Development of molecular subtype and prognostic model related to metabolism-related genes in high-grade serous ovarian cancer: A study based on ten cohorts

Tianren Li
The First Hospital of China Medical University
Dan Sun
The First Hospital of China Medical University
Hui Xue (cmuxuehui@163.com)
The First Hospital of China Medical University

Research Article

Keywords: HGSOC, Model, Subtype, Metabolism, Prognostic.

Posted Date: April 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1538028/v1

Abstract

Background

High-grade serous ovarian cancer (HGSOC) is a subtype of ovarian cancer with poor survival. At present, there are no good prognostic markers to guide treatment. Tumor metabolism plays a vital role in HGSOC; however, current knowledge on this topic is incomplete.

Method

We downloaded ten ovarian cancer cohorts from The Cancer Genome Atlas and Gene Expression Omnibus databases and identified 1525 platinum-treated HGSOC samples. Metabolic-related genes were identified in the Molecular Signatures Database, and consensus clustering was used to identify HGSOC metabolic subtypes. We used the least absolute shrinkage and the selection operator to construct a metabolic-related prognostic model. Finally, we analyzed the potential biological characteristics of high- and low-risk groups by calculating the degree of immune cell infiltration and enrichment analysis.

Results

We identified two metabolic subtypes (clusters 1 and 2) in the cohort of TCGA-OV. Cluster 2 was poor, primarily enriched in carbohydrate processes, and cluster 1 was enriched primarily in respiratory processes. We constructed a 16-metabolic gene-related model that successfully predicted the overall survival rate in TCGA-OV, GPL96-OV, GPL570-OV, and GPL6480-OV cohorts. The model was suitable for predicting progression-free survival. Patients at high risk were prone to infiltration with activated mast cells, while patients at low risk were prone to infiltration of T cell follicular helper cells. Pathway enrichment analysis showed significant DNA replication enrichment in the low-risk group.

Conclusion

We performed a comprehensive analysis of metabolic genes that will aid the study of the metabolic mechanisms of ovarian cancer. Identifying metabolic gene-related subtypes and generating a prognostic model will help predict the risk of HGSOC and identify appropriate treatments to prolong survival and improve quality of life in HGSOC patients.

Background

Ovarian cancer is generally diagnosed in the advanced stage and is the leading cause of death from gynecological cancer (1). The average risk of ovarian cancer in a woman's lifetime is 1.3%, equivalent to one in 78 women (2). There were 313959 new cases of ovarian cancer in 2020, and the number of deaths was 207252 (3). High-grade serous ovarian cancer (HGSOC) is the most common and malignant subtype

of ovarian cancer, causing 70–80% of deaths (4). There are two main reasons for the poor outcome. One is that it is discovered late in the course. Ovarian cancer exhibited a high heterogeneity on the molecular level, which led to difficulties in standard and accurate treatment. Cancer antigen 125 (CA125) is a specific diagnostic marker of ovarian cancer. Outcomes are predicted using CA125 (5) and the ovarian cancer risk algorithm (6); however, these markers have not been widely promoted because of their poor predictive ability. Now, more accurate prognostic markers for ovarian cancer are needed.

Many studies showed that metabolic changes in cancer are associated with the initiation and progression of cancer (7). Tumor metabolism is different from that of healthy tissues. Cancer cells distinguish themselves by participating in numerous aerobic glycolysis and pentose phosphate pathways and have a high degree of metabolic flexibility (8). Targeting metabolic genes is one of the most promising cancer therapies. Metabolic phenotype also provides information concerning outcomes and treatment. Classical metabolic pathways play essential roles in OV; these include PI3K/AKT/mTOR signaling (9). This participates significantly in OV tumorigenesis, chemotherapy, and radiotherapy resistance (10). Under aerobic conditions, ovarian cancer cells display high levels of the Warburg effect, which significantly accelerates disease progression and chemical resistance. Studies showed that PARP inhibitors could treat platinum-sensitive HGSOC (11). Members of the PARP family participate in metabolic reprogramming. In the Anna study, low serum phospholipids and essential amino acids were associated with poor outcomes in ovarian cancer (12). Research regarding OV metabolism is of great significance and might improve outcomes.

With the development of bioinformatics technology, much sample sequencing data has been generated. These data are critical to the work of mining for biomarkers. In the present study, ten ovarian cancer cohorts were downloaded to mine potential metabolic prognostic markers to construct molecular subtype and prognostic models related to OV metabolic genes.

Materials And Methods

Data download and collation

We downloaded the transcriptome and clinical data of the ovarian cancer cohort TCGA-OV from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) database. We also downloaded the transcriptome and clinical data from GSE53963 (13), GSE26712 (14), GSE30161 (15), GSE9891 (16), GSE17260 (17), GSE63885 (18), GSE32062 (19), GSE32063 (19), and GSE14764 (19) from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). We screened samples from the downloaded cohort, which needed to meet the following conditions simultaneously: HGSOC subtype in OV; contains both transcriptional data and survival data; patients are mainly treated with platinum chemotherapy. We used the R package "sva" to merge queues from the same platform and remove the batch effect. Finally, TCGA-OV (N = 363), GPL96-OV (N = 251), GPL570-OV (N = 353), and GPL6480-OV (N = 558) cohorts were obtained. For the data processing and analysis flow chart (Figure 1).

Genes related to metabolism

We found five sets of metabolism-related genes from "hallmark gene sets" in the Molecular Signatures Database (http://www.gsea-msigdb.org/gsea/msigdb/index.jsp): "HALLMARK_BILE_ACID_METABOLISM", "HALLMARK_CHOLESTEROL_HOMEOSTASIS", "HALLMARK_FATTY_ACID_METABOLISM", "HALLMARK_GLYCOLYSIS", "HALLMARK_HEME_METABOLISM", "HALLMARK_OXIDATIVE_PHOSPHORYLATION", and "HALLMARK_XENOBIOTIC_METABOLISM". A total of 855 genes were identified.

Consensus clustering

Consensus clustering determines the number and members of possible clusters in a data set. This method is widely used in cancer genomics research to discover molecular subtypes of new diseases. We use the R package "ConsensusClusterPlus" for this analysis.

Survival analysis

The Kaplan-Meier method is a univariate survival analysis using the log-rank test. The analysis was performed using the R package "survival" that divides patients into two groups by a factor to observe a difference in survival time. Receiver operating characteristic (ROC) curves reflect the continuous variables of sensitivity and specificity. We drew ROC curves using the R package "survivalROC" and calculated the area under the curve (AUC) values. The univariate and multivariate Cox analyses used in this study were based on R package "survival".

Differential gene expression

Using the R package "limma" package for difference analysis, we defined | logFC | > 2 and adjusted p-value > 0.05. We used the R package "ggplot2" to generate volcano plots.

Enrichment analysis

Metascape (http://metascape.org)) is an online enrichment analysis tool (20), integrated with a variety of enrichment methods and a high degree of recognition. Gene Set Enrichment Analysis (GSEA) is used to evaluate the distribution trend of genes in a predefined gene set in a gene table sorted by phenotypic correlation to calculate their contribution to the phenotype. We used the "c2.cp.kegg.v7.0. Symbols" gene set for pathway enrichment analysis (21).

Calculation of immune infiltration degree

ssGSEA is a type of GSEA used for enrichment analysis of a single sample. Here, we calculated the enrichment scores of immune cells and immune-related pathways based on the set of immune-related genes. CIBERSORT is a deep learning algorithm for deconvolution that quantifies 22 kinds of immune cells infiltration in the sample (22) using the existing immune cell and gene matrix.

