



## OPEN ACCESS

## EDITED BY

Vinay Kumar,  
The Pennsylvania State University,  
United States

## REVIEWED BY

Hina Sultana,  
University of North Carolina System,  
United States  
Simna Saraswathi Prasannakumari,  
University of North Carolina at Chapel Hill,  
United States

## \*CORRESPONDENCE

Lianfang Zhao,  
✉ zhaolianfang0913@163.com

<sup>†</sup>These authors have contributed equally  
to this work

RECEIVED 17 November 2024

ACCEPTED 20 January 2025

PUBLISHED 03 February 2025

## CITATION

Chen J, Li Q, Liu X, Lin F, Jing Y, Yang J and  
Zhao L (2025) Potential biomarkers and  
immune infiltration linking endometriosis with  
recurrent pregnancy loss based on  
bioinformatics and machine learning.  
*Front. Mol. Biosci.* 12:1529507.  
doi: 10.3389/fmolb.2025.1529507

## COPYRIGHT

© 2025 Chen, Li, Liu, Lin, Jing, Yang and Zhao.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Potential biomarkers and immune infiltration linking endometriosis with recurrent pregnancy loss based on bioinformatics and machine learning

Jianhui Chen<sup>1†</sup>, Qun Li<sup>2†</sup>, Xiaofang Liu<sup>1</sup>, Fang Lin<sup>1</sup>, Yaling Jing<sup>1</sup>,  
Jiayan Yang<sup>1</sup> and Lianfang Zhao<sup>1\*</sup>

<sup>1</sup>Prenatal Diagnosis Center, Center of Reproductive Medicine, Suining Central Hospital, Suining, Sichuan, China, <sup>2</sup>Department of Radiology, Suining Central Hospital, Suining, Sichuan, China

**Objective:** Endometriosis (EMs) is a chronic inflammatory disease characterized by the presence of endometrial tissue in the non-uterine cavity, resulting in dysmenorrhea, pelvic pain, and infertility. Epidemiologic data have suggested the correlation between EMs and recurrent pregnancy loss (RPL), but the pathological mechanism is unclear. This study aims to investigate the potential biomarkers and immune infiltration in EMs and RPL, providing a basis for early detection and treatment of the two diseases.

**Methods:** Two RPL and six EMs transcriptomic datasets from the Gene Expression Omnibus (GEO) database were used for differential analysis via limma package, followed by weighted gene co-expression network analysis (WGCNA) for key modules screening. Protein-protein interaction (PPI) network and two machine learning algorithms were applied to identify the common core genes in both diseases. The diagnostic capabilities of the core genes were assessed by receiver operating characteristic (ROC) curves. Moreover, immune cell infiltration was estimated using CIBERSORTx, and the Cancer Genome Atlas (TCGA) database was employed to elucidate the role of key genes in endometrial carcinoma (EC).

**Results:** 26 common differentially expressed genes (DEGs) were screened in both diseases, three of which were identified as common core genes (MAN2A1, PAPSS1, RIBC2) through the combination of WGCNA, PPI network, and machine learning-based feature selection. The area under the curve (AUC) values generated by the ROC indicates excellent diagnostic powers in both EMs and RPL. The key genes were found to be significantly associated with the infiltration of several immune cells. Interestingly, MAN2A1 and RIBC2 may play a predominant role in the development and prognostic stratification of EC.

**Conclusion:** We identified three key genes linking EMs and RPL, emphasizing the heterogeneity of immune infiltration in the

occurrence of both diseases. These findings may provide new mechanistic insights or therapeutic targets for further research of EMs and RPL.

#### KEYWORDS

endometriosis, immune infiltration, recurrent pregnancy loss, biomarker, endometrial cancer

## 1 Introduction

Endometriosis (EMs) is an estrogen-dependent chronic gynecological inflammatory disease. Its most typical clinical symptoms are endometrial tissues in non-uterine cavity (Mikhaleva et al., 2020). About 10% of childbearing women suffer from EMs. The reason for EMs currently accepted is that the endometrium tissue is retrograde in the pelvis with menstruation, scattered to different areas, producing local inflammatory reactions, causing local tissue adhesion, and then dysmenorrhea, and dysmenorrhea. EMs can lead to multiple clinical symptoms such as pelvic pain and infertility (Barnhart et al., 2002). The prevalence of EMs among women of reproductive age is largely due to the lack of effective screening methods in the early days. EMs often has two treatments in clinical practice: first, drug treatment, oral hormone drug treatment, this treatment can only help EMs patients to relieve symptoms. When drug treatment is interrupted, the probability of recurrence. Second, surgical treatment, an endometrium tissue with abnormal hyperplasia through laparoscopic surgery, separated the adhesive part, but surgery may cause the ovarian reserve function to decrease, and the risk of recurrence is higher. Although EMs is a benign gynecological disease, the pain caused to patients and the risk of a certain vicious transformation of patients brings a lot of trouble to many patients with childbearing age (Barreta et al., 2018). At present, in clinical practice, an efficient diagnosis and treatment plan is needed.

The stability of the function and environment of the uterus and its accessories is related to the stability of pregnancy and the quality of pregnancy. Clinically, embryo or fetal loss of twice or more than two or more pregnancy failure of pregnancy before 20–24 weeks of pregnancy is diagnosed with recurrent pregnancy loss (RPL, Recurrent Pregnancy Loss), and 5% of women try to experience this pain when they are pregnant (Bohmann et al., 2010; Practice Committee of the American Society for Reproductive Medicine, 2020). The literature reports about 23 million abortions each year, and the prevalence of recurrent abortion is about 2.6% (Quenby et al., 2021). RPL not only brings tremendous psychological pressure to both husband and wife, but it also brings a heavy burden on family, social, and public health systems. The cause of RPL is complicated, and about 50% of the cause of RPL is unknown (Li et al., 2023). The onset of RPL includes many factors such as uterus and accessories, endocrine, immune, genetic, environmental, lifestyle, psychological conditions, etc., (Dimitriadis et al., 2020; Liu et al., 2021). The normal and stability of the environment in the uterus and the accessories are crucial. Due to the lack of a clear understanding of the RPL mechanism, the treatment of patients with RPL is more complicated. A survey of epidemiological diseases found that nearly half of the patients were related to EMs, and many patients caused the failure of auxiliary reproduction (Homer, 2019).

To explore the risk factors of RPL and tap their pathogenesis, more and more researchers have explored the relationship between genetic factors and RPL. However, there is no high-quality monitoring and treatment method for RPL. We focus on EMs, using multiple bioinformatics and machine learning approaches, such as WGCNA, PPI network construction, function enrichment, immune infiltration deconvolution, etc., To reveal the common key biomarkers and their functions in EMs and RPL. A total of eight microarray datasets including 291 samples were enrolled from the gene expression Omnibus (GEO) database. The potential upstream regulatory factors of the biomarkers, i.e., miRNAs or transcription factors were predicted. Finally, we unveiled the pivotal role of two key genes during the development and prognosis of endometrial carcinoma. All these data offer an insightful understanding of molecular mechanisms in EMs and RPL, providing promising directions for early detection of EM, and future extensive studies of these biomarkers may shed light on drug candidate discovery and optimized clinical plans for system risk assessment and personalized therapy of RPL.

## 2 Materials and methods

### 2.1 Data source

We retrieved eight datasets from the GEO database, including two RPL datasets (accessing ID: GSE26787 and GSE165004) and six EMs datasets (accessing ID: GSE51981, GSE6364, GSE7305, GSE120103, GSE23339, and GSE7846). Detailed information on these datasets can be found in [Supplementary Table S1](#). All microarray data preprocessing was performed as previously described (Zhang et al., 2021; Tang et al., 2022b). The “Combat” function implemented in the “SVA” package (v3.42.0), which is a popular algorithm for adjusting systematic variations between arrays, was utilized to generate a combined meta-dataset by removing the batch effects of different sets for both RPL and EMs (Zhang et al., 2023; Guo et al., 2024), respectively, which was then used to perform subsequent bioinformatics analysis. The online TIMER2.0 database (<http://timer.cistrome.org/>) was used to compare the expression levels of the key genes. Additionally, the RNA-seq profile and corresponding clinical information of EC patients (TCGA-UCEC) were acquired from the UCSC XENA database (<https://xena.ucsc.edu>) (Guo et al., 2021; Zhao et al., 2023), and log2 transformed Transcripts per kilobase million (TPM) value was computed for normalization (Guo et al., 2022; Tang et al., 2022b).



## 2.2 DEGs screening

DEGs in EMs and RPL were screened using the R package “Limma” (v3.50.3) with the threshold of  $|\log FC| > 0.585$  and  $\text{adj.P.Val} < 0.05$  (Sun W. et al., 2023). Considering the imbalanced sample composition for EMs, we utilized the subset-based approach to create the 1:1 ratio for the disease and control groups (Tang et al., 2022a). The common DEGs derived from all five subsets were intersected by drawing a Venn diagram, which were defined as the DEGs of EMs. Subsequently, we obtained the intersection of DEGs from EMs and RPL and visualized their expression patterns using heatmap via the “pheatmap” package (v1.0.12).

