Identification of Key genes associated with polycystic ovary syndrome (PCOS) and ovarian cancer using an integrated bioinformatics analysis

Juan Zou
University of South China
Yukun Li
university of south china
Nianchun Liao
University of South China
Jue Liu
University of South China
Qunfeng Zhang
University of South China
Min Luo
University of South China
Jiao Xiao
University of South China
Yanhua Chen
University of South China
Mengjie Wang
University of South China
Kexin Chen
University of South China
Juan Zeng (1076006678@qq.com)
University of South China
Zhongcheng Mo
Guilin Medical University

Research Article

Keywords: Ovarian cancer, Polycystic ovary syndrome, Bioinformatic analysis, OGN, Prognostic marker

Posted Date: September 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-934214/v1

Abstract

Background

Accumulating evidence suggested a strong association between polycystic ovary syndrome (PCOS) and ovarian cancer (OC), but the potential molecular mechanism is still unclear. In this study, we identify unrecognized but significant genes correlated to PCOS and OC via bioinformatics.

Materials and methods

Multiple bioinformatic analysis, such as Differential expression analysis, Univariate Cox analysis, functional and pathway enrichment analysis, protein–protein interaction (PPI) network construction, survival analysis, and Immune infiltration analysis were utilized. We further evaluated the effect of OGN on FSHR expression via immunofluorescence.

Results

The TCGA-OC dataset, GSE140082 (for OC) and GSE34526 (for PCOS) dataset were downloaded. 12 genes, RNF144B, LPAR3, CRISPLD2, JCHAIN, OR7E14P, IL27RA, PTPRD, STAT1, NR4A1, OGN, GALNT6 and CXCL11, were recognized as signature genes. Drug sensitive analysis was showed that OGN might be a hub gene in the progression of PCOS and OC. Experimental analysis found OGN could increase the FSHR expression, indicating OGN could regulate the hormonal response in PCOS and OC. Furthermore, correlation analysis indicated that the function of OGN might be closely related with m6A and ferroptosis.

Conclusions

Our study indicated 12 signatures that might involved in the prognosis significance of OC, and closely related the correlation between OC and PCOS. Furthermore, the hub gene, OGN, was a significant gene in the OC and PCOS progression via regulating the hormonal response.

Introduction

Polycystic ovary syndrome (PCOS), a multi-system reproductive metabolic disease of the reproductive system, is characterized by the pathological accumulation of non-maturating and atretic follicles, ovarian and stromal abnormal hyperplasia, hyperandrogenemia (HA), hyperinsulinemia, insulin resistance (IR), metabolic aberrant, an imbalance ratio of luteinizing hormone (LH) to follicle-stimulating hormone (FSH), and polycystic ovary(1). The mRNA and microRNA profiles of PCOS patients were extremely similar with ovarian cancer (OC) patients, which indicated that the same molecular mechanisms might be between OC and PCOS patients(2, 3).

In the progression of PCOS patients, HA is an important factor for promoting ovulatory dysfunction(4), increasing frequency and amplitude of LH and GnRH pulse secretion(5), inducing lipid metabolism disorders(6), mediating hyperinsulinemia and insulin resistance(7), and dysregulating the ratio of LH to FSH(8).

For the clinical management of patients with PCOS, the anti-androgen therapy is the first line of treatment for patients diagnosed with PCOS(9). Over recent years, the relationship between the PCOS and OC progression has been a hot topic for these study due to AR signaling axis and metabolic disorders are correlated with a high risk of OC(3, 10, 11). Both OC and PCOS are multifactorial diseases, with genetic, endogenous, endocrine maladjusted, metabolically disturbed and environmental factors. Therefore, a better understanding of the physiopathologic mechanism regulating these complex molecular effects is urgently needed to promote the research and development of new drugs, and to improve these patient's prognosis.

With the development of bioinformatic analysis and public databases, such as The Cancer Genome Atlas (TCGA)(12) and Gene Expression Omnibus (GEO)(13), understanding the molecular mechanisms of currently available treatments against PCOS and OC provides a means to emphasize targets for effective treatments. For example, Surleen Kaur found that in PCOS tissues there is amount of differentially expressed genes correlated to metabolic disorders and oxidative stress, and with a potential relationship with cancer as well(14). HSA2 and CBLN1 were all identified in an PCOS dataset(15). Another study identified 36 highly altered genes, among them 10 were shared among endometrial cancer (EC), OC and breast cancer (BC), resulting in cell proliferation, hormone response, and endogenous stimulation(16). A series of bioinformatics tools were used for integrated analysis and detection of metabolism related genes (MRGs) in OC. For example, we found that ENPP1, FH, CYP2E1, HPGDS, ADCY9, NDUFA5, ADH1B and PYGB were correlated with the underlying mechanisms of metabolic reprogramming in OC progression(17). Yang et al. found that CCNB2, TYMS, KIF11, KIF4A, BIRC5, BUB1B, FOXM1, and CDC20 might be potential therapeutic targets for OC patients(18). Nevertheless, it is still unclear as to whether these hub genes are uniquely involved in individual disease progression or not.

To determine potentially molecular mechanisms between PCOS and OC, we integrated two datasets, including PCOS and OC. Utilizing multiple bioinformatic and experimental analysis, we hence sought to validate hub genes and pathway of interest, and to search for potential therapeutic drugs or targets in PCOS and OC.

Materials And Methods

Data extraction

TCGA database(12) (https://portal.gdc.cancer.gov/) is the largest cancer gene information database, and it includes data concerning gene expression. We extracted data for 374 cases of patients with OSC. Moreover, we downloaded level three FPKM data for subsequent analysis. The transcriptome RNA-sequencing and clinical information of 88 normal ovarian samples were extracted from the GTEX

database (https://www.gtexportal.org/)(19). Furthermore, the GSE140082 and GSE34526 dataset was downloaded from GEO database(13).

Functional enrichment analyses

The functional enrichment analyses was also performed as previously published(17).

PPI network construction

GeneMANIA (http://genemania.org)(20) was used to construct the PPI network for 128 DEGs.

Establishing prognostic indicators based on DEGs

Univariate Cox analysis was used to select genes associated with prognosis, and prognostic correlation model was further constructed. After incorporating the expression value of each specific gene, a risk score formula was constructed for each patient. According to the risk score formula, patients were divided into low-risk group and high-risk group, with the median risk score as the cut-off point. Kaplan-meier analysis was used to evaluate the survival difference between the two groups, and log-rank statistical method was used to compare. Finally, receiver operating characteristic (ROC) curve was used to study the accuracy of model prediction.

Protein expression of OGN by bioinformatic analysis

The protein expression of OGN in OC was confirmed by HPA database (https://www.proteinatlas.org/)(21) and CPTAC database (https://cptac-data-portal.georgetown.edu/)(22).

The Relationship of key Genes and Immune infiltration

The correlation between immune cell content and the level of 5 key genes (JCHAIN, CXCL11, OGN, STAT1, and GALNT6) was confirmed by TIMER database (https://cistrome.shinyapps.io/timer/)(23).

The GSVA and GSEA analysis

Gene set variation analysis (GSVA) is a nonparametric and unsupervised method for assessing the enrichment of transcriptome gene sets. GSVA converts gene-level changes into pathway-level changes by comprehensively scoring the sets of genes of interest to judge the biological function of the samples. In this study, gene sets will be downloaded from The Molecular Signatures Database (v7.0), and each gene set will be comprehensively scored by GSVA algorithm to evaluate the potential biological function changes of different samples.