Weighted gene co-expression network analysis (WGCNA)

WGCNA aims to identify co-expressed gene modules and explore the relationship between the module and the phenotype of concern and the core genes in the network. The first step is to calculate the correlation coefficient between any two genes (Pearson coefficient). The second step involves the correlation coefficient weighting method to build a scale-free network. The third step is to use the dynamic cut tree method to divide the genes into different modules. Each gene module has a similar expression pattern or function. In the fourth step, Pearson correlation analysis was carried out by combining gene module and clinical traits. The process is based on the R package "WGCNA" (23). We used Cytoscape to map the gene co-expression network (24).

Least absolute shrinkage and selection operator (LASSO)

LASSO is a method of data dimensionality reduction, which adds a regular term based on general linear regression, ensuring the best fitting error and making the parameters as simple as possible to strengthen the model's generalizability. We used LASSO to construct a prognostic model to achieve an optimal solution, which was based on the R package "glmnet" and "survival" (25).

Statistical analysis

The statistical analysis was based on R Programming Language software (Rx64 3.5.1). We used online tools to draw Venn diagrams (http://bioinformatics.psb.ugent.be/webtools/Venn/).

Results

Identification of molecular subtype related to HGSOC metabolism

The transcriptome data of 855 metabolic genes were extracted from TCGA-OV, and unsupervised clustering was used to construct molecular typing. When K = 2, the slope of CDF (Consistent cumulative distribution function) decreased slightly (Figure 2A-B), the sample size of each subgroup was more balanced, the intra-group correlation was high, and the inter-group correlation was low (Figure 2C). The other subtype when k at other value was showed at Supplementary figure 1. We divided the HGSOC samples of TCGA-OV into cluster 1 and cluster 2. Kaplan-Meier analysis showed that outcomes in cluster 2 were worse than cluster 1 ($p = 0.025$, hazard ratio = 1.395, Figure 2D). We verified the metabolic typing in GPL96-OV, GPL570-OV, and GPL6480-OV and found that these metabolic genes could still divide the samples into two groups, but there was no statistical significance in prognostic guidance. Only in the GPL96-OV and GPL570-OV cohorts, cluster 2 show a trend of poor prognosis (Supplementary figure 2).

Characteristics of two metabolic subtypes

To determine the metabolic characteristics of the two subtypes, we identified differentially-expressed metabolic genes of the two subtypes. We found that 141 genes were overexpressed in cluster 2 ($|logFC| > 1$, adjust p-value > 0.05, Figure 2E) and 160 genes were overexpressed in cluster 1 ($|logFC| > 1$, adjust p-value > 0.05). We placed these genes into Metascape for enrichment analysis and found that cluster 2 was enriched in carbohydrate processes (Figure 2F), while cluster 1 was enriched in respiratory processes

(Figure 2G). Then, we carried out ssGSEA analysis of immune cells and found higher immune infiltration in cluster 1 than in cluster 2 (Figure 2H). Finally, according to the subtype grouping, GSEA analysis showed that cluster 2 was enriched in "LYSINE_DEGRADATION", while cluster1 was enriched in "OXIDATIVE_PHOSPHORYLATION", "ALZHEIMERS_DISEASE", "PARKINSONS_DISEASE", and "HUNTINGTONS_DISEASE" (Figure 2I).

Construction of co-expression network related to metabolic subtypes

To build a scale-free network, it is necessary to select a suitable Soft threshold. When the Soft threshold = 3/4/5/6/8, R-square > 0.9 (Figure 3A). The mean connectivity is the largest among these values when the Soft threshold = 3 (Figure 3B). Therefore, we chose the Soft threshold = 3 to build a scale-free network. Then, using the method of the dynamic cutting tree, the genes were divided into five modules (Figure 3C). We took the futime, fustat, age, grade, stage, and metabolic subtype of the sample as character data and analyzed the correlations with the module. The correlation between the turquoise module and the cluster was the highest (Figure 3D, cor = -0.47, p = 8E-21). We drew a scatter plot of the turquoise module's gene significance and module membership (Figure 3E). Finally, we used Cytoscape to construct a co-expression network (Figure 3F) for the gene significance > 0.3 and module membership > 0.3. Finally, we identified five hub genes (NDUFA2, COX8A, COX5B, UQCR11, UQCRQ, and ATP5MF) with the highest degree of connection in this network.

Construction of the metabolic-related prognostic model

Univariate Cox analysis of 564 genes in the turquoise module showed that 48 genes were statistically significant (Figure 4A, p < 0.05). The 855 metabolic genes in the GSEA database were overlapped with the entire genomes of GPL96-OV, GPL570-OV, and GPL6480-OV. Then GPL96-OV, GPL570-OV, and GPL6480-OV obtained 783, 832, and 846 metabolic genes. To enable our model to be successfully verified by the data of other cohorts, we intersected these 48 genes with the metabolic genes of GPL96-OV, GPL570-OV, and GPL6480-OV (40 genes) (Figure 4B). To optimize our model, we used LASSO for dimensionality reduction analysis. With the increase of lambda value, the coefficients of some genes gradually become 0 (Figure 4C), indicating that these genes have little effect on the model and should be abandoned. Then the 10X cross-validation method was used to calculate the partial likelihood deviance of the model. When the gene number was 16, the model's performance was the best (Figure 4D). Figure 4E shows the 16 genes and their corresponding coefficients (Table 1). We named the model 16-MGM (16 metabolic gene-related model). The risk score of each patient was calculated using 16-MGM.

Verifying the prognostic performance of 16-MGM in training set TCGA-OV

Kaplan-Meier analysis of TCGA-OV showed that 16-MGM had a significant prognostic ability (Figure 4F, p < 0.05), and the prognosis was poor when the score was high. Figure 4G showed the score distribution and survival status distribution of each patient. ROC curve showed that 16-MGM had a certain prognostic ability after nine years (AUC > 0.65). Univariate and multivariate Cox analysis of risk score, age, stage, and grade showed that the model's risk score was independent of age, stage, and grade for predictive

analysis (Figure 4I). It has prognostic ability in many clinical subgroups, including age < = 60, age > 60, grade 3, stage 3, and stage 4 (Figure 4J).

Verify the prognostic performance of 16-MGM in the verification sets

We verified the prognostic ability of 16-MGM in verification sets GPL96-OV (N = 251) and GPL570-OV (N = 353), and GPL6480-OV (N = 558) cohorts. Kaplan-Meier analysis showed that, in the GPL96-OV cohort, the risk score predicted the overall survival rate (Figure 5A, p = 0.43), and the AUC values of 3 and 5 years were greater than 0.5. In the GPL570-OV cohort, the risk score predicted the overall survival rate (Figure 5B, p = 0.019), and the AUC values of 3 and 5 years were greater than 0.5. In the GPL6480-OV cohort, the risk score predicted the overall survival rate (Figure 5C, p < 0.001), and the AUC value of 3 and 5 years was greater than 0.5. 16-MGM also predicted progression-free survival (PFS), which was statistically significant in the cohort TCGA-OV and GPL570-OV cohorts (Figure 5D-E, p < 0.05), but not in the GPL6480-OV cohort (Figure 5D-E).

Difference of immune infiltration level between high and low-risk score samples

We used the CIBERSORT algorithm to calculate the infiltration percentage of 22 immune cells per sample in the TCGA-OV, GPL96-OV, GPL570-OV, and GPL6480-OV cohorts. There was no significant difference in infiltration between patients with high- and low-risk scores (Figure 6A). However, there were significant differences in the infiltration of some immune cells. We selected the cells with different immune infiltration in each cohort (Figure 6B) and found that there was significant infiltration of "mast cell activated" in the high-risk group of the GPL96, GPL70, and GPL6480 cohorts and "T cell follicular helper" in the low-risk group of the GPL96, GPL70, and TCGA-OV cohorts (p < 0.05).

GSEA analysis of samples with high and low-risk scores

We performed GSEA analysis of the KEGG pathway in TCGA-OV, GPL96-OV, GPL570-OV, and GPL6480-OV. The high-risk group did not enrich the pathway, while the low-risk group enriched several pathways (Figure 7A-D), including metabolic-related pathways. We intersected the results (Figure 7E) and found that "DNA replication" was enriched in all four cohorts.