## 2.3 WGCNA analysis

Based on the combined datasets of RPL and EMs, the “WGCNA” package (v1.71) was applied to define functional transcriptomic co-expression modules shared by both diseases. Outlier samples were excluded according to the sample clustering plot, followed by the selection of the best soft thresholding power ( $\beta$ ) and the establishment of a scale-free network. Dynamic and merged gene modules were generated according to the module size and similarity. The module-trait relationships were determined by Pearson correlation analysis to identify important gene modules, and gene significance (GS) and module membership (MM) were evaluated and scatter plots were depicted.

## 2.4 PPI network construction and function enrichment

PPI analysis was conducted by the STRING database (version 12.0, <http://string-db.org/>) with the confidence score of  $\geq 0.25$  and was further visualized by the Cytoscape software (v3.2.1). The MCODE app was used for cluster analysis. The topological parameters of the network including node degree and maximal clique centrality (MCC) were computed by the “cytohubba” application, which were used to pick out the core genes of the network.

For a better understanding of the potential biological function of identified gene modules from WGCNA, GO and KEGG pathway enrichment analyses were performed using the Metascape (<http://metascape.org>) with default settings.

## 2.5 Machine learning algorithms to discover potential diagnostic markers

Two machine learning algorithms including the least absolute shrinkage and selection operator (LASSO) regression and support vector machine-based recursive feature elimination (SVM-RFE) were adopted for feature selection to identify candidate diagnostic biomarkers in both RPL and EMs. The expression of the common DEGs and core genes of the PPI network in the combined dataset served as the input matrices for both RPL and EMs. LASSO was carried out by the “glmnet” package (v4.1.8) and SVM-RFE

was carried out by the “e1071” package (v1.7.16). Ten-fold cross-validation was applied to calculate the misclassification error of the candidate model and to determine the optimal lambda.  $\lambda_{1se}$  was used to pick out the least number of variables for LASSO. For SVM-RFE, a ranked gene list was created using the SVM-RFE algorithm, followed by the optimal subset acquisition with a linear kernel. We obtained the intersection of LASSO and SVM-RFE to define the core genes for RPL and EMs, respectively, and the overlapping core genes of the two diseases were further regarded as common core genes representing diagnostic biomarkers of patients with EMs and RPL.

## 2.6 Diagnostic/therapeutic value assessment and immune infiltration analysis

Based on the combined datasets, the “pROC” package (v1.18.4) was used to depict the ROC curves to assess the diagnostic values of the common core genes in patients with EMs and RPL. Besides, Drug Signatures DataBase (DSigDB, v1.0), which is a comprehensive gene set resource consisting of 17,389 unique compounds covering 19,531 genes, was used to predict potential drugs/compounds that may disturb the expression of the identified core biomarkers in both EMs and RPL. For immune infiltration estimation, we utilized the CIBERSORTx algorithm (Newman et al., 2019) to obtain the abundance of immune cells in RPL and EMs based on the combined datasets of the two diseases. The immune infiltration landscape in different groups of RPL or EMs was presented by a heatmap plot. Spearman correlation analysis was used to assess the relationship between the expression of the common core genes and immune cell infiltration.

## 2.7 Upstream regulation prediction

MiRNet 2.0 is an up-to-date integrated platform that illustrates “multiple-to-multiple” relationships and functional interpretation. In this study, MiRNet 2.0 was applied to examine the upstream regulation of the abovementioned common core genes. Predicted transcription factor-key gene interactions were inferred from the Jasp database that plugged in MiRNet 2.0. Moreover, miRNAs that may target these common core genes were predicted with miRTarBase 9.0, an online database to predict miRNA-target interactions (MTIs) which were verified by series types of cell-based experiments including microarray, CLASH, PAR-CLIP, HITS-CLIP, and proteomics. The interactions between predicted transcription factors or miRNAs with common core genes were displayed by Cytoscape.

## 2.8 Comprehensive exploration of the common core genes in EC

TIMER 2.0 was used to compare the expression levels of the common core genes between tumor and adjacent normal tissues in pan-cancer based on the whole TCGA cohort. The TCGA-UCEC dataset was used to assess the correlation between gene expression and clinicopathological features including age, BMI, grade,

histology, stage, and tumor burden. Immune score and stromal score were computed with the ESTIMATE algorithm (Yoshihara et al., 2013) via the “estimate” package (v1.0.13). Kaplan-Meier survival plots for overall survival (OS) and disease-free survival (DFS) were drawn to evaluate the prognostic power of common core genes in EC.

## 2.9 Human samples collection

Three samples of ectopic endometrial lesions were collected from women aged 18–37 years, and the diagnosis was confirmed by an experienced histopathologist. Three proliferative-phase endometrial specimens were collected from regular cycling women undergoing hysterectomy for non-endometrial conditions. The tissues were fixed in formalin and embedded in paraffin for immunohistochemistry. The written informed consent for use of human tissues was obtained from each subject, and ethical approval was obtained from the Suining Centre Hospital (No. KYLLKS20240183).

## 2.10 Immunohistochemistry

Tissue sections (5 µm) were dewaxed, EDTA antigen repair solution (MXB, MVS-0099) for antigen thermal repair, 3% H<sub>2</sub>O<sub>2</sub> quenches endogenous peroxidase, 1% BSA blocked 20 min at room temperature, washed with ddH<sub>2</sub>O and PBS, and then probed with the primary antibody (RIBC2, Proteintech, Rosemont, IL 60018, United States; MAN2A1, HUABIO Hangzhou, Zhejiang P.R.C; PAPSS1, Proteintech, Rosemont, IL 60018, United States) at 1/100 dilution overnight at 4°C. Following washing with 0.025% Tween/Tris, sections were incubated with horseradish peroxidase (HRP)-labeled streptavidin-biotin detection kits according to the manufacturer's instructions (Dako, Carpinteria, CA). The color was developed by incubation with DAB chromagen (3, 3'-diaminobenzidine; Dako, Carpinteria, CA) for 5 min.

## 3 Results

We downloaded two RPL datasets and six EMs datasets from the GEO database, respectively. After batch removal in individual diseases, we obtained two combined datasets for RPL and EMs, which were further used for downstream analysis. The principal component analysis (PCA) plots indicated that batch effects were well corrected for both EMs and RPL (Supplementary Figure S1). The combined RPL set was composed of 29 RPL and 29 control tissue samples, while the combined EMs set consisted of 141 EMs and 92 control tissue samples. The complete flowchart of the current research is presented in Figure 1.

### 3.1 DEGs screening in RPL and EMs

Using the combined datasets of RPL and EMs, we screened the DEGs in these two diseases with the predefined filtering criteria ( $\text{abs}(\log_2\text{FC}) > 0.585$ ,  $\text{adj.P.Val} < 0.05$ ). For RPL, a total of 330 DEGs were

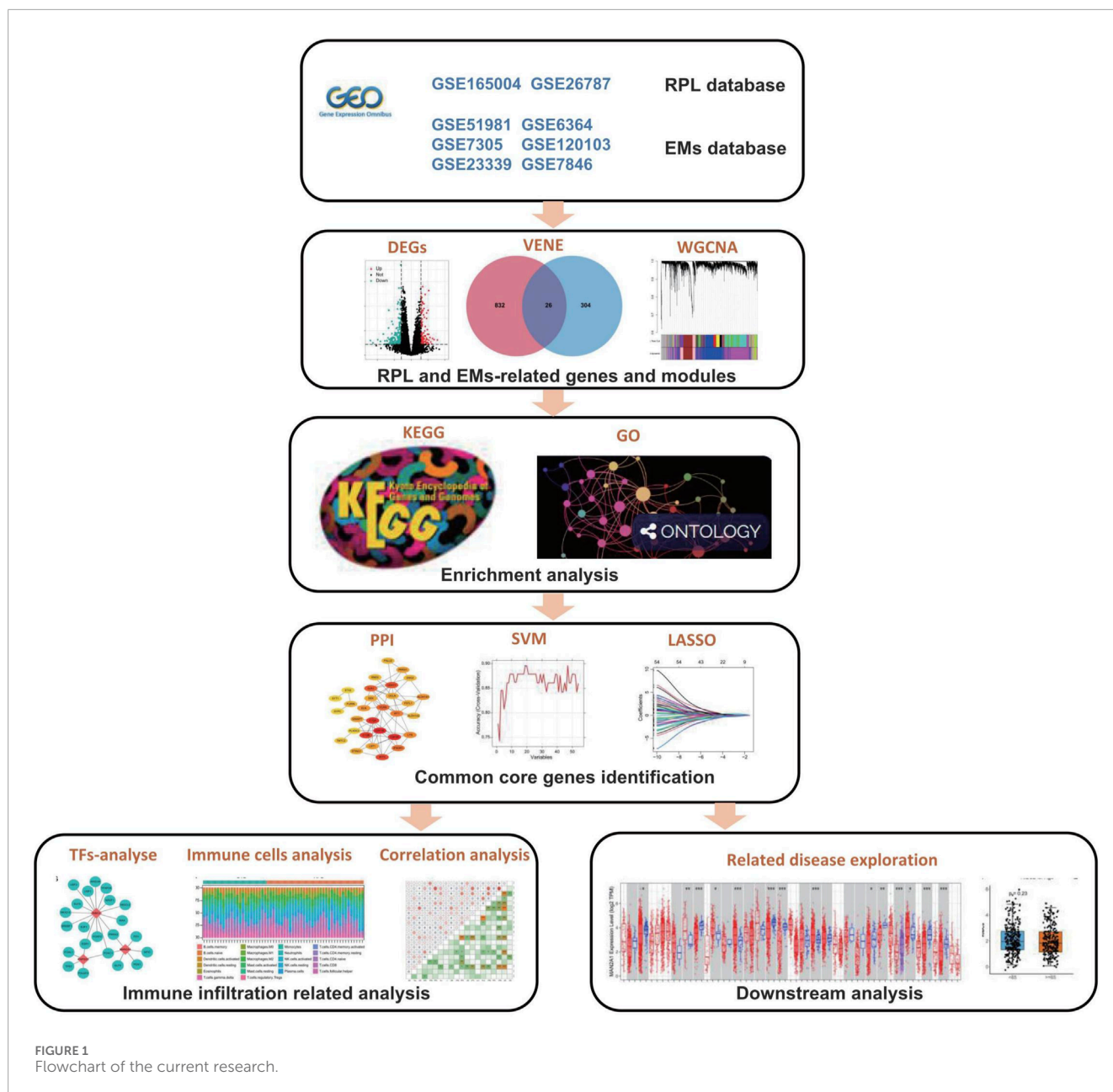
achieved between 29 pairs of RPL and control samples (Figure 2A). For EMs, we first screened the DEGs between five subsets EMs ( $n = 92$ ) and control ( $n = 92$ ). Then, we intersected the DEGs to get 858 DEGs (Figures 2B, C). Next, overlapping analysis of the Venn diagram resulted in 26 common DEGs of RPL and EMs (Figure 2D). The expression patterns of these common DEGs in two combined datasets of RPL and EMs were shown by heatmap (Figures 2E, F).