GSEA analysis uses a predefined set of genes, sequencing genes according to how differentially expressed they are in the two types of samples, and then testing whether the predefined set of genes is enriched at the top or bottom of the sequencing table. In this study, the possible molecular mechanism of the difference in prognosis of different patients in ovarian cancer was explored by comparing the

difference in signal pathways between the high-expression group and the low-expression group of GSEA, in which the number of replacements was set to 1000 and the replacement type to phenotype.

Cell culture and transfection

Human ovarian cancer cell line (SKOV3 and KGN) were bought from the American Type Culture Collection (ATCC, VA, USA). Dulbecco's Modified Eagle's Medium (DMEM) containing 10% (v/v) fetal bovine serum (FBS; Gibco, Invitrogen, Carlsbad, CA, USA) and 1% penicillin/streptomycin (GIBCO, CA, USA,) growth media was used for SKOV3 and KGN cell culture. All cells were incubated at 37°C and 5% CO2. The over-expressed OGN and empty vector plasmids were purchased from GeneCopoeia Biotechnology (GeneCopoeia Biotechnology, MD, USA). For transient cell transfection, SKOV3 and KGN cells were seeded in 6-wll plates for 24 hours. After incubation, cells were transfected with 3 μg empty vector and 3 μg over-expressed OGN plasmid using Lipfectamine 3000 (Invitrogen, Carlsbad, CA, USA) according to the protocol to establish a cell line with upregulated OGN expression.

qRT-PCR analysis

Total RNA was extracted by TRIzol (Invitrogen, CA, USA) according to the manufacturer's instructions. cDNA was produced by reverse transcription kit (TaKaRa, Dalian, China). PCR were carried out in ABI 7500 fast system (Applied Biosystems, CA, USA). Primer sequences were as follows: OGN Forward: 5´-TCTACACTTCTCCTGTTACTGCT-3´, OGN Reverse: 5´-GAGGTAATGGTGTTAT TGCCTCA-3´;

Immunofluorescence

The immunofluorescence assays were performed with anti-FSHR (Abcam, 1:300) according to the protocol. The primary antibodies used in this study were against FSHR (ab113421). Incubated with the corresponding FITC conjugated secondary antibodies (Abcam, 1:200). Two hours later, 0.1%DAPI stained the nucleus for 30 minutes. Images were detected by confocal microscopy (Leica, Jena, Germany).

Statistical analysis

All statistical analyses were performed in the R language (Version 3.6). All statistical tests were bilateral, and $P < 0.05$ was statistically significant

Results

Identification of common 128 significant differentially expressed genes（DEGs）in PCOS and OC

Firstly, we found 1061 DEGs in the PCOS patients compared to normal women based on GSE34526 dataset of GEO database, and 2254 DEGs in the OC patient samples compared to normal ovary samples based on OC dataset of TCGA database (Figure1 A&B). Moreover, we found that there were 128 common DEGs in PCOS and OC progression (Figure1 C). We also constructed the protein–protein interactions (PPI) network to identify all of 120 genes in the dataset of PCOS and OC, which were visualized by GeneMANIA

database (Figure1 D), which indicated that these genes have closely interactions. The PCA analysis found that these DEGs expression could well discriminates between ovarian cancer (blue) and normal (red) (Figure1 E). We extracted GO and KEGG pathway data for these genes based on DAVID database. In the term of GO enrichment, these genes were enriched in cell adhesion molecule binding, actin binding, cadherin binding, actin filament binding, cell-substrate junction, cell-substrate adherens junction, focal adhesion, collagen-containing extracellular matrix, antigen processing and presentation, and so on (Figure1 F). In the term of KEGG enrichment, these genes were enriched in cell adhesion molecules, staphylococcus aureus infection, hematopoietic cell lineage, viral myocarditis, and Asthma (Figure1 G). In summary, these results indicated that common DEGs highlighted the significant role of cell adhesion in the relationship between PCOS and OC.

Evaluation of clinical outcomes in OC based on the common 128 DEGs

The common 128 DEGs were used to analyse the prognosis in OC patients by the univariate Cox method. A number of twelve key genes were closely associated to the prognosis of OC patients, including RNF144B, LPAR3, CRISPLD2, JCHAIN, OR7E14P, IL27RA, PTPRD, STAT1, NR4A1, OGN, GALNT6 and CXCL11 (Figure 2A). Then, we used these expression profile to construct the prognosis model, and the risk score formula was as follow: Risk score = RNF144B*(-0.1441) + LPAR3*(-0.0187) + CRISPLD2*0.0701 + IL27RA*0.2226 + PTPRD*0.0055 + STAT1*(-0.0988) + NR4A1*0.0369 + OGN*0.0590 + GALNT6*(-0.0718) + CXCL11*(-0.0886). Next, we could divide these OC patients into high risk and low risk group with the median risk score based on the risk score formula (Figure 2B).

The survival score and status of the two groups in the training cohort based on TCGA database OC datasets were shown in Figure 3A&B. These twelve key genes expression profiles were shown by the heatmap (Figure 3C). Moreover, we used the GSE140082 dataset as a test cohort to validate the risk score formula, which survival score and status of high risk and low risk group were shown in Figure 3D&E. These key genes expression files in GSE140082 datasets was also visualization by the heatmap (Figure 3F).

In the training cohort, the survival time and rate were significantly decreased with the risk score increased (Figure 4A). The AUC at 1, 2, and 3 years under the ROC curve were 0.571, 0.607, and 0.554, respectively, indicating that a moderate incubation period could be utilized as a prognostic marker of twelve key genes expression profiles in survival monitoring (Figure 4B). However, t-SNE analysis showed that OC patients in different risk groups were not distributed in two group based on TCGA database, which suggested that the 12 signatures could not be an excellent subtype marker (Figure 4C). To validate the efficiency of the prognosis model constructed from the TCGA-OC cohort, we used the median value of training cohort to divide the OC patients into high risk and low risk group based on the GSE140082 cohort. Similar to the results of training cohort, OC patients with high risk had a poor prognosis compared to other OC patients in the low risk group (Figure 4D). The AUC value in 1, 2, and 3 years were 0.617, 0.682, and 0.651 in the test cohort (Figure 4E). The t-SNE analysis was also similar to the training cohort (Figure 4F).

The ectopic expression and prognosis significance of 12 signatures in OC patients

Next, we used the boxplot to visualize the mRNA level of 12 signatures in OC samples, which indicated that LPAR3, JCHAIN, IL27RA, GALNT6, CXCL11, RNF144B, STAT1, OR7E14P were significantly increased in OC patients, but CRISPLD2, PTPRD, OGN, NR4A1 were obviously decreased in patients with OC (Figure 5A). We also confirmed the overall survival rate of 12 signatures in OC patients based on TCGA database, suggesting that OGN was significantly and negatively correlated with OC patient's prognosis, but JCHAIN, GALNT6, CXCL11, STAT1 were significantly and positively correlated with prognosis in OC patients (Figure 5B). These results suggested that JCHAIN, GALNT6, CXCL11, STAT1, and OGN might play a key role in the progression of OC patients.