Discussion

Ovarian cancer is usually diagnosed in an advanced stage and is the most common cause of cancer death in women. HGSOC is a more malignant clinical subtype of ovarian cancer characterized by late discovery and poor outcome. Surgical cell reduction to R0 is the pillar of the treatment of HGSOC, followed by platinum chemotherapy (26). The selection of appropriate treatment is critical to improving outcomes. Tumor metabolism plays a vital role in the occurrence, development, and treatment of ovarian cancer. In this study, we downloaded ten ovarian cancer cohorts from TCGA and GEO databases, performed a comprehensive analysis of metabolism-related genes, and constructed a metabolism-related molecular typing and prognosis model of ovarian cancer (16-MGM). This model might facilitate a

broader understanding of the metabolic mechanisms of ovarian cancer and the realization of proper treatment.

We identified molecular typing related to HGSOC metabolism. We used consistent clustering in the TCGA-OV cohort to classify the patients into two types (cluster 1 and cluster 2). Cluster 2 was characterized by enrichment in the carbohydrate metabolism process. Cluster 1 was enriched in the respiratory process, and the outcome of cluster 2 was worse than that of cluster 1. Cells produce ATP via glycolysis and oxidative phosphorylation. Normal cells primarily use oxidative phosphorylation to obtain energy, while tumor cells primarily use aerobic glycolysis. Our findings suggest that when glycolysis is strengthened, outcomes are worse. We found six network center genes (NDUFA2, COX8A, COX5B, UQCR11, UQCRQ, and ATP5MF) related to metabolic typing using a gene-weighted co-expression network. A study showed that NDUFA2 is significantly overexpressed in head and neck paragangliomas and is located in the cytoplasm of PGL-626 cells (27). COX5B is a prognostic marker for renal clear cell carcinoma (28), glioma (29), breast cancer 28592145, and other cancers. UQCR11 was significantly overexpressed in lung adenocarcinoma (30). UQCRQ plays an essential role in the oxidative metabolism of cancer (31). These six genes may play an essential role in the metabolism of ovarian cancer.

We constructed a metabolism-related prognostic model (16-MGM) for HGSOC and successfully predicted the overall survival rate in four cohorts, with a sample size of 1525. We used the LASSO algorithm in TCGA-OV to construct the 16-MGM model. By calculating risk scores using this model, we divided patients into high- and low-risk groups. Outcomes in the high-risk group were significantly worse than those in the low-risk group. The model had strong predictive ability for 1, 3, 5, 7, and 9 years and predicted outcomes independently of other clinical features. We validated 16-MGM in the GPL96/GPL570/GPL6480 OV cohorts and found that the model robustly predicted outcomes. The model also successfully predicted disease-free survival of HGSOC in the GPL570/GPL6480/TCGA OV cohort. Many studies constructed the prognostic model of HGSOC; however, its accurate performance was low (32) and has not been verified by other independent cohorts (33); therefore, it cannot be used clinically. The sample size of our study was 1525, suggesting that our model has high robustness and credibility. The verification samples of the model came from various regions around the world, suggesting that 16-MGM has a wide range of applicability and possesses clinical significance.

Patients with a high risk of HGSOC metabolism showed high infiltration in activated mast cells and low infiltration in follicular helper T cells. Histamine is synthesized and released by mast cells. High levels of histamine are associated with the biological characteristics of ovarian cancer (34). T cell follicular helper cells help recruit CD8 + T cells, NK cells, and macrophages to participate in anti-tumor immunity and support B cells' anti-tumor antibody responses (35). Activated mast cells may act as accomplices of ovarian cancer cells, while follicular helper T cells are the opposite. The GSEA analysis found that the "DNA replication" signal pathway was significantly enriched in the low-risk group. This suggests that the "DNA replication" signal pathway may play an essential role in inhibiting the malignant progression of ovarian cancer.

Our research also has some limitations. First, our study does not include transcriptional data from normal ovaries, therefore, we are unsure whether these are cancer-specific genes among the 16 metabolism-related genes used to build the model. Second, our metabolic subtype is only successfully verified in TCGA-OV, and it does not perform well in other cohorts, and more clinical studies are needed to verify it in the future.

Conclusions

We have made a comprehensive bioinformatic analysis of the metabolism-related genomics of ovarian cancer, which provides more insights for exploring the metabolism-related mechanisms of ovarian cancer. The metabolism-related prognosis model we constructed, 16MGM, successfully predicted 1525 HGSOC samples, is a robust and powerful prognostic tool for HGSOC. In the future, more samples and in-depth experiments are needed to explore the applicability and effectiveness of 16-MGM.

Abbreviations

HGSOC
High-grade serous ovarian cancer
TCGA
The Cancer Genome Atlas
GEO
Gene Expression Omnibus
AUC
Area Under the Curve
ROC
Receiver Operating Characteristic
LASSO
Least absolute shrinkage and selection operator
coef
Regression coefficient
GSEA
Gene set enrichment analysis
PFS
Progression-free survival
OS
Overall survival.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

You can obtain the TCGA-OV cohort from The Cancer Genome Atlas database. And obtain GSE53963, GSE26712, GSE30161, GSE9891, GSE17260, GSE63885, GSE32062, GSE32063, and GSE14764 cohorts from the Gene Expression Omnibus.

Competing interests

The authors declare that they have no competing interests

Funding

Not applicable

Authors' contributions

Tianren Li conducted data analysis and written manuscripts; Dan Sun assisted in drawing pictures; Hui Xue provided ideas and designed experiments.

Acknowledgements

We are grateful for the free use of TCGA and GEO databases.

References

1. Lheureux S, Gourley C, Vergote I, Oza Am. Epithelial ovarian cancer. Lancet (London, England) (2019) 393(10177):1240–53. doi: 10.1016/s0140-6736(18)32552-2. PubMed PMID: 30910306.
2. Torre La, Trabert B, Desantis Ce, Miller Kd, Samimi G, Runowicz Cd, et al. Ovarian cancer statistics, 2018. CA Cancer J Clin (2018) 68(4):284–96. doi: 10.3322/caac.21456. PubMed PMID: 29809280.
3. Sung H, Ferlay J, Siegel Rl, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin (2021) 71(3):209–49. doi: 10.3322/caac.21660. PubMed PMID: 33538338.
4. Bowtell Dd, Böhm S, Ahmed Aa, Aspuria Pj, Bast Rc, Beral V, et al. Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. Nat Rev Cancer (2015) 15(11):668–79. doi: 10.1038/nrc4019. PubMed PMID: 26493647.
5. Karam Ak, Karlan By. Ovarian cancer: the duplicity of CA125 measurement. *Nat Rev Clin Oncol* (2010) 7(6):335-9. doi: 10.1038/nrclinonc.2010.44. PubMed PMID: 20368726.