### 3.2 WGCNA to identify key modules in RPL and EMs

WGCNA network was constructed to unveil the co-expression gene patterns and identify key modules related to EMs or RPL. Outlier samples were screened by sample clustering for both EMs and RPL (Supplementary Figure S2), followed by the determination of optimal  $\beta$  values for meaningful scale-free topology (Figures 3A, B). Subsequently, genes in the RPL and EMs dataset were clustered into 5 and 7 modules, respectively. Pearson correlation analysis revealed that the greenyellow gene module was mostly positively related to RPL ( $r = 0.37$ ,  $p = 0.004$ ), while the brown and blue gene modules were mostly negatively connected with EMs ( $r = -0.39$ ,  $p = 6 \times 10^{-10}$ ;  $r = -0.38$ ,  $p = 4 \times 10^{-10}$ ) (Figures 3C, D). Besides, scatter plots of these modules were presented with GS and MM in Figures 3E–G.

### 3.3 Function enrichment and pathway analysis of the three key gene modules in RPL and EMs

We obtained three key gene modules in RPL and EMs via WGCNA, including the greenyellow gene module in RPL and the brown and blue modules of EMs. To understand the biological functions of three key gene modules in RPL and EMs, we performed GO and KEGG analysis. We displayed the top 20 results of GO biological process enrichment analysis with the green-yellow gene module related to RPL and the brown and blue gene modules related to EMs, respectively (Figures 4A–C). To further investigate the link between these GO enrichment results of RPL and EMs, we took the intersection. A total of 67 were retrieved, focused on three GTPase-related results, including GTPase regulator activity, GTPase activator activity, GTPase activity and protein tyrosine kinase activity, ribonucleoside triphosphate phosphatase activity, glycosyltransferase activity, scaffold protein binding, and transcription factor binding, etc. On the other hand, KEGG pathway enrichment analysis found that the greenyellow gene module was mostly involved in 96 pathways, while the blue gene module was major related to 76 pathways. Moreover, the brown gene module was closely related to 69 pathways. The top 20 enriched KEGG pathways of three key gene modules in RPL and EMs are shown in Figures 4D–F. The intersection of KEGG results of three key gene modules indicated that the PI3K-Akt signaling pathway, pathways in cancer, p53 signaling pathway, MAPK signaling pathway, AMPK signaling pathway, Wnt signaling pathway, and other important KEGG term might be involved in the pathological process of RPL and EMs.

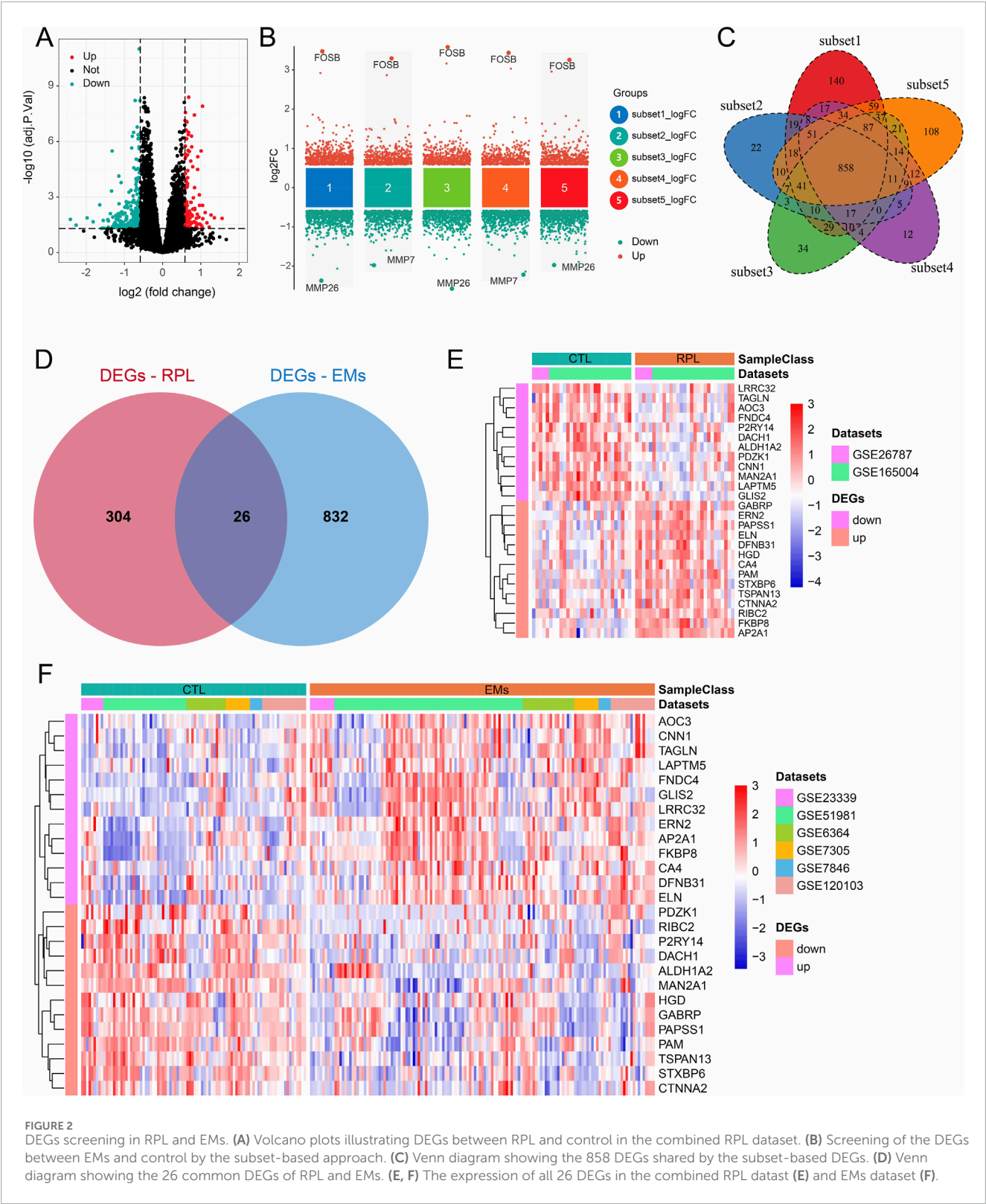


### 3.4 PPI network analysis

From the three gene modules related to RPL with EMs, 98 key genes were identified and used to construct a PPI network with the STRING database (Figure 5A). Then, the MCC of each node was calculated and the top 30 genes with the highest MCC were identified as potential hubs which might be closely related to RPL and EMs (Figures 5B, C). Next, we investigated the expression correlation of the 30 core genes in RPL and EMs. The correlation matrices in Figure 5D suggested that most of the 30 genes were positively related in both RPL and EMs.

### 3.5 Identification of the common core genes in RPL and EMs

Based on the 30 core genes of the PPI network and the 26 shared DEGs of RPL and EMs, we obtained 54 genes as the core genes for RPL associated with EMs (two duplicate genes were removed), which were then used to perform LASSO and SVM-RFE approaches to get diagnostic biomarkers for disease diagnosis. The workflow for screening design is presented in Figure 6A. Following the 10-fold cross-validation procedure, LASSO regression identified 17 candidate core genes of RPL (Figures 6B, C), and the SVM-RFE method screened 19 candidate core genes in RPL



(Figure 6D). After taking the intersection (Figure 6H), a total of 12 core genes were retrieved in RPL. On the other hand, through LASSO regression and SVM-RFE, we obtained 16 and 20 candidate EMs core genes, respectively (Figures 6E–G). The 14 overlapping genes were identified as core genes in EMs (Figure 6I). Finally, we obtained three common core genes in both RPL and EMs by intersecting LASSO regression and SVM-RFE including MAN2A1, PAPSS1, and RIBC2 (Figure 6J).



