value with an independence degree of up to 0.95 (Figure 1B). Nine modules were identified (Figure 1C), of which the red module was positively associated with the stage of OSCC, and the yellow with the overall survival (OS) of OSCC patients (Figure 1D and E).

Construction of a Prognostic Signature Based on IRGs
A total of 114 hub genes from the red and yellow modules were subjected to univariate Cox regression analysis, identifying 10 IRGs as significant (P < 0.05) (Table 1). Among them, six genes (ARHGEF18, PLEKHA7, MTOR, VPS45, BRCA1, and HINT2) were identified as model characteristic genes (Figure 2A) and their characteristic coefficients were calculated to obtain the signature index. Analysis of tROCs revealed that the area under the curve (AUC) tended to 1 for two- and five-year survival in the GSE53963 training set, suggesting that the signature index is an effective prognostic marker for OSCC (Figure 2B). Patients in the high-signature index group had a significantly lower survival rate than patients in the low-signature index group (Figure 2C), and similar survival results were obtained in the TCGA validation dataset (Figure 2D). Furthermore, analysis of the TCGA dataset based on another optimized cut-off to distinguish between high- and low-signature index groups showed that the signature was an independent predictor of survival after adjustment for routine clinicopathological features (Figure 2E).

Model Characteristic Genes Highly Expressed at the Protein Level in OSCC
The six identified model characteristic genes were present in The Human Protein Atlas, where they were expressed at higher levels in OSCC samples than in normal ovary samples (Figure 3), strongly supporting our results. The immunohistochemical image credit: Human Protein Atlas. The image available from Human Protein Atlas, the URL of each gene was shown in Supplementary Materials.

Molecular Function in the High- and Low-Signature Index Groups
GSEA indicated that several BPs, including “Golgi organization” and “autophagosome organization” (Figure 4A), as well as several KEGG pathways, including “inositol phosphate metabolism” and “oxidative phosphorylation”, were significantly enriched in the GSE53963 training set (Figure 4B).
In the TCGA validation set, enriched BPs included “activation of GTPase activity” and “activation of Jun kinase activity” (Figure 4C), while enriched KEGG pathways included “adherens junctions” and the “phosphatidylinositol signaling system” (Figure 4D).

**Immune Infiltration Score in the High- and Low-Signature Index Groups**

A total of 24 types of immune cells were analyzed: memory and naive B cells; activated and resting dendritic cells; endothelial cells; eosinophils; fibroblasts; M0-, M1-, and M2-
Table 1 Immune-Related Genes in the GSE53963 Dataset

| Gene    | HR (95% CI)          | Lasso Coefficient |
|---------|----------------------|-------------------|
| BRCA1   | 2.698 (1.141–6.377)  | 0.450422377       |
| BCL2    | 0.339 (0.132–0.868)  | 0.691152409       |
| ARHGEF18| 5.047 (1.853–13.745) | 0.364578189       |
| RASA1   | 3.848 (1.236–11.98)  | 0.140692263       |
| PLEKHA7 | 4.392 (1.447–13.33)  | 3.848 (1.236–11.98)|
| FAM3D   | 0.507 (0.261–0.984)  | 10.012592712      |
| TOP1    | 4.255 (1.12–16.167)  | 0.063 (0.015–0.265)|
| MTO1    | 4.287 (1.163–15.803) | 0.339 (0.132–0.868)|
| VPS45   | 13.4 (2.421–74.161)  | 0.024              |
| HINT2   | 0.063 (0.015–0.265)  | 0.024              |

Abbreviations: CI, confidence interval; HR, hazard ratio; Lasso, least absolute shrinkage and selection operator.

Discussion

OSCC is a deadly gynecological cancer associated with high rates of recurrence and mortality. Immunotherapy has recently been applied for personalized treatment of serous ovarian neoplasms based on the association of immune cells in the TME with OSCC. The combination of poly (ADP-ribose) polymerase (PARP) and immune checkpoint inhibition have also been proposed as a therapeutic approach against ovarian cancer, while immune checkpoint blockade therapy, PARP inhibitors, neoantigen vaccines, and personalized T-cell therapy have been applied in the clinic. Several studies have also reported gene expression characteristics for OSCC, though the clinical relevance of disease models based on such expression profiling remains unclear.

Here, in order to improve the prognosis of patients with OSCC, we developed a novel gene-based signature index to identify new immune-related prognostic biomarkers. IRGs were extracted from the InnateDB database, which includes a comprehensive list of genes related to innate immunity, and their expression profiles in GSE53963 were extracted by WGCNA to identify gene modules associated with OSCC. Subsequent LASSO analysis identified ARHGEF18, PLEKHA7, MTO1, VPS45, BRCA1, and HINT2 as characteristic genes, and their coefficients were used to calculate the signature index.