6. Karlsen Ma, Sandhu N, Høgdall C, Christensen Ij, Nedergaard L, Lundvall L, et al. Evaluation of HE4, CA125, risk of ovarian malignancy algorithm (ROMA) and risk of malignancy index (RMI) as diagnostic tools of epithelial ovarian cancer in patients with a pelvic mass. Gynecol Oncol (2012) 127(2):379–83. doi: 10.1016/j.ygyno.2012.07.106. PubMed PMID: 22835718.
7. Vander Heiden Mg, Deberardinis Rj. Understanding the Intersections between Metabolism and Cancer Biology. *Cell* (2017) 168(4):657 – 69. doi: 10.1016/j.cell.2016.12.039. PubMed PMID: 28187287.
8. Kreuzaler P, Panina Y, Segal J, Yuneva M. Adapt and conquer: Metabolic flexibility in cancer growth, invasion and evasion. Molecular metabolism (2020) 33:83–101. doi: 10.1016/j.molmet.2019.08.021. PubMed PMID: 31668988.
9. Ediriweera Mk, Tennekoon Kh, Samarakoon Sr. Role of the PI3K/AKT/mTOR signaling pathway in ovarian cancer: Biological and therapeutic significance. *Semin Cancer Biol* (2019) 59:147 – 60. doi: 10.1016/j.semcancer.2019.05.012. PubMed PMID: 31128298.
10. Mabuchi S, Kuroda H, Takahashi R, Sasano T. The PI3K/AKT/mTOR pathway as a therapeutic target in ovarian cancer. Gynecol Oncol (2015) 137(1):173–9. doi: 10.1016/j.ygyno.2015.02.003. PubMed PMID: 25677064.
11. Scott Cl, Swisher Em, Kaufmann Sh. Poly (ADP-ribose) polymerase inhibitors: recent advances and future development. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* (2015) 33(12):1397 – 406. doi: 10.1200/jco.2014.58.8848. PubMed PMID: 25779564.
12. Bachmayr-Heyda A, Aust S, Auer K, Meier Sm, Schmetterer Kg, Dekan S, et al. Integrative Systemic and Local Metabolomics with Impact on Survival in High-Grade Serous Ovarian Cancer. Clinical cancer research: an official journal of the American Association for Cancer Research (2017) 23(8):2081–92. doi: 10.1158/1078-0432.ccr-16-1647. PubMed PMID: 27797973.
13. Konecny Ge, Wang C, Hamidi H, Winterhoff B, Kalli Kr, Dering J, et al. Prognostic and therapeutic relevance of molecular subtypes in high-grade serous ovarian cancer. J Natl Cancer Inst (2014) 106(10). doi: 10.1093/jnci/dju249. PubMed PMID: 25269487.
14. Bonome T, Levine Da, Shih J, Randonovich M, Pise-Masison Ca, Bogomolniy F, et al. A gene signature predicting for survival in suboptimally debulked patients with ovarian cancer. Cancer Res (2008) 68(13):5478–86. doi: 10.1158/0008-5472.can-07-6595. PubMed PMID: 18593951.
15. Dugalić D, Colović R, Savić M. [Duodenal obstruction caused by gallstones (Bouveret syndrome)]. Acta Chir Iugosl (1990) 37(1):75–82. PubMed PMID: 2248014.
16. Canossa M, Twiss Jl, Verity An, Shooter Em. p75(NGFR) and TrkA receptors collaborate to rapidly activate a p75(NGFR)-associated protein kinase. The EMBO journal (1996) 15(13):3369–76. PubMed PMID: 8698038.
17. Yoshihara K, Tajima A, Yahata T, Kodama S, Fujiwara H, Suzuki M, et al. Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets. PLoS ONE (2010) 5(3):e9615. doi: 10.1371/journal.pone.0009615. PubMed PMID: 20300634.
18. Lisowska Km, Olbryt M, Dudaladava V, Pamuła-Piłat J, Kujawa K, Grzybowska E, et al. Gene expression analysis in ovarian cancer - faults and hints from DNA microarray study. Frontiers in

oncology (2014) 4:6. doi: 10.3389/fonc.2014.00006. PubMed PMID: 24478986.

19. Yoshihara K, Tsunoda T, Shigemizu D, Fujiwara H, Hatae M, Fujiwara H, et al. High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by downregulation of antigen presentation pathway. Clinical cancer research: an official journal of the American Association for Cancer Research (2012) 18(5):1374–85. doi: 10.1158/1078-0432.ccr-11-2725. PubMed PMID: 22241791.
20. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi Ah, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6. PubMed PMID: 30944313.
21. Subramanian A, Tamayo P, Mootha Vk, Mukherjee S, Ebert Bl, Gillette Ma, 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(43):15545–50. doi: 10.1073/pnas.0506580102. PubMed PMID: 16199517.
22. Newman Am, Liu Cl, Green Mr, Gentles Aj, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Meth (2015) 12(5):453–7. doi: 10.1038/nmeth.3337. PubMed PMID: 25822800.
23. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics (2008) 9:559. doi: 10.1186/1471-2105-9-559. PubMed PMID: 19114008.
24. Shannon P, Markiel A, Ozier O, Baliga Ns, Wang Jt, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res (2003) 13(11):2498–504. doi: 10.1101/gr.1239303. PubMed PMID: 14597658.
25. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of statistical software (2010) 33(1):1–22. PubMed PMID: 20808728.
26. Kuroki L, Guntupalli Sr. Treatment of epithelial ovarian cancer. *BMJ* (2020) 371:m3773. doi: 10.1136/bmj.m3773. PubMed PMID: 33168565.
27. Wang Z, Chen H, Xue L, He W, Shu W, Wu H, et al. High throughput proteomic and metabolic profiling identified target correction of metabolic abnormalities as a novel therapeutic approach in head and neck paraganglioma. Transl Oncol (2021) 14(8):101146. doi: 10.1016/j.tranon.2021.101146. PubMed PMID: 34118692.
28. Stein J, Tenbrock J, Kristiansen G, Müller Sc, Ellinger J. Systematic expression analysis of the mitochondrial respiratory chain protein subunits identifies COX5B as a prognostic marker in clear cell renal cell carcinoma. International journal of urology: official journal of the Japanese Urological Association (2019) 26(9):910–6. doi: 10.1111/iju.14040. PubMed PMID: 31280487.
29. Hu T, Xi J. Identification of COX5B as a novel biomarker in high-grade glioma patients. *OncoTargets and therapy* (2017) 10:5463-70. doi: 10.2147/ott.s139243. PubMed PMID: 29180880.
30. Xu H, Ma J, Wu J, Chen L, Sun F, Qu C, et al. Gene expression profiling analysis of lung adenocarcinoma. Brazilian journal of medical and biological research = Revista brasileira de

pesquisas medicas e biologicas (2016) 49(3). doi: 10.1590/1414-431x20154861. PubMed PMID: 26840709.

31. Perry Mc, Dufour Cr, Tam Is, B'chir W, Giguère V. Estrogen-related receptor-α coordinates transcriptional programs essential for exercise tolerance and muscle fitness. Molecular endocrinology (Baltimore, Md) (2014) 28(12):2060–71. doi: 10.1210/me.2014-1281. PubMed PMID: 25361393.
32. Zhang Y, Liu J, Raj-Kumar Pk, Sturtz La, Praveen-Kumar A, Yang Hh, et al. Development and validation of prognostic gene signature for basal-like breast cancer and high-grade serous ovarian cancer. *Breast Cancer Res Treat* (2020) 184(3):689 – 98. doi: 10.1007/s10549-020-05884-z. PubMed PMID: 32880016.
33. Wu C, He L, Wei Q, Li Q, Jiang L, Zhao L, et al. Bioinformatic profiling identifies a platinum-resistant-related risk signature for ovarian cancer. Cancer medicine (2020) 9(3):1242–53. doi: 10.1002/cam4.2692. PubMed PMID: 31856408.
34. Faustino-Rocha Ai, Ferreira R, Gama A, Oliveira Pa, Ginja M. Antihistamines as promising drugs in cancer therapy. Life Sci (2017) 172:27–41. doi: 10.1016/j.lfs.2016.12.008. PubMed PMID: 27986539.
35. Crotty S. T Follicular Helper Cell Biology: A Decade of Discovery and Diseases. Immunity (2019) 50(5):1132–48. doi: 10.1016/j.immuni.2019.04.011. PubMed PMID: 31117010.

Tables

Table 1. The basic information of genes in the model.