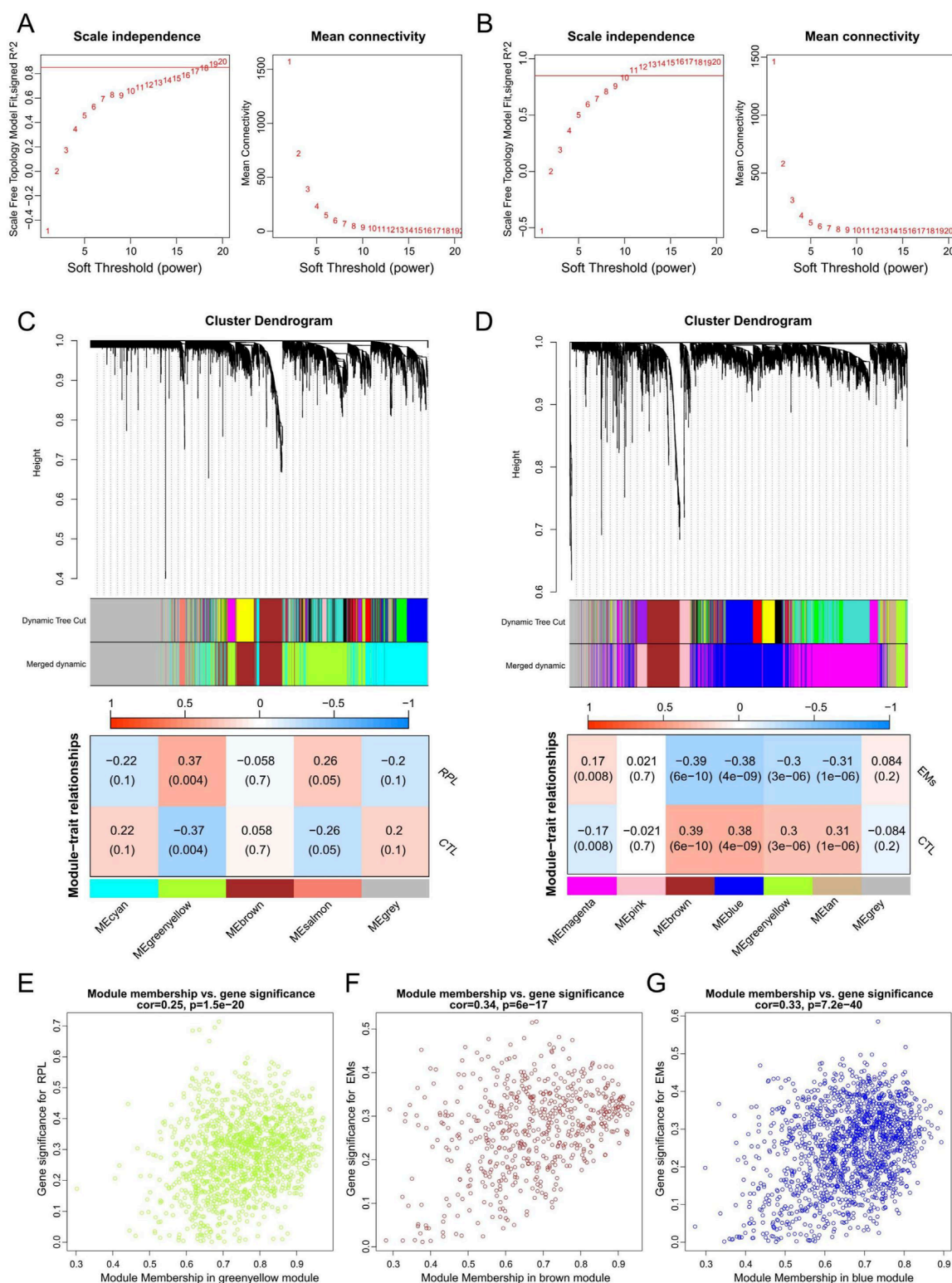
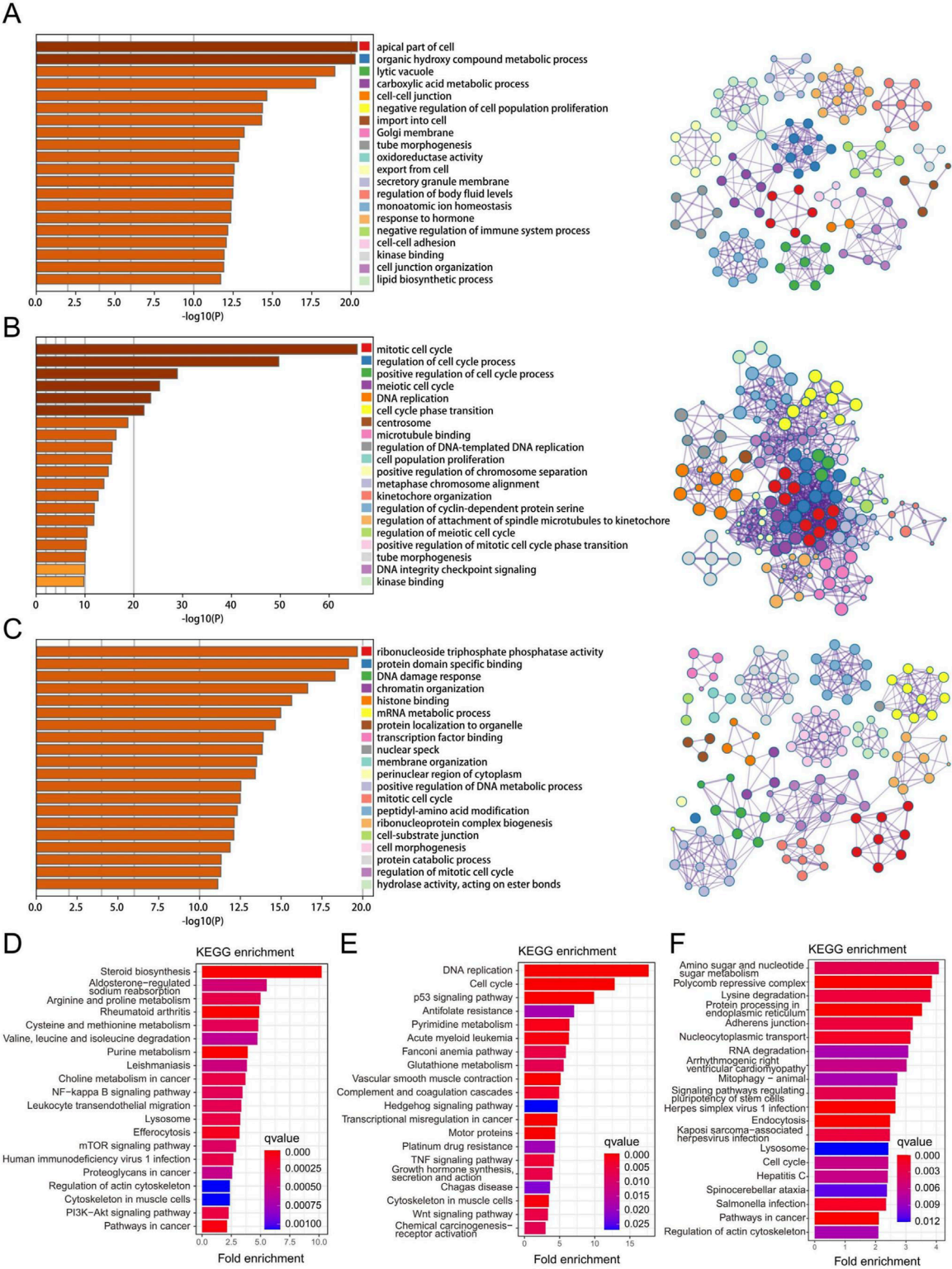


FIGURE 3

WGCNA to identify key gene modules in RPL and EMs. (A, B) The selection of the optimal soft-threshold power in RPL (A) and EMs (B). (C, D) Gene cluster dendrograms and correlation analysis of module eigengenes of RPL (C) and EMs (D). (E) Scatter plot illustrating the relationship between GS and MM in the greenyellow module of the combined RPL dataset. (F, G) Scatter plots of the brown (F) or blue (G) gene modules illustrating the relationship between GS and MM in the combined EMs dataset, respectively.