The DNA alteration and immune infiltration of 5 key genes in OC progression

We found the 5 genes were genetically altered, such as missense mutation, amplification and deep deletion (Figure 6A). The CNV of JCHAIN was significantly correlated with CD8+ T cell, Neutrophil, and Dendritic cell. CXCL11 CNV was closely associated with CD8+ T cell, CD4+ T cell, Neutrophil, and Dendritic cell. The CNV level of OGN was markedly related to Macrophage. STAT1 CNV level had a closely relationship with CD8+ T cell and Dendritic cell. The CNV of GALNT6 was significantly associated with B cell, CD8+ T cell and CD4+ T cell (Figure 6B). Furthermore, we found the mRNA expression of GALNT6 was not obviously correlated with immune infiltration in any immune cell types. JCHAIN level was closely associated with purity, CD8+ T cell, CD4+ T cell, Neutrophil, and Dendritic cell. CXCL11 expression was correlated with the infiltration of purity, B cell, CD8+ T cell, CD4+ T cell, Neutrophil, and Dendritic cell. OGN level was significantly correlated with purity. STAT1 mRNA level had a closely relationship with purity, CD8+ T cell, Neutrophil, and Dendritic cell (Figure 6C). Taken together, the expression and alteration of these 5 key genes was involved in the immune infiltration progression of OC.

The drug sensitivity of hub gene

We further used the drug sensitivity analysis to confirm these 5 key genes. The result showed that OGN was closely correlated with chemotherapy resistance based on GSCALite database (Supplementary Figures S1). Therefore, targeting OGN could be a potential target in the treatment of patients with OC or PCOS.

The characteristics of OGN in OC and PCOS

For elucidate the expression, function and structure of OGN, we used the PDB databased to confirm the OGN structure, as shown in Figure 7A. The OGN has an LRR_8 domain and multiple phosphorylation, acetylation and N-linked glycosylation site. The protein expression was significantly decreased in OC tissues samples compare to normal ovary samples (Figure 7B&C). We further utilized GSVA and GSEA analysis to predict the potential function of OGN, as shown in Figure 7D&E. OGN might be involved in the progression of steroid hormone biosynthesis and steroid hormone response. Furthermore, we found the level of OGN was significantly and positively correlated with the level of FSHR in OC (Figure 7F). We overexpressed OGN in KGN and SKOV3 cell lines (Figure 7G), and confirmed the effect of OGN on FSHR

expression by IF (Figure 7H). The results indicated that OGN played a key role in PCOS and OC progression by upregulating FSHR level.

OGN level is correlated with regulators of ferroptosis and m6A methylation in OC

The ferroptosis and m6A methylation was involved in the development and progression of OC. We firstly used the TCGA database to analyze whether OGN level is correlated with ferroptosis. We make a correlation analysis between the OGN expression and 25 ferroptosis genes in OC and ovary tissue samples based on TCGA and GTEx database (Figure 8A). The results showed that the expression of 25 ferroptosis genes were significantly between OC and normal ovary. Furthermore, we also confirmed the correlation of ferroptosis genes with OGN in OC samples. We found that the level of OGN was positively correlated with MT1G, HSPB1, GPX4, FDFT1 and ATP5MC3, but negatively correlated with CDKN1A, HSPA5, SLC1A5, NCOA4, LPCAT3, DPP4, ALOX15, ACSL4, and ATL1 (Figure 8B). Then, we extracted the expression profiles for the 20 m6A methylation genes between OC and normal ovary samples based on TCGA and GTEx database (Figure 8C), which indicated that these m6A methylation regulators were played a key role in OC progression. We further made a correlation analysis between these m6A regulators and OGN expression. The result showed that METTL14, WTAP, VIRMA, RBM15, RBM15B, ZC3H13, YTHDC1, YTHDC2, YTHDF3, YTHDF1, IGF2BP2, HNRNPA2B1, FTO, and ALKBH5 were positively and significantly associated with OGN expression in OC patients (Figure 8D). Taken together, OGN might be had another important function on OC ferroptosis and m6A methylation modifications.

Discussion

OC is a lethal malignancy in gynecological diseases, which is a complex disease, and multiple metabolic enzymes and pathways are involved in its pathophysiology(24). PCOS is a benign gynecological disease with multi-system metabolic disorder, which is characterized mainly by HA, IR, LH and FSH ratio imbalance, infertility, endometrial disorder, obesity, and polycystic ovary(1). We found that these MRGs, such as ENPP1, FH, CYP2E1, HPGDS, ADCY9, NDUFA5, ADH1B and PYGB, were involved in the development and progression of OC, which could construct to a prognosis model to predict the postoperative risk for OC patients(17). Zhou et al also supports this view, and they found that PYGB could promote the development of OC via activating Wnt/β-catenin signaling pathway(25). Moreover, many studies suggested that alternative AR activating signals, including both ligand dependent or independent, were involved in OC progression(26, 27). Therefore, the underlying pathophysiological association between PCOS and OC might be a key way to formulate clinical treatment strategy for OC or PCOS.

In this study, we found 128 shared DEGs between PCOS and OC when compared to corresponding normal tissue samples. These 128 genes were significantly enriched in cell adhesion molecule binding via GO and KEGG analysis. Twelve genes, RNF144B, LPAR3, CRISPLD2, JCHAIN, OR7E14P, IL27RA, PTPRD, STAT1, NR4A1, OGN, GALNT6, and CXCL11, were validated as key genes related to the prognosis of OC patients. A prognostic score based on the twelve genes obviously classified OC patients into high risk or low risk group. Moreover, we found high risk group had a poor prognosis compared to low risk group.

Given the underlying molecular mechanisms of these MRGs, studies on the functions and mechanisms of RNF144B, JCHAIN, OR7E14P, IL27RA, PTPRD, and OGN have not been confirmed in OC progression. But another six genes have been elucidated in OC development, namely, LPAR3, CRISPLD2, STAT1, NR4A1, GALNT6, and CXCL11. The expression of LPAR3 is significantly increased in OC tissue samples compared to normal ovary tissue samples, which might play a role in carcinogenesis of ovarian cancer(28). High expression of CRISPLD2 was correlated with the worse prognosis in OC patients(29). High levels of STAT1 was associated with the improved prognosis in OC patients, which might be using for the development of new immunomodulator drugs to OC treatment(30). Ectopic expression of NR4A1 protein was significantly correlated with poor prognosis in patients with OC(31), which could modulate platinum resistance in OC(32). GALNT6 could modify O-glycans on EGFR to increase its activation, which was able to significantly enhance the OC cell viability, migration, and invasion(33). Many studies indicated that CXCL11 could promote the OC progression via mediating angiogenesis(34), lymph node metastasis(35), and immune infiltration(36, 37). Moreover, we found RNF144B, JCHAIN, OR7E14P, IL27RA, PTPRD, OGN, CRISPLD2, and GALNT6 have not been confirmed in PCOS progression. STAT1 could co-interact with CD44-OPN adhesion complex, ERα, and NF-κB to formulate a significant crosstalk, resulting in modulating endometrial receptivity(38). Insulin could enhance the expression of STAT1 and STAT3 to repress the level of miR-27a-3p which could decrease granule cells proliferation ability and apoptosis escaping ability(39). High level of androgen could mediate a series of important genes, including TFAP2A, ETS1, ELK1, ERG, FLI1 and SPI1, to increase the level of NR4A1 in PCOS(40). CXCL11 expression was significantly and obviously correlated with Prolactin and 17-OH-progesterone level in PCOS(41). Based on previous studies and our study, we were able to obtain only limited information about the 12 key genes involved in OC patient survival and in PCOS patients pathophysiological change.