Gene	Coef	Protein-name	Subcellular location
EZH1	0.006000718	Histone-lysine N-methyltransferase	nucleus
ABCA1	0.00537628	Phospholipid-transporting ATPase	plasma membrane/endosome
PYGB	0.005146045	Glycogen phosphorylase	_
MVK	0.004671453	Mevalonate kinase	cytosol
CITED2	0.002112938	Cbp/p300-interacting transactivator 2	nucleus
MRPS12	0.002038517	28S ribosomal protein S12	mitochondrion
ALDOC	0.000541047	Fructose-bisphosphate aldolase C	cytosol
UQCRFS1	9.46E-06	Cytochrome b-c1 complex subunit Rieske	mitochondrion
TSTA3	-3.69E-05	GDP-L-fucose synthase	cytosol
CXCR4	-0.00115564	C-X-C chemokine receptor type 4	plasma membrane/extracellular/endosome/lysosome
GSTK1	-0.001912715	Glutathione S-transferase kappa 1	peroxisome
NDUFV2	-0.005135994	NADH dehydrogenase flavoprotein 2	mitochondrion
TCEA1	-0.006150921	Transcription elongation factor A protein 1	nucleus
GLRX5	-0.006483673	Glutaredoxin-related protein 5	mitochondrion
MRPS11	-0.023664211	28S ribosomal protein S11	mitochondrion
FBXO9	-0.036556168	F-box only protein 9	_

Figures

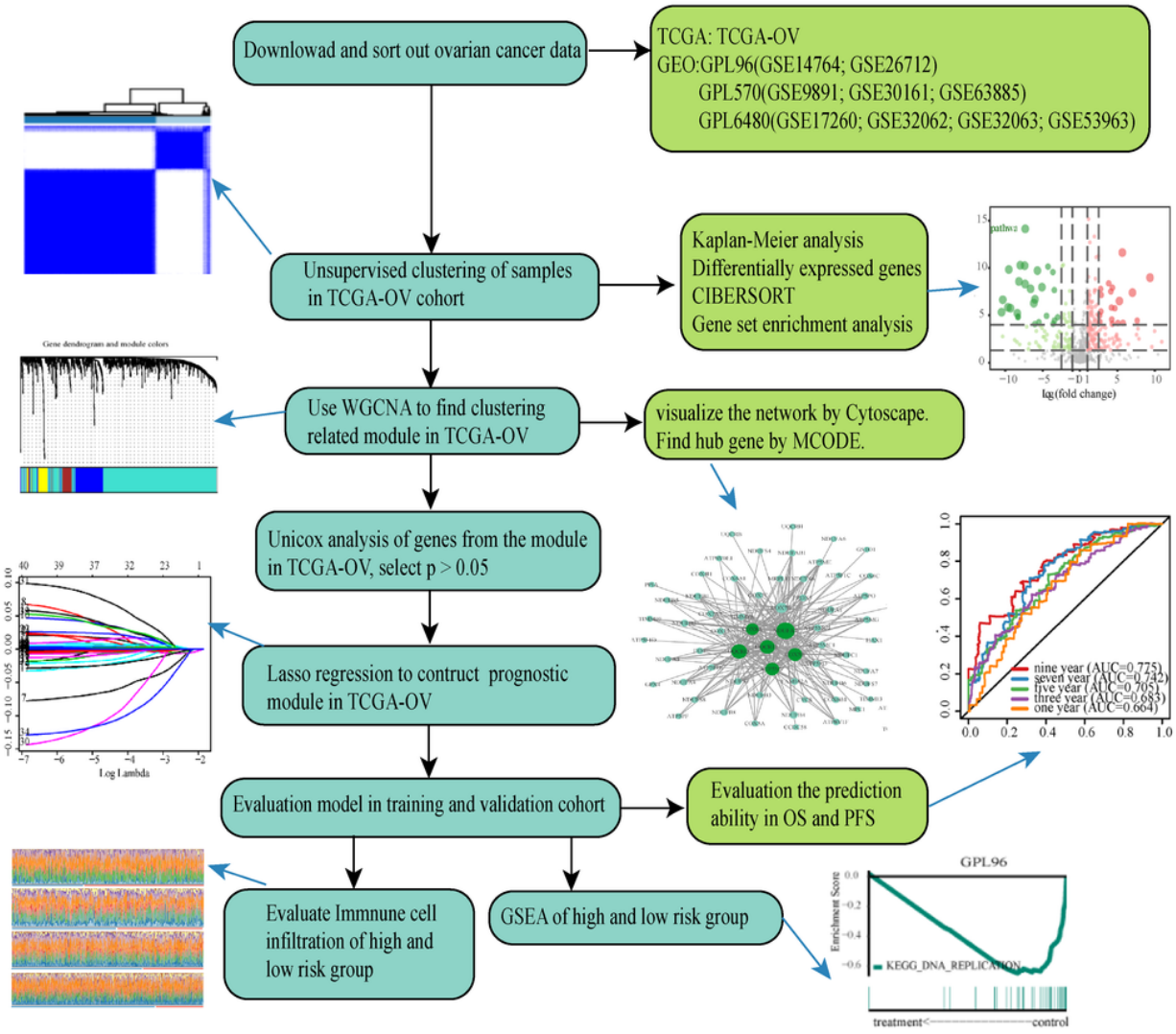

Figure 1

Data analysis and research process.