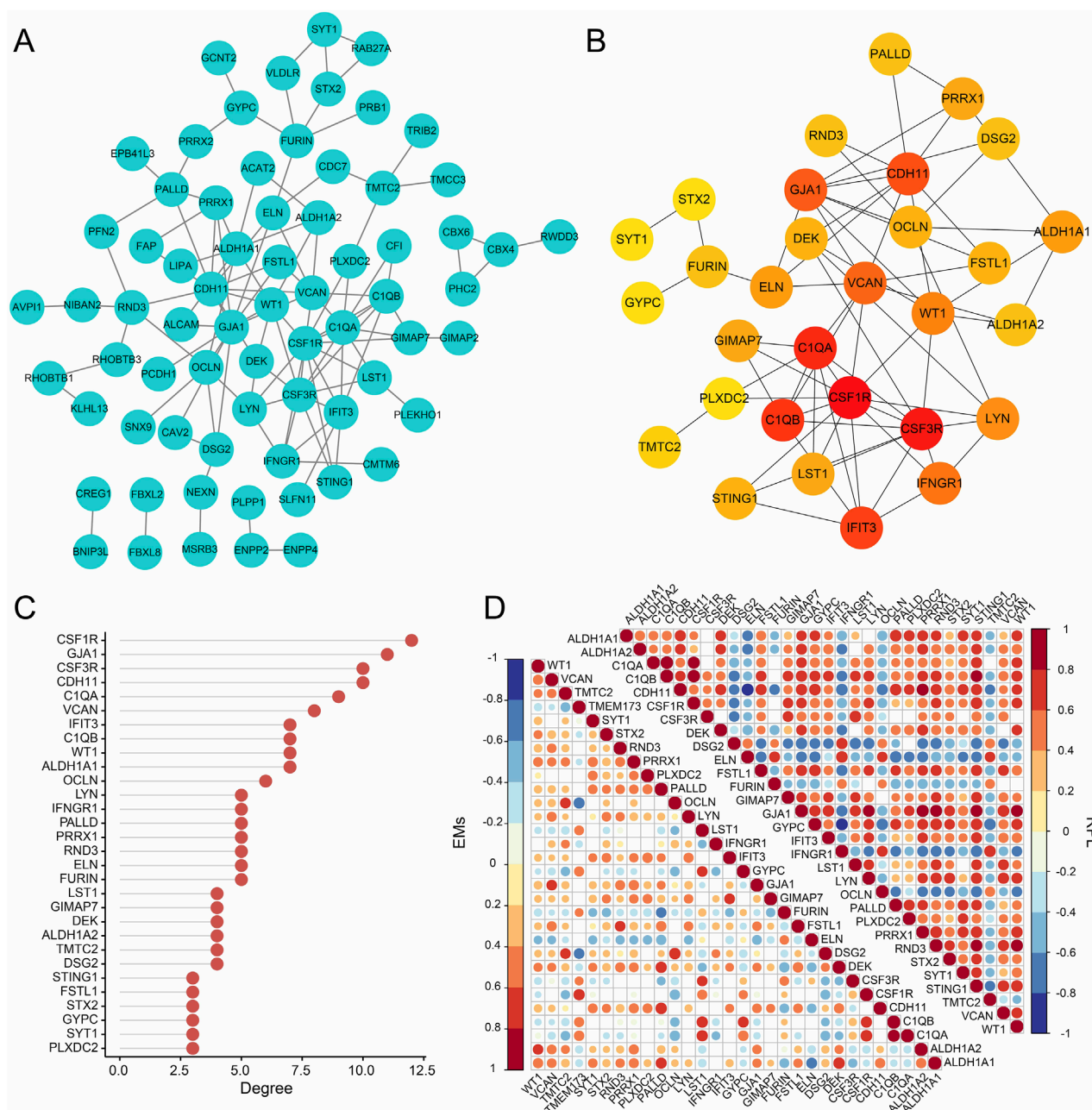


FIGURE 5

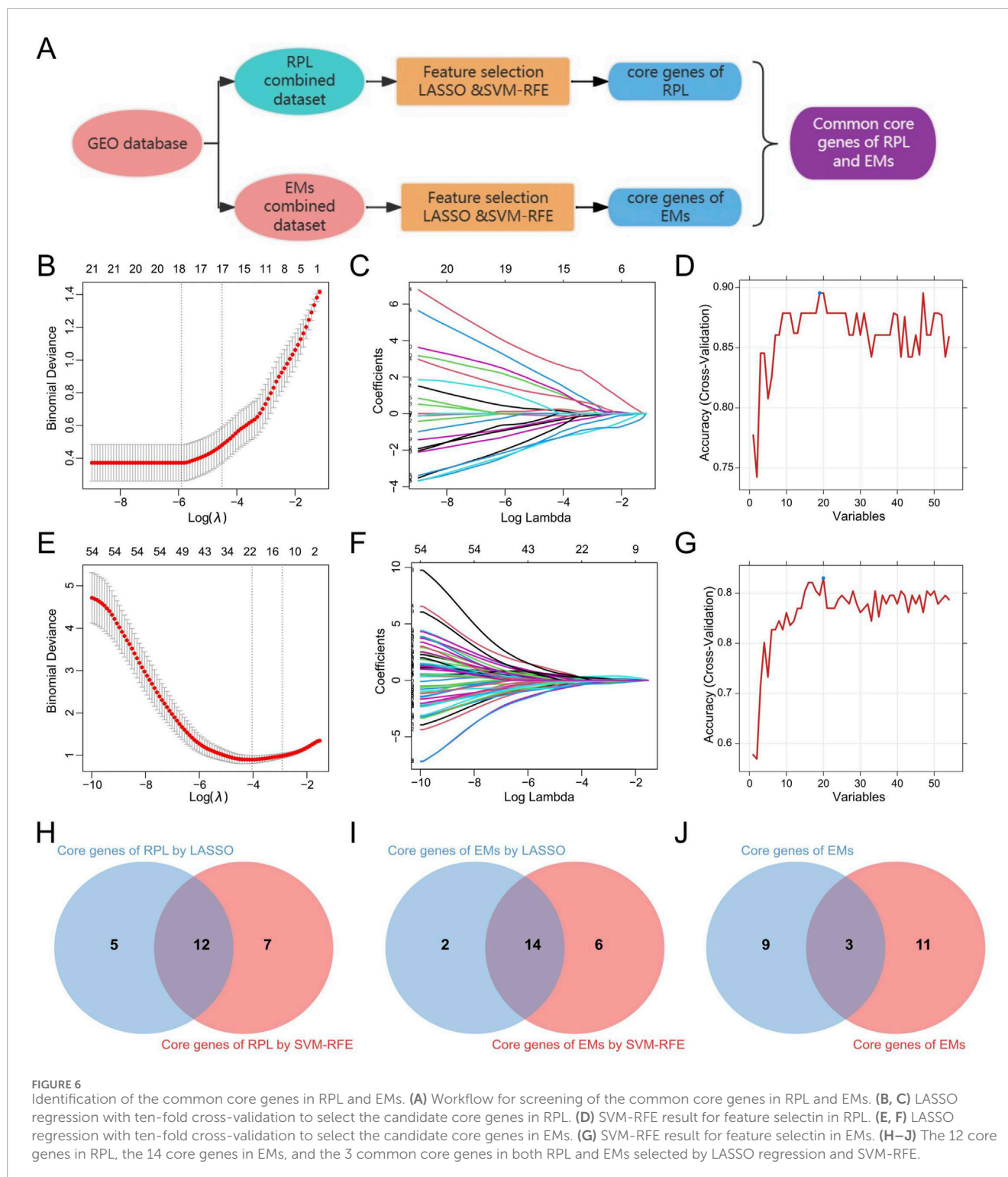
PPI network analysis of the intersection (98 genes) of the key module of RPL and that of EMs (the blue and the brown modules). (A) PPI network construction of the 98 genes by the STRING database. (B) The subnetwork of the core genes of EMs with comorbid RPL (Top 30 genes by the MCC algorithm). (C) The ranking scores of the 30 core genes. (D) The expression correlation of the 30 core genes in the combined RPL dataset and the combined EMs dataset.

### 3.6 Diagnostic/therapeutic values in RPL and EMs and the upstream regulations prediction of the three common core genes

We drew ROC curves of the three common core genes to evaluate their diagnostic value in RPL and EMs (Figures 7A–F). The area under the curves (AUCs) of individual key genes exceeds 0.7, which denotes all three common core genes had significant

diagnostic value in disease discrimination. For therapeutic implications, we inferred the putative small molecules that may perturb the gene expression levels of MAN2A1, PAPSS1, and RIBC2. Consequently, all these three genes were predicted to be targeted by more than 10 therapeutic drugs. A drug-gene interaction network was constructed by Cytoscape for MAN2A1, PAPSS1, and RIBC2, respectively (Figures 7G–I).