To seek efficient therapeutic agents for the treatment of OC and PCOS, we further screened these 12 key genes by expression, survival significance, immune infiltration, and drug sensitivity. Osteoglycin (OGN), a small proteoglycan with tandem leucine-rich repeats (LRR), is overexpressed in blood vessels and bone, mediating bone formation mediating by transforming growth factor beta(42). The function of OGN is involved in the progression of extracellular matrix and remodeling and tissue development(42, 43). Katja Hummitzsch et al found that the expression of OGN mRNA was significantly upregulated in ovary theca internal compared to stroma(44). Hao and his colleagues further found that OGN protein expression was markedly associated with signaling pathway related to follicular development, especially in estrogen, insulin, and PI3K-Akt signaling pathway(45). We found that OGN mRNA and protein expression were significantly decreased in OC/PCOS compared to normal ovary tissues or granule cells from woman with non-PCOS. Taken together, these results indicated OGN might be a significant factor in the progression of OC and PCOS. Furthermore, our GSVA and GSEA analysis showed that highly expressed OGN could also enhance steroid hormone biosynthesis, and correlation analysis indicated that OGN was significantly associated with the level of FSHR. Many studies have been indicated that FSHR was a significant oncogene in OC development, which could mediate the cell proliferation(46, 47), metastasis(48), and apoptosis escaping(49). Moreover, the FSHR level was significantly decreased in PCOS granulosa cells compared to normal granulosa cells(50). FSHR is critical for FSH mediated follicle growth and

development, and the decrease of FSH/FSHR pathway might be induce the follicular growth arrest, promoting the progression of PCOS(51, 52). Therefore, OGN might upregulate the FSHR to sensitize the steroid hormone response, which could accelerate the OC formation and progression, but reverse the progression of PCOS. The low expression of OGN is an important feature from PCOS to OC, indicating that the PCOS patient with low level of OGN may has the greater OC risk. The aberrant activation of OGN has been associated with high tumor invasiveness and poor prognosis in the OC patients(53). Moreover, the function of OGN in OC progression is closely related with m6A modification and ferroptosis. In summary, we infer that OGN might be a new risk marker for PCOS to OC, although this assumption needs to be verified in further studies.

Conclusions

In conclusion, this study provided evidence about the association between PCOS and OC. Through the functional analysis of identified DEGs, we found that cell adhesion was significantly enriched in PCOS and OC datasets. We also confirmed 12 key genes to construct the prognosis model, which divide OC patients into high risk and low risk group for the prognosis prediction. Moreover, we found OGN might be a key biomarker, indicating the greater OC risk for PCOS patients. Nevertheless, further experimental verification is required.

Abbreviations

polycystic ovary syndrome (PCOS), hyperandrogenemia (HA), insulin resistance (IR), luteinizing hormone (LH), follicle-stimulating hormone (FSH), ovarian cancer (OC), The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), endometrial cancer (EC), breast cancer (BC), metabolism related genes (MRGs), protein–protein interactions (PPI), Osteoglycin (OGN)

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The data used to support the findings of this study are available from the corresponding author upon request

Competing interests

The authors declare that they have no competing interests

Funding

The present study was supported by the Hunan Provincial Health Commission 2019 Annual Scientific Research (Grant No. C2019098).

Authors' contributions

Conception and design: Yukun Li, Juan Zou and Nianchun Liao. Collection and assembly of data: Jue Liu, Qunfeng Zhang, and Jiao Xiao. Data analysis and interpretation: Yanhua Chen, Mengjie Wang, Kexin Chen, and Min Luo. Manuscript writing: Yukun Li and Juan Zou. Paper revision: Juan Zeng and Zhongcheng Mo. Final approval of manuscript: all authors.

Acknowledgements

Not applicable.

Figures

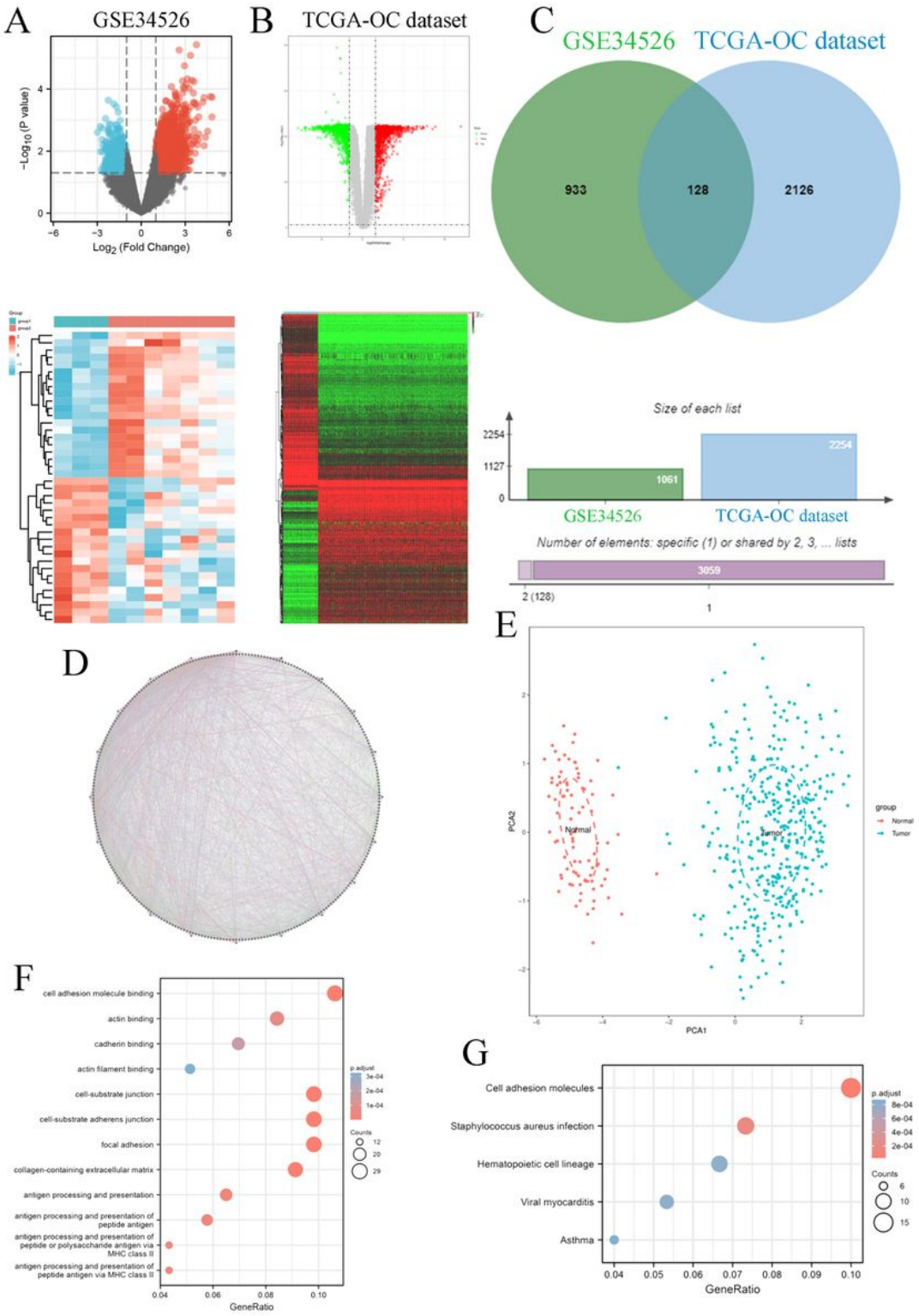

Figure 1

Identification of the DEGs associated PCOS and OC. A. The DEGs in PCOS based on GSE34526 datasets. B. The DEGs in OC based on TCGA-OC datasets. C. The common DEGs in PCOS and OC. D. The PPI network of 128 common DEGs in PCOS and OC. E. The PCA analysis between OC patient samples (TCGA-OC dataset) and normal ovary samples (GTEx-ovary datasets) based on 128 DEGs. F. The GO functional enrichment analysis for 128 DEGs. G. The KEGG enrichment analysis for 128 DEGs.