PLEKHA7 can alter E-cadherin-containing cell–cell contacts, while high PLEKHA7 expression has been associated with good prognosis in epithelial ovarian cancer. MTO1 and VPS45 have also been related to prognosis in OSCC, while ARHGEF18 has been identified as an indicator of genetic susceptibility to non-idiopathic pulmonary arterial hypertension. BRCA1 mutations have been found in invasive ovarian cancer patients, while we are unaware of studies reporting associations between HINT2 and OSCC. By modulating HINT2 mRNA translation, dynamic N6-methyladenosine (m6A) RNA modification can suppress ocular melanoma. Thus, the signature developed in the present study using the six characteristic genes may serve as a new prognostic biomarker for OSCC. Using the median value of this signature as a cut-off, we were able to stratify patients into groups showing different survival.

Our GSEA analysis revealed that oxidative phosphorylation, an active metabolic pathway in serous ovarian neoplasms, is significantly enriched in the GSE53963 training set. Moreover, we identified adherens junctions, which are involved in the Mical-L2 regulated epithelial–mesenchymal transition (EMT) in ovarian cancer cells, as an enriched KEGG pathway in the TCGA validation set. The loss of E-cadherin during the EMT disrupts adherens type macrophages; activated and resting mast cells; monocytes; neutrophils; activated and resting natural killer (NK) cells; plasma cells; memory activated, memory resting, and naive CD4 T cells; CD8 T cells; follicular helper T cells; gamma delta T cells; and transferred regulatory T cells.

Abbreviations: CI, confidence interval; HR, hazard ratio; Lasso, least absolute shrinkage and selection operator.
Figure 2 Construction of the immune-related prognostic signature for ovarian serous cystadenocarcinoma. (A) Determination of the six model characteristic genes by least absolute shrinkage and selection operator analysis. (B) Time-dependent receiver operating characteristic (tROC) curves to evaluate the ability of the signature index to predict 2-year and 5-year survival in the GSE53963 dataset. (C and D) Survival probability in the GSE53963 and The Cancer Genome Atlas datasets. (E) Forest plot of the signature index.

Abbreviation: AUC, area under the tROC curve.
junctions, thus preventing anoikis and promoting metasta-
sis and chemoresistance.\textsuperscript{32}

Furthermore, 24 immune cell types were identified in the TME of OSCC by ssGSEA analysis. Among them, dendritic and NK cells, which are an integral part of the innate immune response to prevent the development of OSCC,\textsuperscript{33} had higher expression in the low-signature index group than in the high-signature index group. Moreover, the survival rate in the high-signature index group was lower than that in the low-signature index group, consistent with the results of a survival analysis of this study. Taken together, our results suggest that the signature index may serve as an immune marker for OSCC, while also providing a foundation for further mechanistic and therapeutic studies.

Our study was limited to the size of the samples in the public databases from which we obtained our data, and we performed analyses based only on bioinformatics. Our
Figure 4 Gene set enrichment analysis. (A and C) Biological processes and (B and D) Kyoto Encyclopedia of Genes and Genomes pathways enriched in the (A and B) GSE53963 and (C and D) The Cancer Genome Atlas datasets.

Figure 5 Abundance of immune cells in ovarian serous cystadenocarcinoma patients in the (A) GSE53963 and (B) The Cancer Genome Atlas datasets. Relative abundance was determined indirectly from the expression level of marker genes corresponding to each type of immune cell.
results should be verified and extended in biological experiments, particularly the hypothesis that the genes in our novel signature are involved in immune responses to ovarian cancer.

**Conclusion**

In summary, we developed an immune-related gene-based signature that may be useful as a prognostic biomarker for OSCC patients. Moreover, the insights from our research may provide leads for further experiments to understand the onset, progression and treatment of OSCC.

**Ethics and Consent to Participate**

The immunohistochemical image credit: Human Protein Atlas. The image available from Human Protein Atlas, the URL of each gene was shown in Supplementary Materials.

**Disclosure**

The authors report no conflicts of interest.

**References**

1. Kurman RJ. WHO classification of tumours of female reproductive organs, 4th ed. France: IARC, Lyon; 2014. Available from: https://www.ncbi.nlm.nih.gov/nlmcatalog/101656343. Accessed July 19, 2021.

2. Pogge von Strandmann E, Reintz S, Wager U, Muller R. Tumor-host cell interactions in ovarian cancer: pathways to therapy failure. Trends Cancer. 2017;3:137–148. doi:10.1016/j.trecan.2016.12.005

3. Yin M, Li X, Tan S, et al. Tumor-associated macrophages drive spheroïd formation during early transcoelomic metastasis of ovarian cancer. J Clin Invest. 2016;126:4157–4173. doi:10.1172/JCI87252

4. Yin M, Shen J, Yu S, et al. Tumor–Associated Macrophages (TAMs): a critical activator in ovarian cancer metastasis. Onco Targets Ther. 2019;12:8687–8699. doi:10.2147/OTT.S16355

5. Kim J, Bae JS. Tumor-associated macrophages and neutrophils in tumor microenvironment. Mediators Inflamm. 2016;2016:6058147. doi:10.1155/2016/6058147

6. Hui L, Chen Y. Tumor microenvironment: sanctuary of the devil. Cancer Lett. 2015;368:7–13. doi:10.1016/j.canlet.2015.07.039