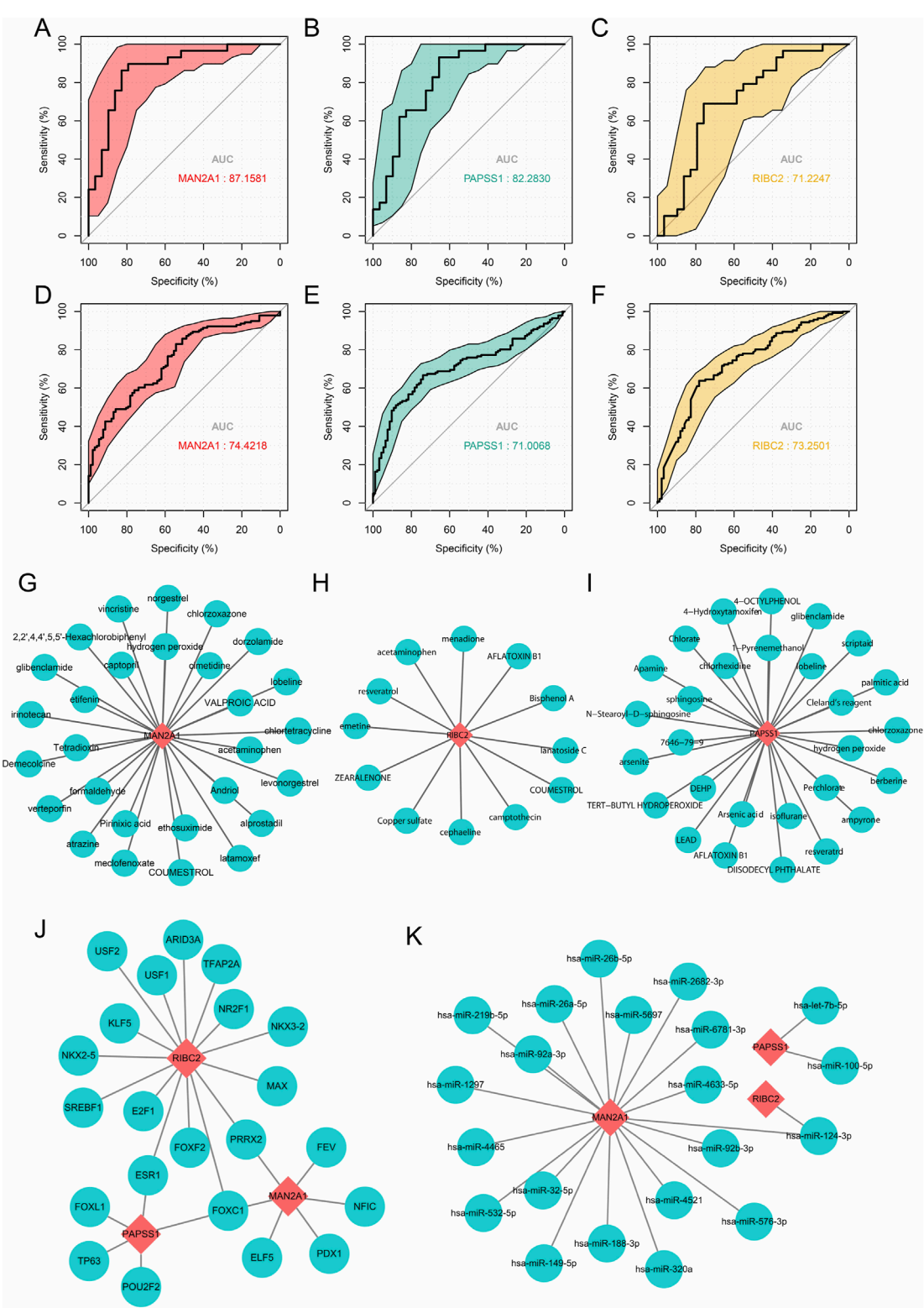
Besides, we predicted potential upstream regulations of the three common core genes related to RPL and EMs,



which involved transcription factors and regulative miRNA (Supplementary Table S2). Predictive transcription factors of these three core genes were derived from the Jaspas database deposited in the miRNet online application, and putative miRNAs targeting these common core genes were obtained from the miRTarget database. A transcription factor-core gene network consisting of 22 transcription factors and a miRNA-target network

consisting of 21 miRNAs were constructed and visualized by Cytoscape (Figures 7J, K). Conspicuously, RIBC2 and MAN2A1 are the most important nodes in the transcription factor-core network or the miRNA-target network, respectively, indicating their crucial role in both RPL and EMs. These findings could provide new insight into the further exploration of the molecular mechanisms of RPL and EMs.





**FIGURE 7** Diagnostic values in RPL and EMs and the upstream regulations prediction of the three common core genes. (A–F) ROC curves showing the AUC values with 95%CI of the three diagnostic markers in RPL and EMs. (G–I) Potential small drugs that may target MAN2A1, PAPSS1, and RIBC2 predicted by the DSigDB database. (J, K) Upstream transcription factors (J) and miRNAs (K) that may regulate the expression of the three common core genes were predicted by miRNet and miRTarBase.

### 3.7 Evaluation of immune cell infiltration in RPL and EMs

Through the CIBERSORTx algorithm, we achieved the landscape of the 22 immune infiltration cells for RPL and EMs (Figures 8A, B). Next, we calculated the correlation of the immune cell infiltration for RPL and EMs. The correlation matrix of immune cell infiltration in RPL and EMs revealed high correlations of the 22 immune infiltration cells in RPL and EMs, which suggested immune infiltration cells may play critical roles in the development of RPL and EMs (Figures 8C, D). The relationship between the three core genes (MAN2A1, PAPSS1, and RIBC2) and immune cells was further investigated using Spearman correlation analysis. As illustrated in Figures 8E–J, distinct results were observed for the correlation analysis of immune cells between the two groups. Specifically, in RPL, MAN2A1 had a positive correlation with Dendritic cells activated ( $r = 0.32$ ,  $p = 0.01$ ) and Macrophages M2 ( $r = 0.28$ ,  $p = 0.04$ ), while it showed a negative correlation with T cells CD8 ( $r = -0.35$ ,  $p = 0.007$ ). PAPSS1 was positively correlated with T cells CD8 ( $r = 0.26$ ,  $p = 0.04$ ), T cells Regulatory Tregs ( $r = 0.28$ ,  $p = 0.03$ ), Eosinophils ( $r = 0.26$ ,  $p = 0.04$ ), and it was negatively correlated with Macrophages M2 ( $r = -0.38$ ,  $p = 0.002$ ), Dendritic cells Activated ( $r = -0.28$ ,  $p = 0.02$ ). RIBC2 exhibited a positive correlation with Plasma Cells ( $r = 0.27$ ,  $p = 0.03$ ), T cells CD8 ( $r = 0.35$ ,  $p = 0.005$ ) and T cells CD4 memory active ( $r = 0.36$ ,  $p = 0.005$ ), and a negative correlation with NK cell activated ( $r = -0.30$ ,  $p = 0.01$ ), B cells Naive ( $r = -0.27$ ,  $p = 0.03$ ), NK cells Resting ( $r = -0.28$ ,  $p = 0.02$ ) (Figures 8E–G; Supplementary Table S3). On the other hand, for EMs, MAN2A1 and PAPSS1 had a positive correlation with Mast cells resting ( $r = 0.31$ ,  $p = 1.36 \times 10^{-6}$  and  $r = 0.36$ ,  $p = 6.55 \times 10^{-9}$ , respectively) and a negative correlation with T cells CD8 ( $r = -0.24$ ,  $p = 0.0001$ ). RIBC2 was positively correlated with T cells CD4 memory resting ( $r = 0.16$ ,  $p = 0.009$ ), and a negative correlation with NK cell activated ( $r = -0.17$ ,  $p = 0.007$ ) (Figures 8H–J; Supplementary Table S4). The above findings indicated that immune cell infiltration had great potential in the occurrence, development, and prognosis of RPL and EMs.

### 3.8 The expression of three hub genes in pan-cancer

Considering the key gene modules of RPL and EMs were enriched in cancer-related KEGG pathways, such as PI3K-Akt signaling pathway, mTOR signaling pathway, p53 signaling pathway, MAPK signaling pathway, cell cycle, and Wnt signaling pathway, etc., which indicate the innate linkage of these diseases and the pathogenesis of cancer, we next analyzed the expression of three core genes in various types of cancer by the TIMER2.0 database (Supplementary Figure S3A–C). Dysregulated expressions were observed for all three core genes in certain cancer types. Notably, MAN2A1 and RIBC2 exhibited significantly differential expression in endometrial cancer compared with adjacent normal tissue, while PAPSS1 had no significant difference. The result suggested that MAN2A1 and RIBC2 may act as a risk or protective factor in the occurrence, development, and prognosis of endometrial cancer.

### 3.9 Comprehensive analysis of RIBC2 and MAN2A1 in endometrial cancer

To investigate the relevance to clinical variables, the RIBC2 expression in the RNA-seq profile of endometrial carcinoma patients was assessed by different age, BMI, grade, histology, stage, and tumor burden status. Consequently, except for age, the RIBC2 expression was significantly associated with BMI, grade, histology, stage, and tumor burden status (Figures 9A–F). Meanwhile, the MAN2A1 expression was significantly associated with age, grade, and histology. There were no significant associations in BMI, stage, and tumor burden status (Figures 9K–P). Moreover, while RIBC2 exhibited a negative correlation with stromal score (Figure 9H), MAN2A1 was negatively associated with immune score (Figure 9Q). The results suggested their predominant roles in the initiation of carcinogenesis and the involvement of the heterogeneous tumor microenvironment in EC. For Kaplan-Meier survival analyses, both the OS survival and the DFS survival were shown to be significantly higher in the high-expression group of RIBC2 (Figures 9I, J). Similarly, the high-expression of MAN2A1 group showed a significantly superior OS survival rate than that of the low-expression group of MAN2A1 (Figures 9S, T).