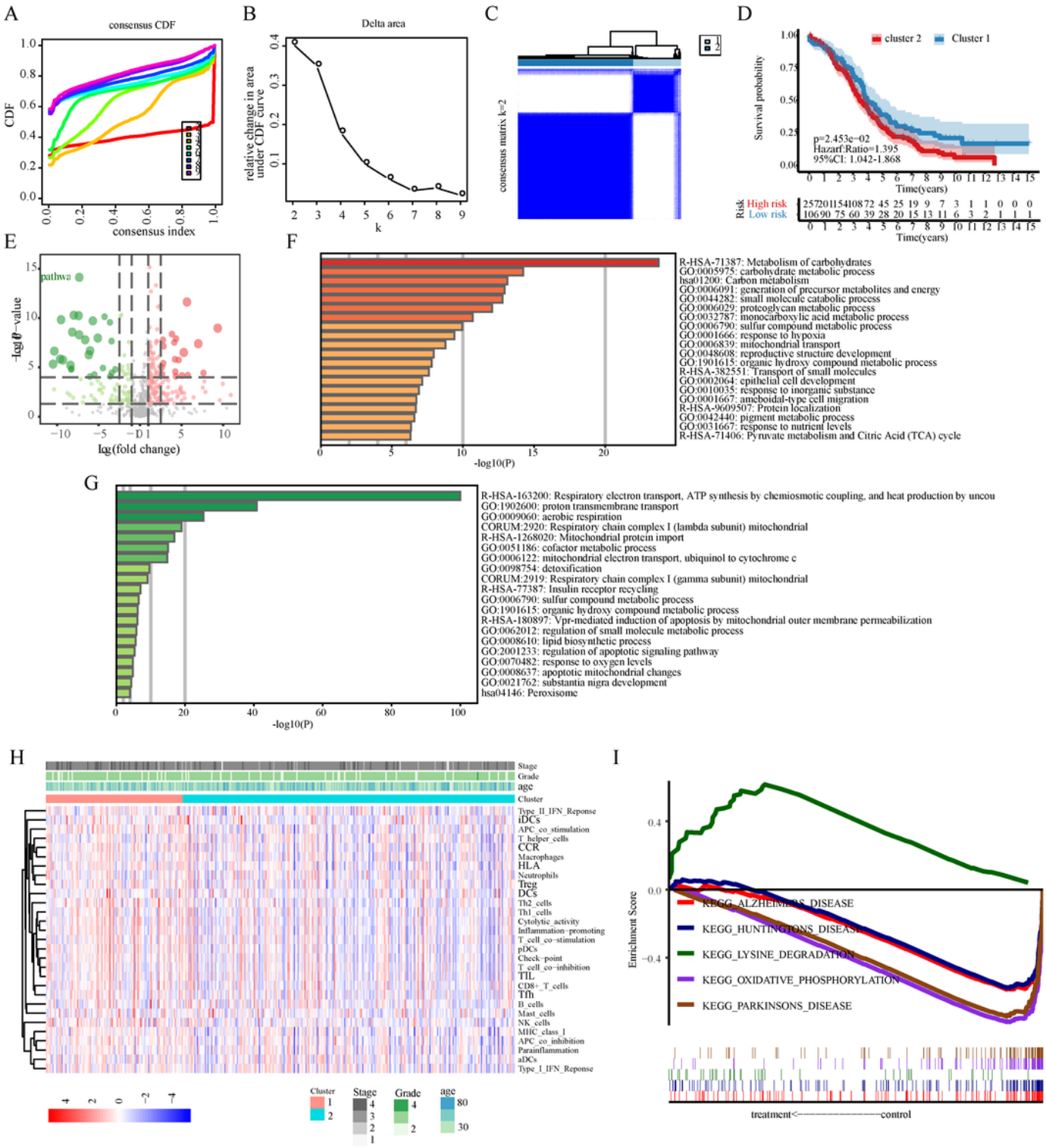

Figure 2

Molecular typing related to metabolism. (A) The trend chart of different k value curves in CDF and consensus index. (B) The relationship between the area under the CDF curve and the k value. (C) When k = 2, the cluster diagram of TCGA-OV. (D) Kaplan-Meier analysis of Cluster 1 and cluster 2. (E) Volcano map of clusters 1 and 2 differentially-expressed genes. Red represents overexpression in cluster 2, and the green represents overexpression in cluster 1. (F) Enrichment analysis of overexpressed genes in

Cluster 2. (G) Enrichment analysis of overexpressed genes in Cluster 1. (H) The infiltration heat map of several immune cells. (I) Enrichment line chart of GSEA.

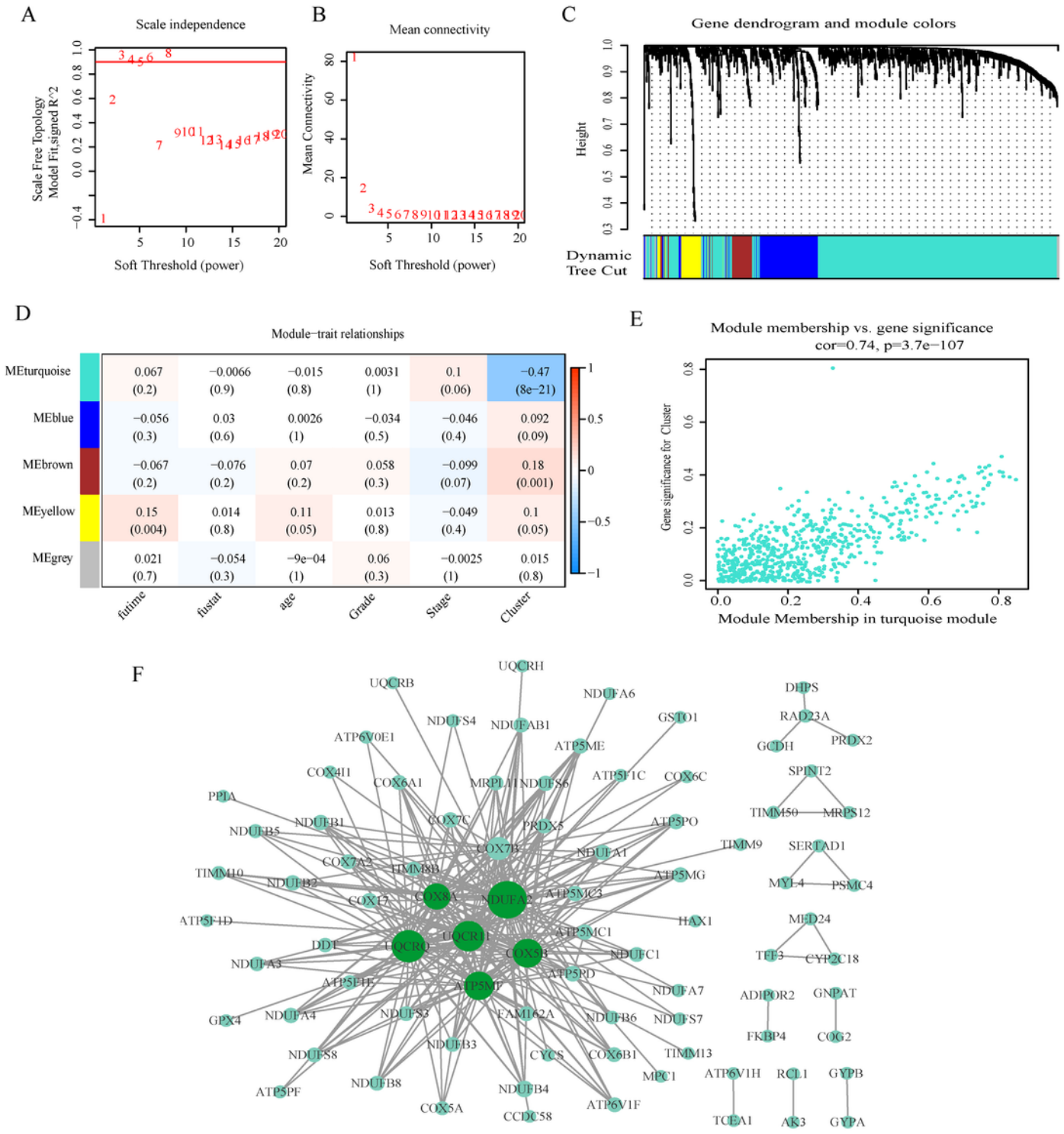

Figure 3

Gene weighted co-expression network analysis. (A) R-square values under different soft thresholds. (B) Average connectivity under different Soft thresholds. (C) Gene clustering tree and gene module. (D)

Correlation analysis between modules and clinical characters. (E) Scatter diagrams of module membership and gene significance for each gene in the turquoise module. (F) The co-expression network of turquoise module genes.