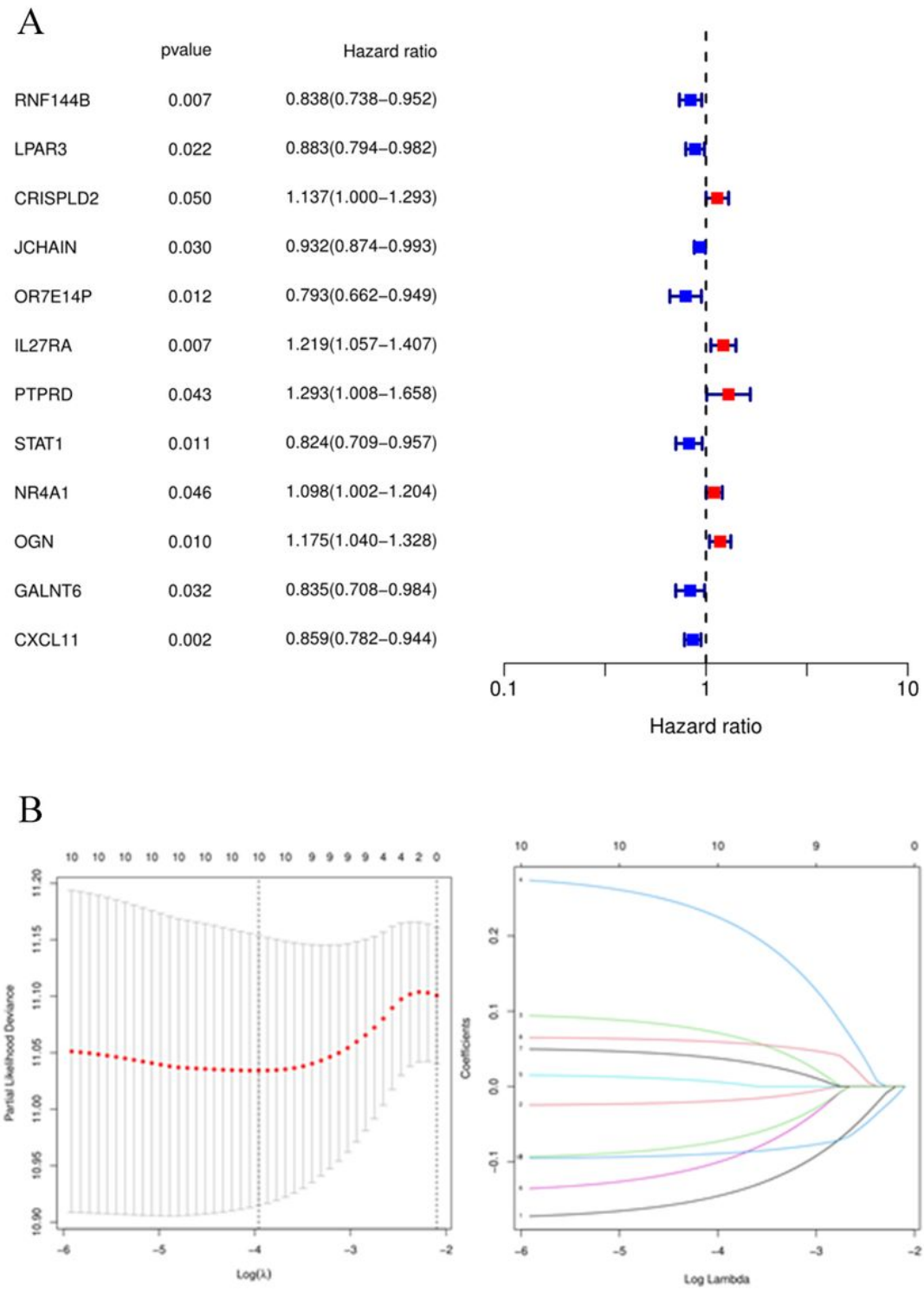

Figure 2

Key genes prognostic values. A. Prognostic values of 12 genes via forest plot. B. Prognostic signatures construction by lasso Cox analysis.