### 3.10 Protein expression of RIBC2, MAN2A1, and PAPSS1 in the clinical samples of EMs

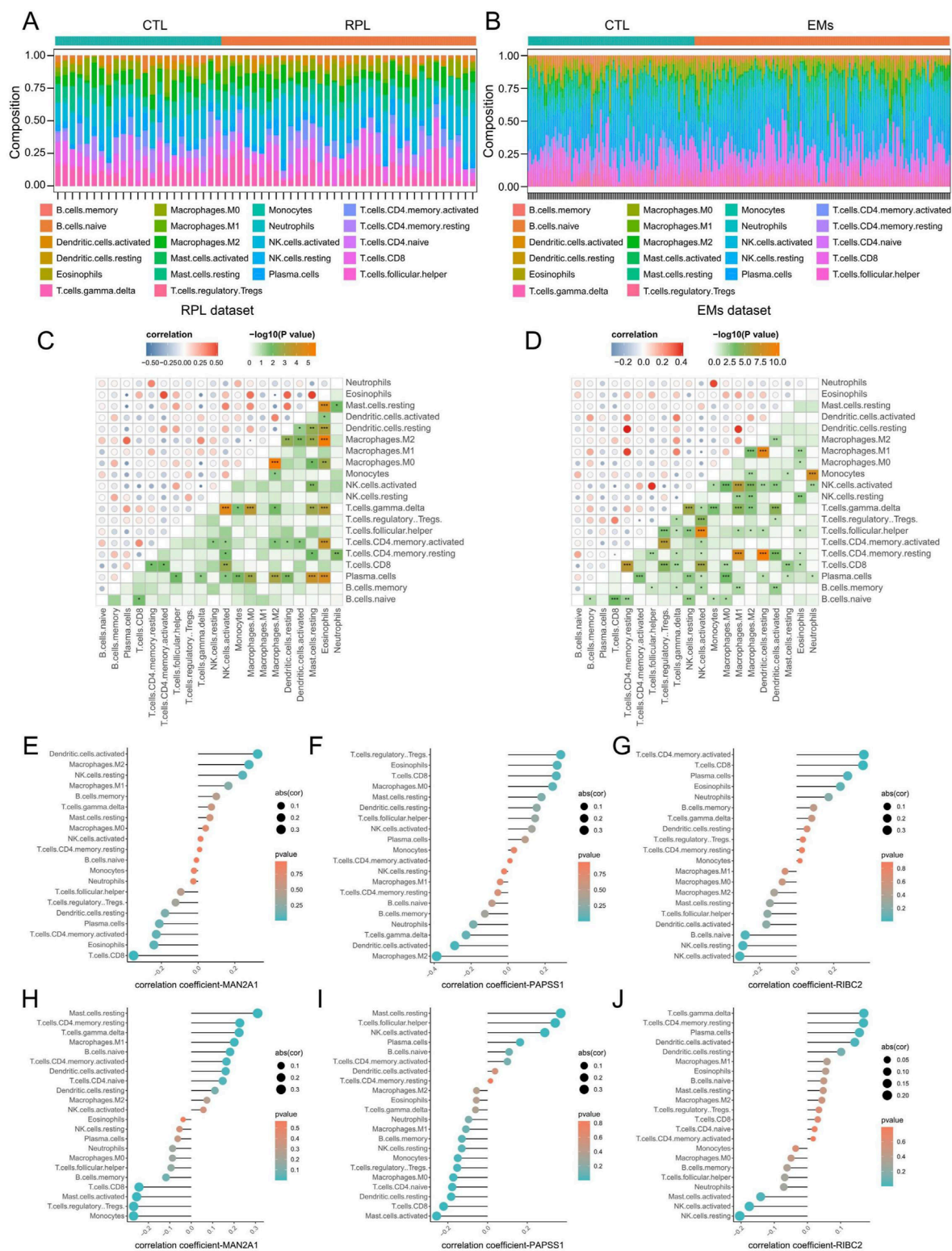
To further verify the expression levels of the abovementioned core genes in clinical samples, we enrolled three ectopic endometriotic lesion samples and three endometrial samples for immunohistochemistry assays. As the result, protein expression analysis of endometrial tissue in the hyperplasia stage and EMs showed that RIBC2 gene was expressed in both glandular cells and stromal cells, MAN2A1 was mainly expressed in the cytoplasm of glandular cells, and PAPSS1 was mainly expressed in the nucleus of glandular cells. All three proteins were abnormally expressed compared to endometrial tissue at the hyperplasia stage (Supplementary Figure S4).

## 4 Discussions

Endometriosis is a chronic inflammatory disease, mainly characterized by the colonization and growth of endometrial cells outside the uterine cavity. This process seriously endangers women's physical and mental health. Despite the existence of several proposed theories, including retrograde menstruation, endometrial stem cell implantation, residual Mullerian abnormalities, and body cavity metaplasia, there is currently no established theory that can fully explain the pathogenesis of all types of endometriosis lesions. Additionally, the definitive diagnosis and treatment of endometriosis often entails a traumatic experience due to the necessity for surgical procedures to visualize the abdominal or pelvic cavity and to remove the lesions. Endometriosis alters the internal milieu of the uterus and pelvis, thereby increasing the risk of early miscarriage and the development of RPL.

Currently, bioinformatics employs a multidisciplinary approach, drawing upon biology, computer science, information engineering,





**FIGURE 8** Evaluation of immune cell infiltration in RPL and EMs. (A, B) Immune cell landscapes of RPL (A) and EMs (B) estimated by the CIBERSORTx algorithm. (C, D) Correlation matrix of immune cell infiltration in RPL (C) and EMs (D). Blue indicates negative correlation, and red indicates positive correlation. (E–G) Correlation coefficients of MAN2A1, PAPSS1, and RIBC2 with immune cell infiltration in RPL. (H–J) Correlation coefficients of MAN2A1, PAPSS1, and RIBC2 with immune cell infiltration in EMs.

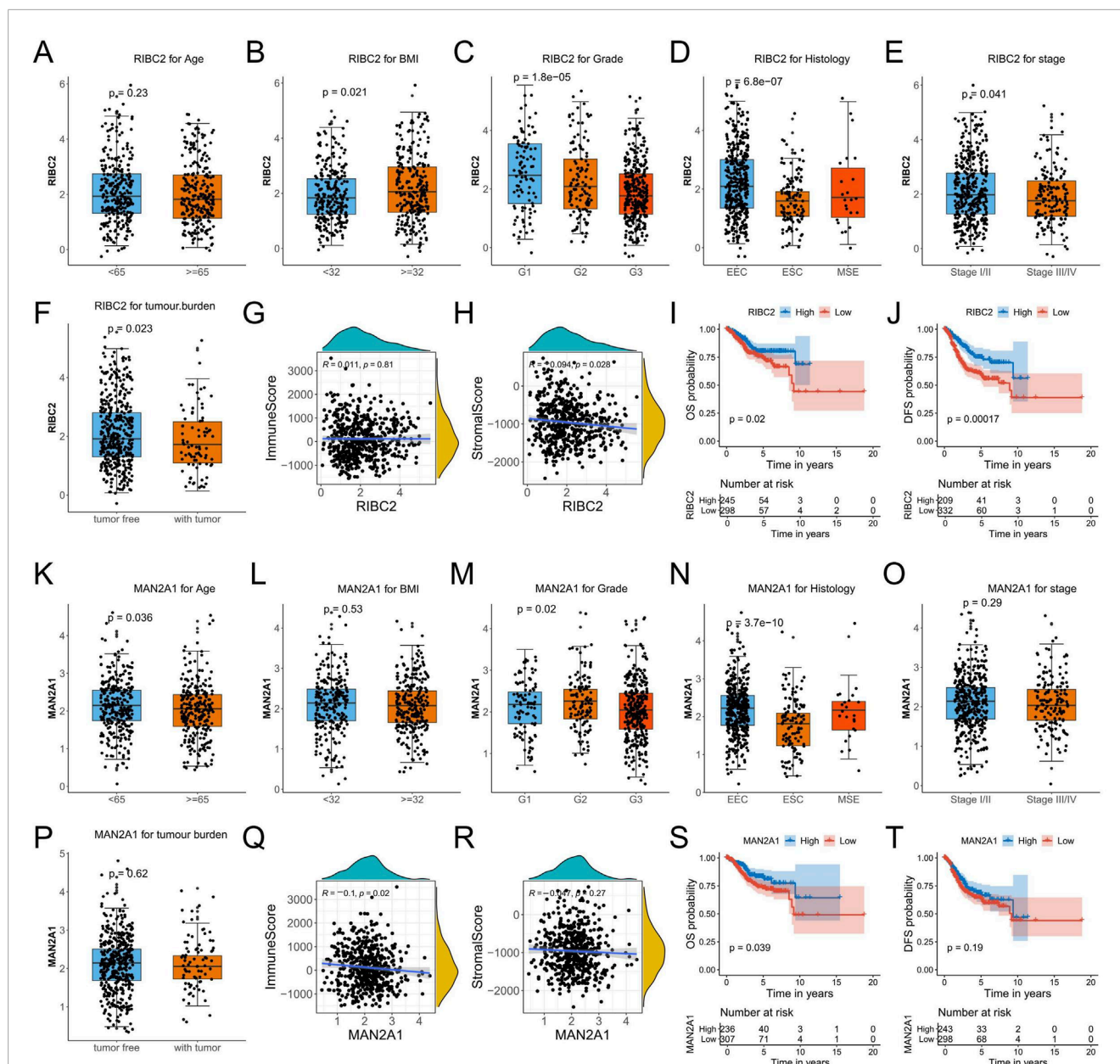


FIGURE 9

Comprehensive analyses of RIBC2 and MAN2A1 based on the TCGA-UCEC dataset. (A–F) Comparison of the RIBC2 expression in subsets of endometrial carcinoma patients with different age, BMI, grade, histology, stage, and tumor burden status. (G, H) Spearman correlation analysis between RIBC2 and immune score (G) and stromal score (H). (I, J) Kaplan-Meier survival plot of OS and DFS for the expression of RIBC2. (K–P) Comparison of the MAN2A1 expression in subsets of endometrial carcinoma patients with different age, BMI, grade, histology, stage, and tumor burden status. (Q, R) Spearman correlation analysis between MAN2A1 and immune score (Q) and stromal score (R). (S, T) Kaplan-Meier survival plot of OS and DFS for the expression of MAN2A1.