7. Reintz S, Schumann T, Finkelkraut F, et al. Mixed-polarization phenotype of ascites-associated macrophages in human ovarian carcinoma: correlation of CD163 expression, cytokine levels and early relapse. Int J Cancer. 2014;134:32–42. doi:10.1002/ijc.28335

8. Buonaguro FM, Pauza D, Tornesello ML, Haimau P, Franco R, Marincola FM. Cancer diagnostic and predictive biomarkers. Biomed Res Int. 2014;2014:980163. doi:10.1155/2014/980163

9. McGranahan N, Rosenthal R, Hiley CT, et al. Allele-specific HLA loss and immune escape in lung cancer evolution. Cell. 2017;171:1259–1271.e1211. doi:10.1016/j.cell.2017.10.001

10. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res. 2012;41(D1):D991–D995. doi:10.1093/nar/gks1193

11. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47. doi:10.1093/nar/gkv007

12. Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol. 2015;19:A68–A77.

13. Breuer K, Foroushani AK, Laird MR, et al. InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. Nucleic Acids Res. 2015;41:D1228–D1233. doi:10.1093/nar/gks1147

14. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics. 2000;56:337–344. doi:10.1111/j.0006-341x.2000.00337.x

15. Uhlen M, Fagerberg L, Hallstrom BM, et al. Tissue-based map of the human proteome. Science. 2015;347:6220. doi:10.1126/science.1260419

16. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545–15550. doi:10.1073/pnas.0506580102

17. Libezron A, Birger C, Thovarldessdot H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MsigDB) hallmark gene set collection. Cell Syst. 2015;1:417–425. doi:10.1016/j.cels.2015.12.004

18. Xiao Y, Ma D, Zhao S, et al. Multi-omics profiling reveals distinct microenvironment characterization and suggests immune escape mechanisms of triple-negative breast cancer. Clin Cancer Res. 2019;25:5002–5014. doi:10.1158/1078-0432.CCR-18-3524

19. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–457. doi:10.1038/nmeth.3337

20. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218. doi:10.1186/s13059-016-1070-5

21. Cojoarcu E, Parkinson CA, Brenton JD. Personalising treatment for high-grade serous ovarian carcinoma. Clin Oncol. 2018;30:515–524. doi:10.1016/j.clon.2018.05.008

22. Lee EK, Konstantinopoulos PA. Combined PARP and immune checkpoint inhibition in ovarian cancer. Trends Cancer. 2019;5:524–528. doi:10.1016/j.trecan.2019.06.004

23. Kandalaft LE, Odunci K, Coukos G. Immune therapy opportunities in ovarian cancer. Am Soc Clin Oncol Educ Book. 2020;e228–e240.

24. Rea K, Roggiani F, De Cecco L, et al. Simultaneous E-cadherin and PLEKHA7 expression negatively affects E-cadherin/EGFR mediated ovarian cancer cell growth. J Exp Clin Cancer Res. 2017;36:148. doi:10.1186/s13046-016-0796-1

25. Wang Y, Zhang X, Tang W, et al. mIR-130a upregulates mTOR pathway by targeting TSC1 and is transactivated by NF-kappaB in high-grade serous ovarian carcinoma. Cell Death Differ. 2017;24:2089–2100. doi:10.1038/cdd.2017.129

26. Yamanot K, Baba T, Abiko K, et al. Acquisition of a side population fraction augments malignant phenotype in ovarian cancer. Sci Rep. 2019;9:14215. doi:10.1038/s41598-019-50794-w

27. Li D, Sun Y, Kong X, et al. Association between a single nucleotide polymorphism in the 3'-UTR of ARHGEF18 and the risk of nonischaemic pulmonary arterial hypertension in Chinese population. Dis Markers. 2018;2018:2461845. doi:10.1155/2018/2461845

28. Zhang S, Royer R, Li S, et al. Frequencies of BRCA1 and BRCA2 mutations among 1342 unselected patients with invasive ovarian cancer. Gynecol Oncol. 2011;121:353–357. doi:10.1016/j.ygo.2011.01.020

29. Jia R, Chai P, Wang S, et al. m6A modification suppresses oral melanoma through modulating HINT2 mRNA translation. Mol Cancer. 2019;18:161. doi:10.1186/s12934-019-1088-x

30. Nayak AP, Kapur A, Barroilhet L, Patankar MS. Oxidative phosphorylation: a target for novel therapeutic strategies against ovarian cancer. Cancers. 2018;10:337. doi:10.3390/cancers10090337
31. Zhu LY, Zhang WM, Yang XM, et al. Silencing of MICAL-L2 suppresses malignancy of ovarian cancer by inducing mesenchymal-epithelial transition. *Cancer Lett*. 2015;363:71–82. doi:10.1016/j.canlet.2015.04.002

32. Wade CA, Kyprianou N. Profiling prostate cancer therapeutic resistance. *Int J Mol Sci*. 2018;19:904. doi:10.3390/ijms19030904

33. Drakes ML, Stiff PJ. Regulation of ovarian cancer prognosis by immune cells in the tumor microenvironment. *Cancers*. 2018;10:302. doi:10.3390/cancers10090302