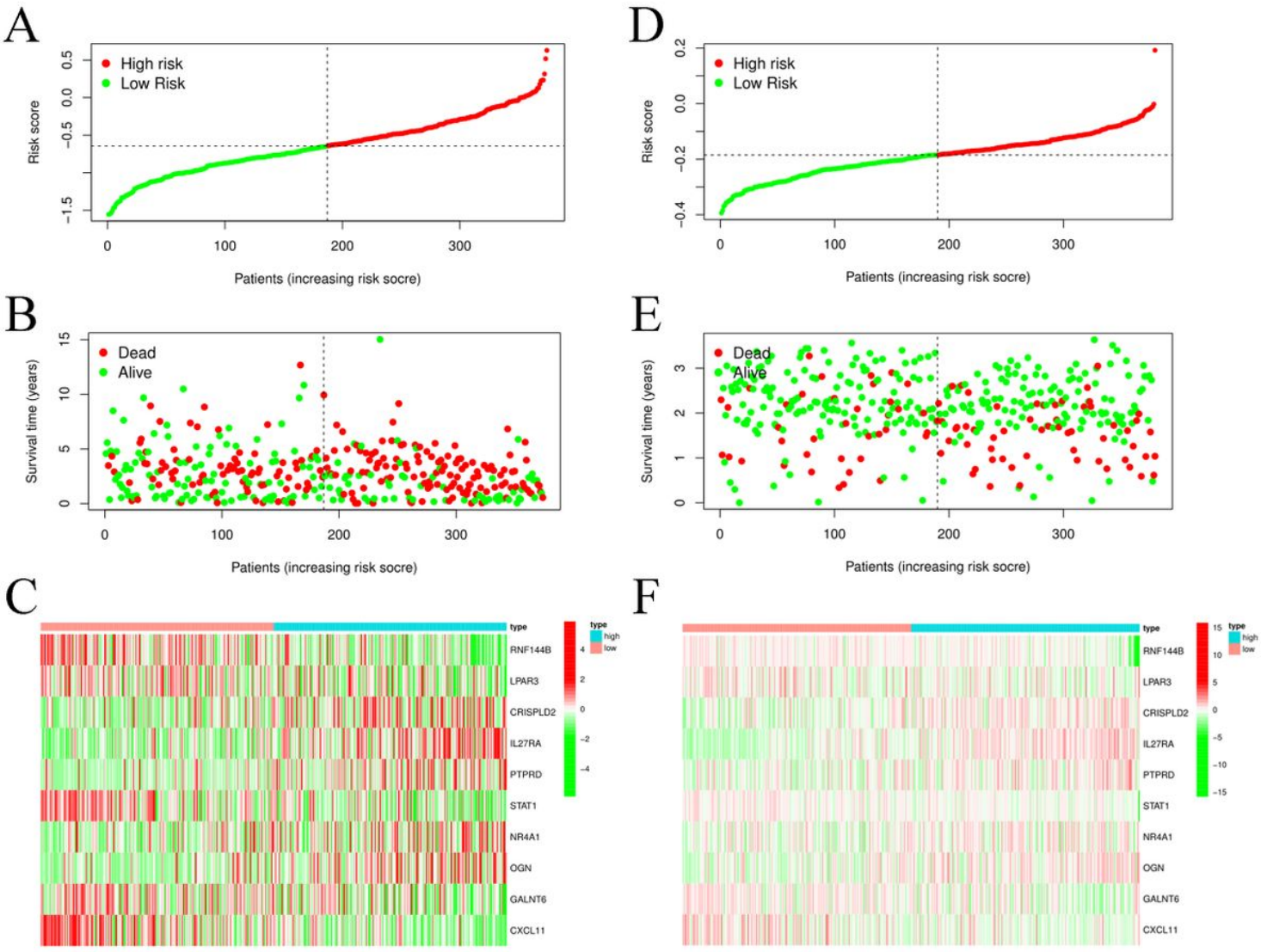

Figure 3

12 prognostic index of OC patients. A. The PI distribution of patients in the training dataset. B. The survival of OC patients in the training dataset. C. The expression profiles of 12 key genes in training dataset. D The PI distribution of patients in the test dataset. E. The survival of OC patients in the test dataset. F. The expression profiles of 12 key genes in test dataset.

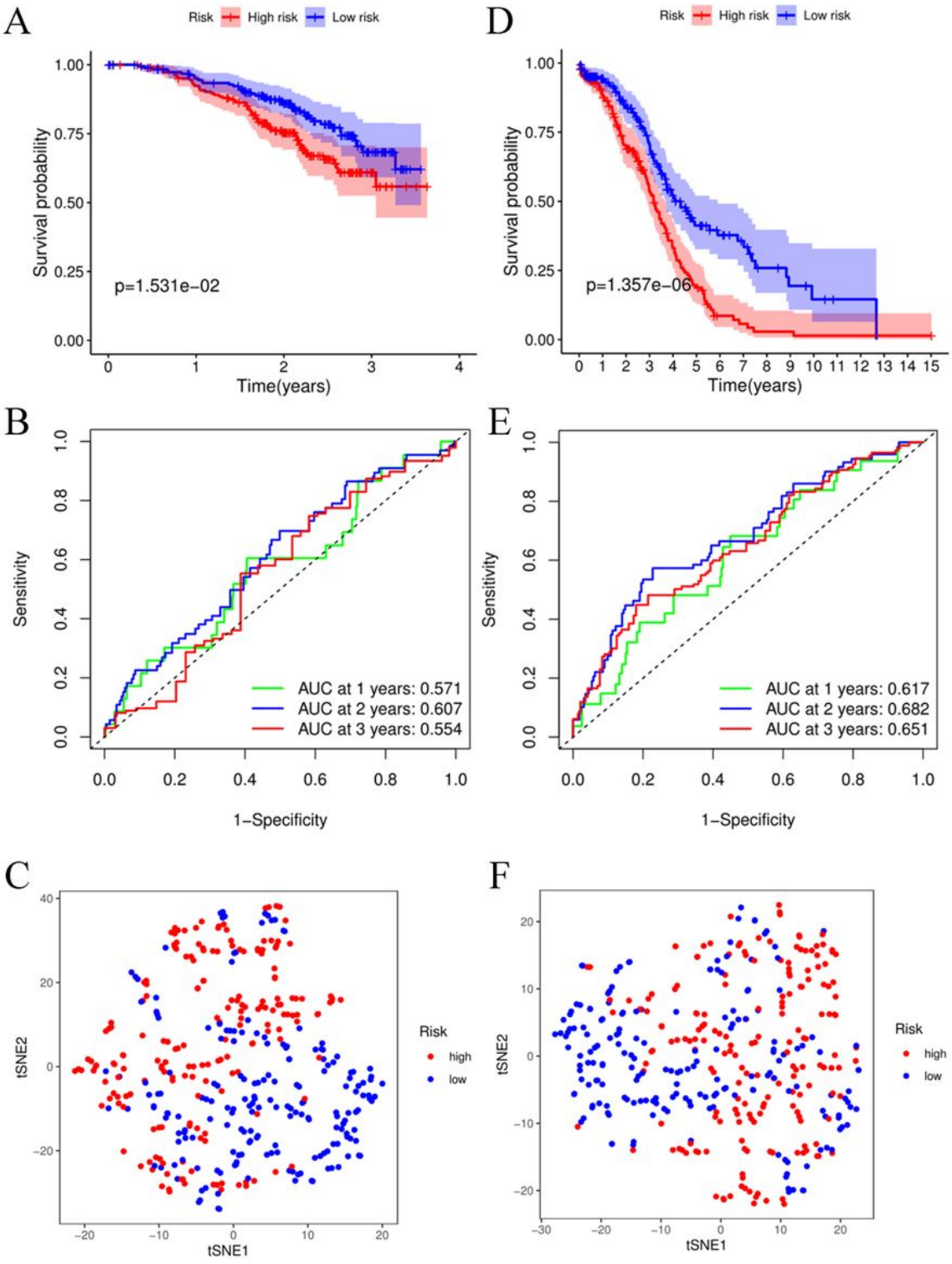

Figure 4

Prognostic analysis of the 12-gene signature model in the training cohort and test cohort. A. Kaplan-Meier curves for OS of patients with high or low risk groups in the training cohort. B. AUC time-dependent ROC curves for OS in the training cohort. C. t-SNE analysis for OS in the training cohort. D. Kaplan-Meier curves for OS of patients with high or low risk groups in the test cohort. E. AUC time-dependent ROC curves for OS in the test cohort. F. t-SNE analysis for OS in the test cohort.