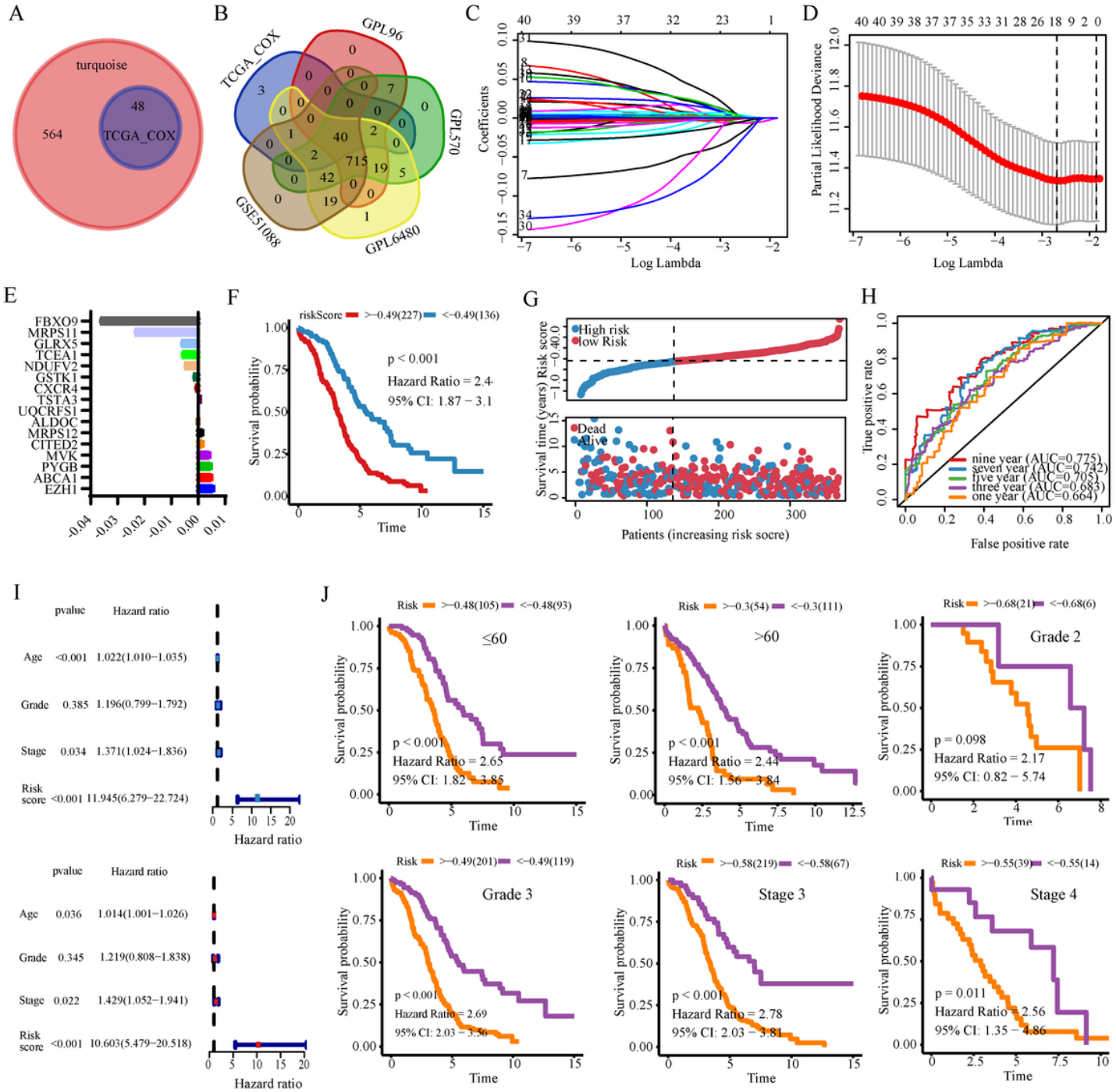

Figure 4

The construction of the prognostic model. (A) The results of univariate Cox analysis. (B) Genes existing in several cohorts. (C) The coefficient of 40 genes and lambda curve. (D) The distribution map of the genes in deviance and lambda. (E) The coefficient corresponding to genes in 16-MGM. (F) Kaplan-Meier

analysis of 16-MGM. (G) Distribution map of risk score and survival status of patients. (H) 16-MGM's ROC in 1, 3, 5, 7, and 9 years. (I) Univariate and multivariate Cox analysis of 16-MGM. (J) Kaplan-Meier analysis of 16-MGM in different clinical subgroups.

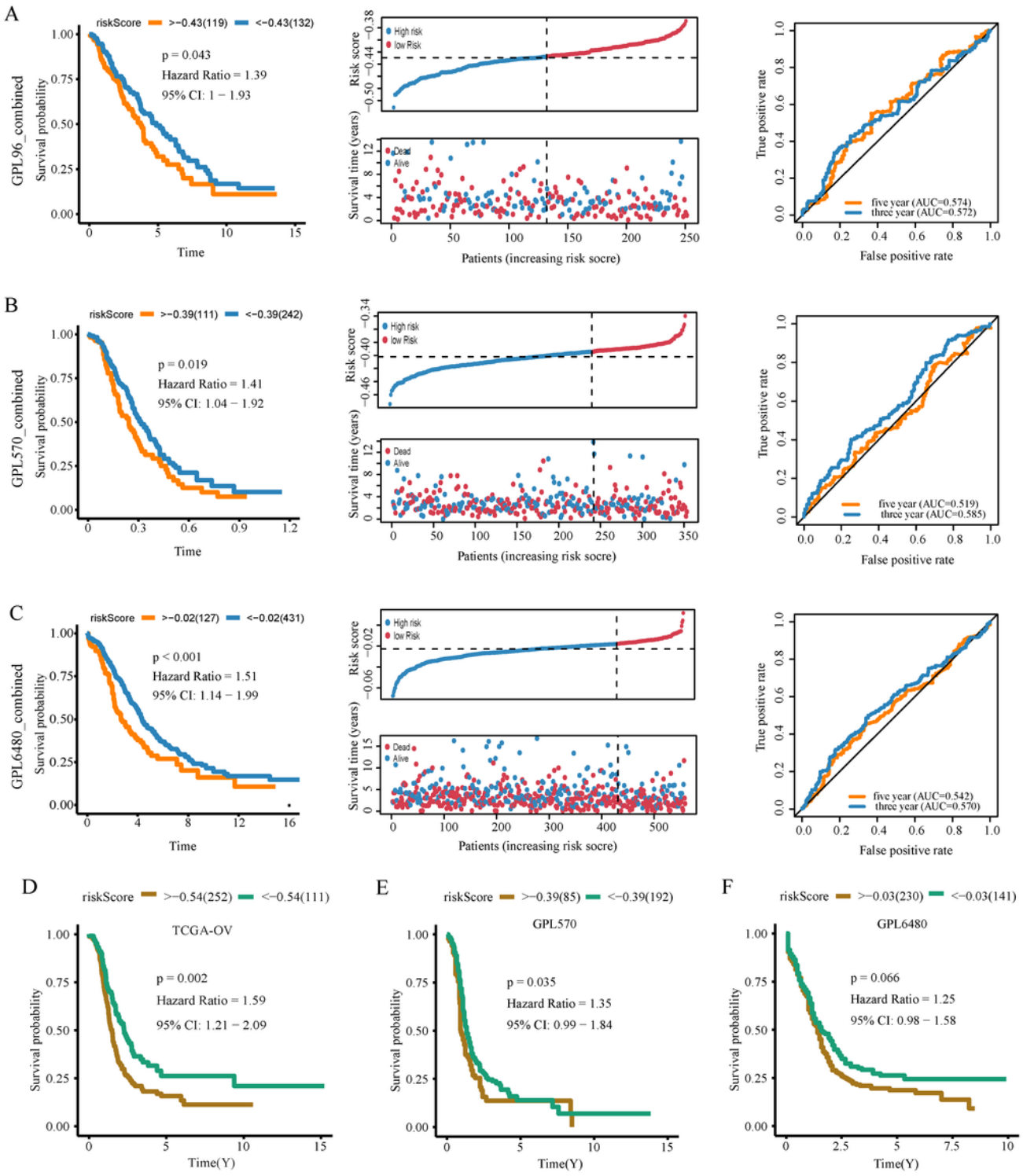

Figure 5

Verification of prognostic model. (A) In GPL96-OV cohort, Kaplan-Meier analysis of 16-MGM predictor OS; risk score and survival status distribution map of patients; 3- and 5-year ROC. (B) GPL570-OV cohort. (C) GPL6480-OV cohort. (D) In the TCGA-OV cohort, the Kaplan-Meier analysis of 16-MGM prediction PFS. (E) GPL570-OV cohort. (F) GPL6480-OV cohort.