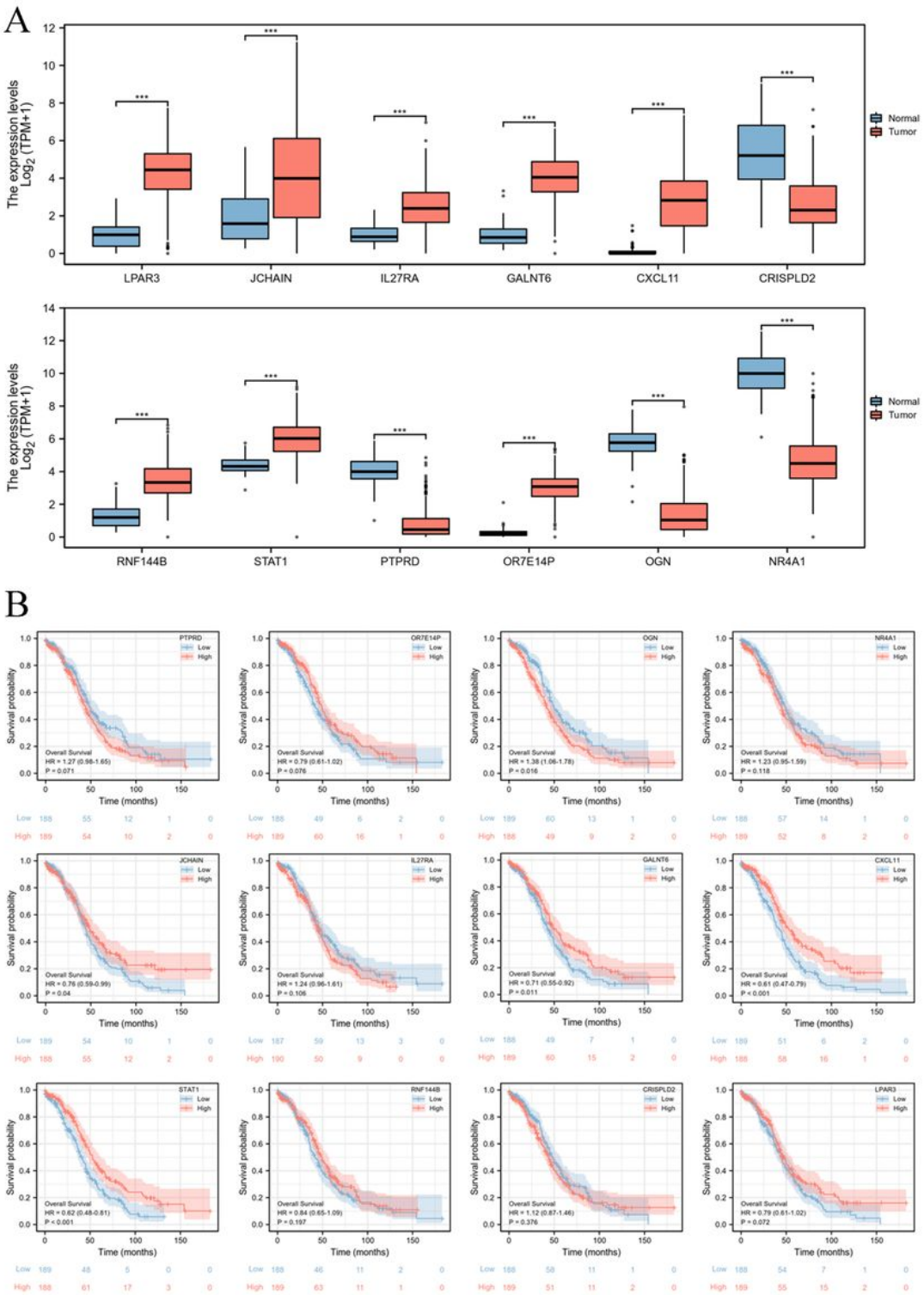

Figure 5

The expression and prognosis significance of 12 signatures. A. The mRNA expression of 12 signatures based on TCGA database. B. The prognosis significance of 12 signatures based on TCGA database.

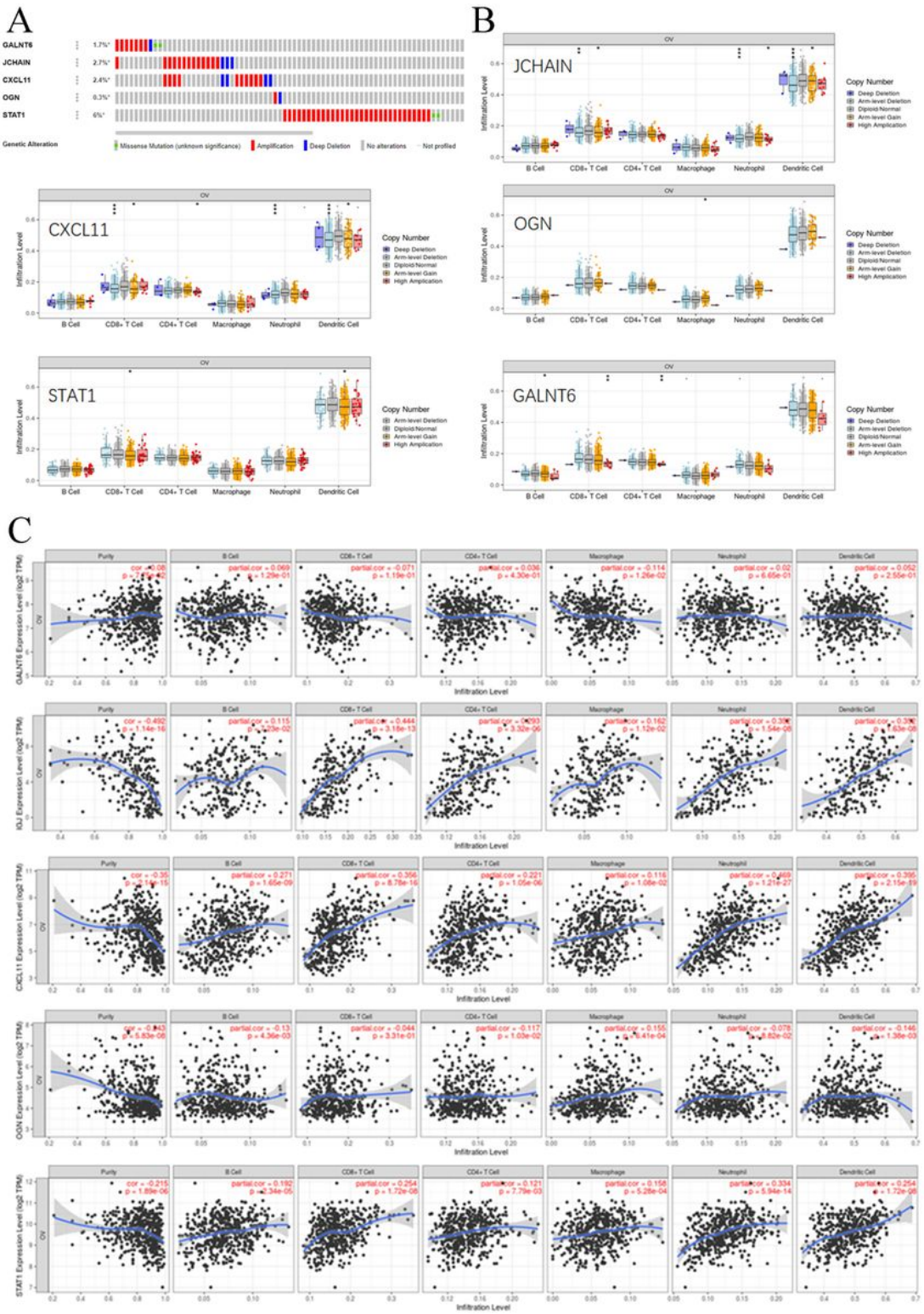

Figure 6

The DNA alteration and immune infiltration of 5 key genes. A. The DNA alteration of 5 key genes. B. The CNV affecting the distribution for the 5 genes. C. The cancer purity and immune infiltration for 5 genes.

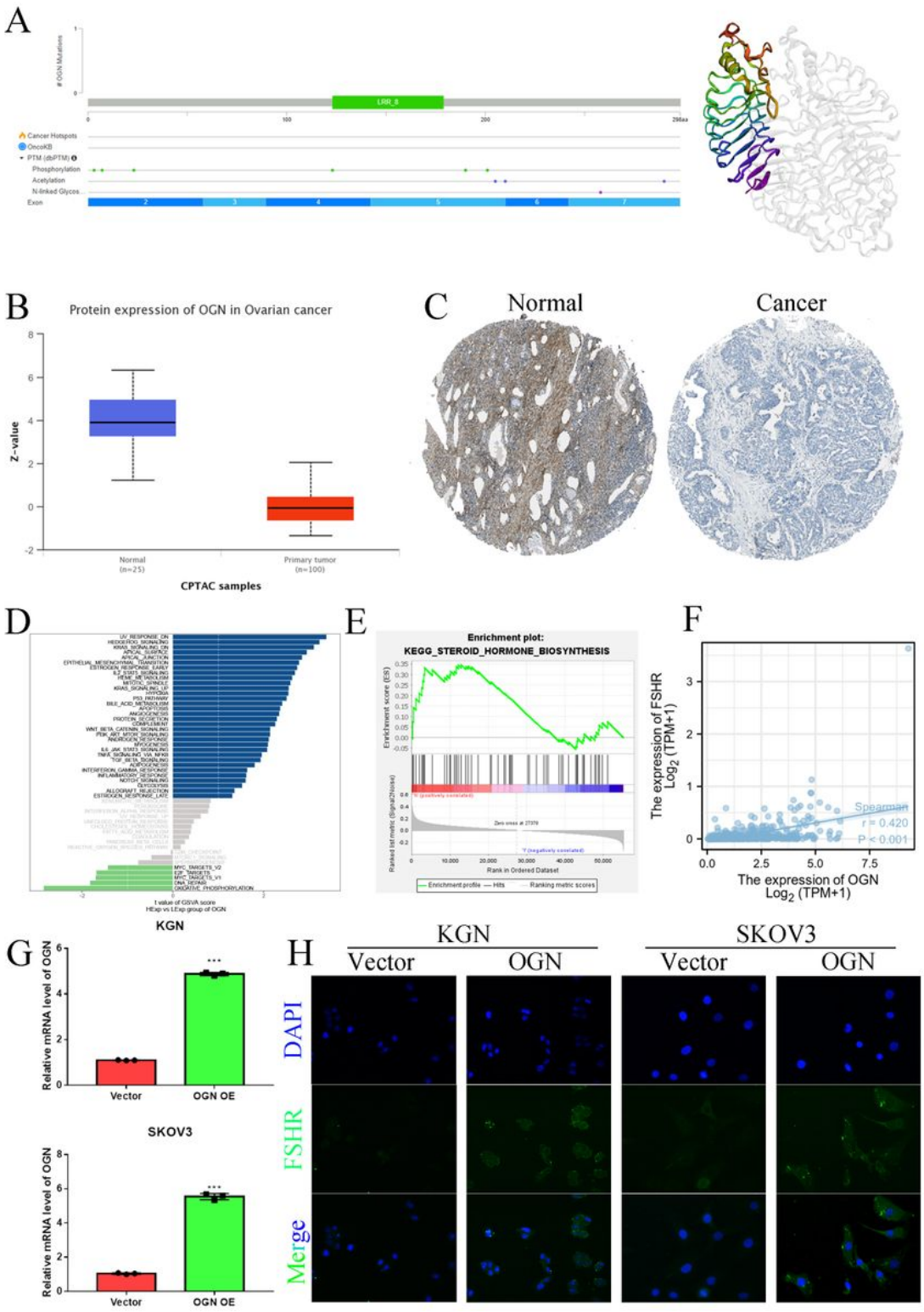

Figure 7

The structure, expression and function of OGN. A. The Structure of OGN. B. The OGN protein expression in OC based on CPTAC database. C. The OGN protein expression in OC based on HPA database. D. GSVA analysis for OGN. E. GSEA analysis for OGN. F. The correlation analysis for OGN and FSHR in TCGA-OC dataset. G. The mRNA levels between vector and OGN overexpression of KGN and SKOV3 cells. H. The effect of OGN on FSHR tested by immunofluorescence.

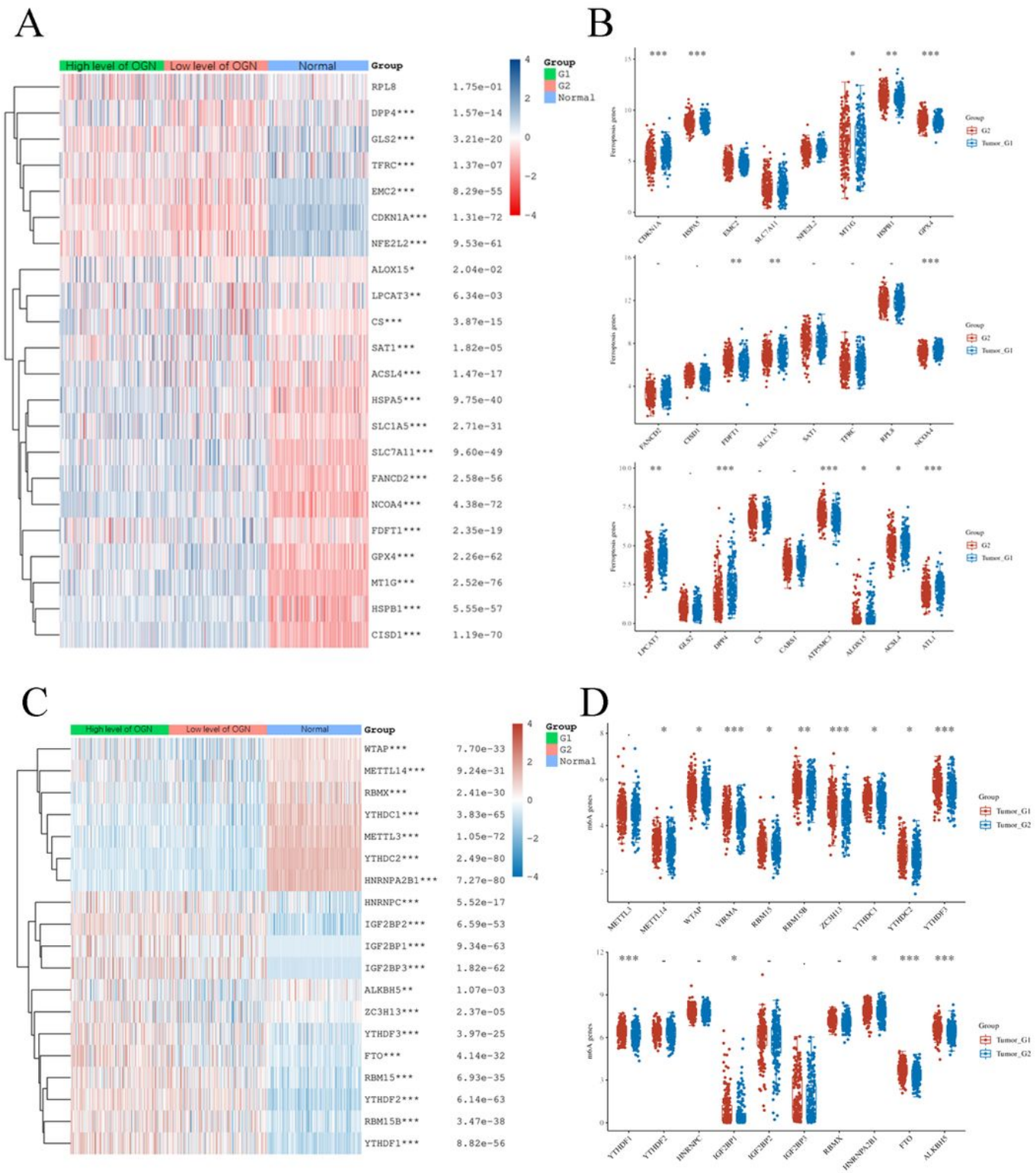

Figure 8

Association of OGN level with ferroptosis and m6A methylation related genes in OC. A. The ferroptosis genes expression in OC with high or low level OGN and normal ovary. B. The ferroptosis genes expression in OC with high or low level OGN. C. The m6A methylation genes expression in OC with high or low level OGN and normal ovary. D. The m6A methylation genes expression in OC with high or low level OGN.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- CTRP.png