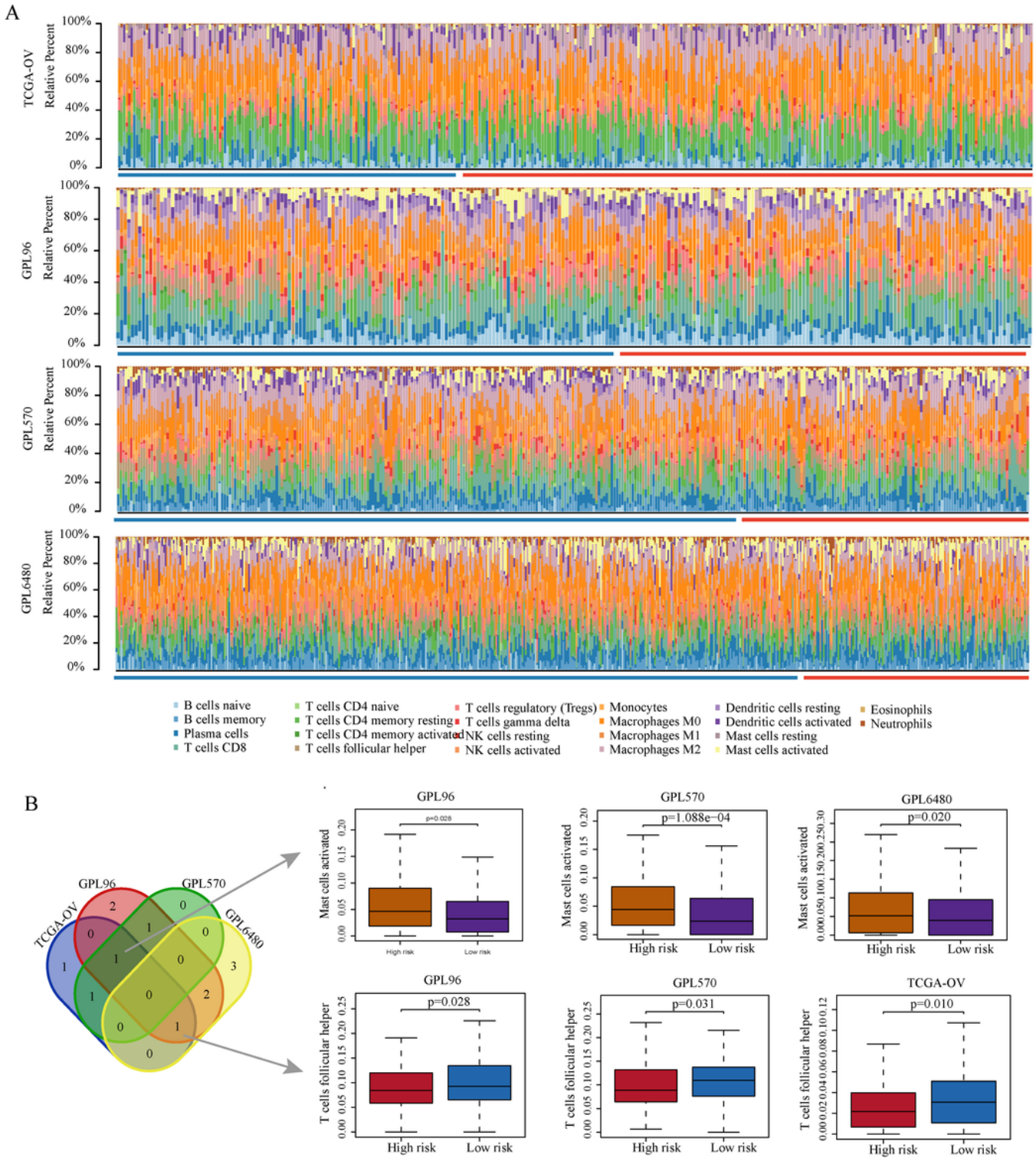

Figure 6

Differential infiltration of immune cells in high and low-risk groups. (A) The infiltration map of 22 immune cells in TCGA-OV, GPL96-OV, GPL570-OV, and GPL6480-OV cohorts. (B) The immune cells with different infiltration in the four cohorts were collected and intersected.

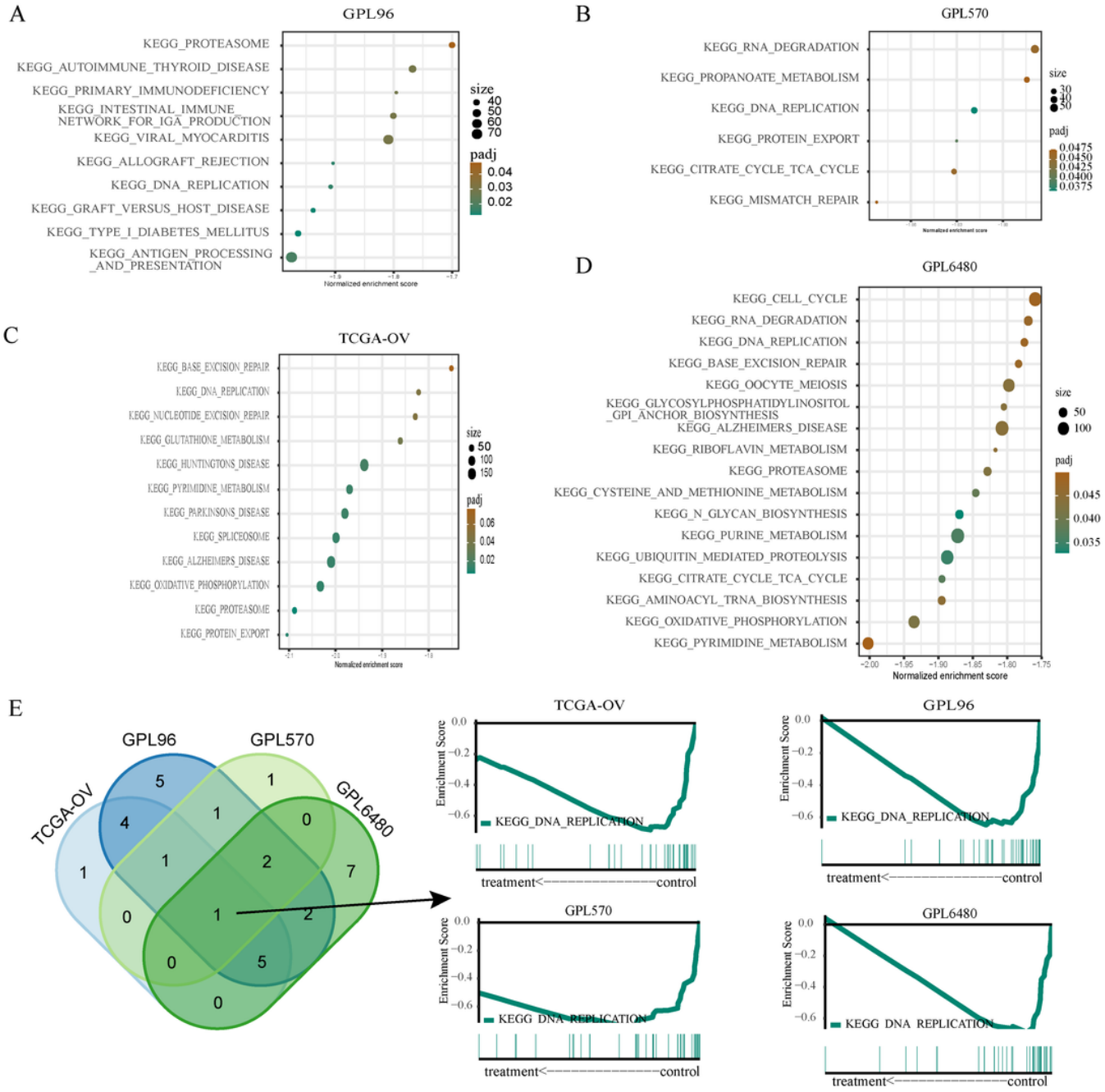

Figure 7

Enrichment analysis of GSEA pathway in high- and low-risk groups. (A) TCGA-OV cohort's GSEA analysis bubble chart. (B) GPL96-OV cohort. (C) GPL570-OV cohort. (D) GPL6480-OV cohort. (E) The GSEA results of the four cohorts are intersected.

Supplementary Files

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

- Supplementaryfigure1.tif
- Supplementaryfigure2.tif