and other fields, to analyze and integrate copious amounts of data, with applications in life science research. In this study, we employed the WGCNA methodology to identify putative biomarkers and therapeutic targets employing clustering highly correlated genes and analyzing the correlation between these clusters and clinical features. The gene modules related to EMs and RPL were retrieved using WGCNA, resulting in the identification of 98 shared genes. The PPI network for these genes was constructed using the STRING database, and the key genes were subsequently identified through

LASSO and SVM-RFE. The intersection of these three methods yielded three common core genes in RPL and EMS, MAN2A1, PAPSS1, and RIBC2.

3'-Adenosine 5'-phosphate sulfate (PAPS) synthetase 1 (PAPSS1) catalyzes the synthesis of PAPS from ATP and inorganic sulfate ( $\text{SO}_4^{2-}$ ) through the reaction of ATP thiolase and adenosine 5-phosphate sulfate kinase, which serves as the substrate for cellular sulfonation reactions. PAPSS is expressed widely throughout the human body and may be involved in the sulfation process of various

organs (Xu et al., 2000). Sulfuration regulates numerous biological effects through the modification of proteins, carbohydrates, and lipids. An illustrative example is its role in mediating the initial interaction between leukocytes and endothelial cells (Pouyani and Seed, 1995; Wilkins et al., 1995). This process may be related to early ectopic growth of endometrium. Sulfation represents a major pathway of estrogen metabolism (Foster, 2021). Estradiol sulfate, which can be reactivated by the desulfurization of estrogen sulfatase. The overexpression of PAPSS1 has been demonstrated to impede estrogen-stimulated cell proliferation (Xu et al., 2012). The sulfation mediated by PAPSS1 is implicated in the process of DNA damage repair (Leung et al., 2015), which, in turn, affects the expression of estrogen-responsive genes (Sun L. et al., 2023). The close relationship between papss1 and hormones also suggests that PAPSS1 plays an important role in EMs and RPL, which are closely involved in hormone regulation. Polymorphisms within the ATP sulfurylase domain of the PAPSS1 gene are also strongly correlated with both liver cancer susceptibility and the efficacy of drug therapy (Shih et al., 2009). It would be of interest to ascertain whether this structural polymorphism exerts disparate effects on estrogen metabolism.

Mannosidase alpha class 2A member 1 (MAN2A1), also referred to as MAN2A1, is a glycosidase found within the Golgi apparatus. Its primary function is to cleave N-glycan precursors into mannose (Man) residues, rendering it an essential substrate for the formation of complex N-glycans (Moremen and Robbins, 1991; Misago et al., 1995). Glycan biosynthesis is essential for the maintenance of cellular structure and function and also plays a pivotal role in cell signaling, immune responses, and inflammatory processes. Concurrently, MAN2A1 is a crucial enzyme in the glycosylation of mature membrane proteins, which is essential for ensuring signal transmission and material transport (Fagerberg et al., 2014; Shi et al., 2020). Abnormal MAN2A1 expression has also been observed to affect pathways related to hedgehog, epidermal growth factor, and transforming growth factor signaling (Anvarian et al., 2019). N-glycans are involved in the expression and function of immune cell surface glycoproteins and affect the binding of immune receptors and ligands (Lyons et al., 2015; Li et al., 2018).

The RIB43A domain with coiled-coils 2 (RIBC2) is also known as TRIB2. The TRIB protein family belongs to a highly conserved type of serine/threonine pseudokinase. In the absence of serine/threonine protein kinase activity, the domain is involved in cell signal transduction, acting as a scaffold protein to regulate protein stability and some protein kinase activities. This affects a variety of cellular processes, including proliferation, differentiation, cell cycling, and cell death (Boudeau et al., 2006; Hill et al., 2015; Monga et al., 2022). The TRIB gene family comprises three homologous genes, TRIB1, TRIB2, and TRIB3. Among them, TRIB2 is mainly located in the cytoplasm and participates in the pathogenesis and progression of various tumors (Monga et al., 2022; Hill et al., 2015). TRIB2 mainly regulates target genes at the post-transcriptional level, such as the modification of ubiquitinated proteins and the subsequent alteration of their functionality, as well as the modulation of target genes through the classical AKT or MAPK pathways (Eyers et al., 2017).

Notably, all three of these core markers (MAN2A1, PAPSS1, and RIBC2) in RPL and EMs showed high potential for early detection of both diseases with AUCs greater than 0.7. Thus, more

future studies are expected to develop non-invasive biomarkers for the diagnosis of RPL and EMs based on the three core genes, including cell-free and exosomal-derived miRNAs, circRNAs or DNA/RNA methylation markers. Moreover, we evaluated the utility of MAN2A1, PAPSS1, and RIBC2 for serving as molecular targets in predicting novel drugs with DSigDB database. Remarkably, all of them were predicted to be targeted by more than 10 small compounds, suggesting the great potential in clinical drug discovery. Taken together, the identified core genes are promising targets with significant implications for both the diagnosis and treatment of RPL and EMs in clinical practice.

The correlation analysis of immune cells exhibited that the infiltration of immune cells was closely related to the development of RPL and EMs. The analysis of MAN2A1, PAPSS1, and RIBC2 genes and immune cell infiltration found that the three genes were also closely associated with immune cell infiltration in the two diseases. MAN2A1 has been previously identified as a regulator of the inflammatory response and autoimmune diseases. For instance, MAN2A1 is involved in the development and progression of glomerulonephritis (Chui et al., 2001). Mice lacking MAN2A1 exhibit a deficiency in N-glycans on red blood cells, resulting in anemia and the development of a late-onset autoimmune disease that closely resembles systemic lupus erythematosus (SLE) (Chui et al., 1997; Chui et al., 2001). At the same time, MAN2A1 may affect vascular development and vascular deformation through NAC (antioxidant like N-acetyl cysteine NAC) (Herrera et al., 2017; Ningappa et al., 2021). It was postulated that the effects of MAN2A1 on endothelial cell growth and vascular deformation may be closely related to ectopic endometrial growth and RPL. Additionally, TRIB2 has been demonstrated to exhibit high expression in T cells within normal hematopoietic cells, particularly in lymphoid tissues. Its elevated expression is linked to the T cell receptor (TCR) signaling pathway, particularly the Notch signaling pathway (Breit et al., 2006; Hannon et al., 2012). TRIB2 has been shown to activate p38, thereby reducing the chemotherapy resistance and disease progression of myeloid leukemia (Salome et al., 2018). Collectively, existing evidence suggests that these three genes play a role in regulating immune cells.

Abnormal immune function has been linked to the occurrence and development of tumors. The infiltration of immune cells has also been implicated in regulating the development of tumors. Given that these three genes are engaged in the regulation of immune cells, as well as the significant enriched KEGG pathways of well-known cancer-related terms that appeared in EMs and RPL, such as PI3K-Akt signaling pathway, p53 signaling pathway, MAPK signaling pathway, cell cycle and Wnt signaling pathway, we undertook a pan-cancer analysis of their expression to ascertain whether endometriosis is associated with endometrial cancer. Our findings elucidated that RIBC2 and MAN2A1 are aberrantly expressed in EC. The TCGA-UCEC dataset analysis revealed a correlation between RIBC2 expression and various clinicopathological variables, including body mass index (BMI), grade, histology, stage, and tumor burden status. The Spearman correlation analysis uncovered a negative correlation between RIBC2 and stromal score. The Kaplan-Meier survival plot of OS and DFS for the expression of RIBC2 was found to be associated with both OS and DFS, which is consistent with the published literature. Studies suggest that TRIB2 overexpression negatively regulates the



activity of the WNT signaling pathway and inhibits WNT-mediated transcriptional activity in hepatocellular carcinoma (HCC), which is closely related to patient prognosis (Wang et al., 2013). In HCC, TRIB2 has proven to inhibit Wnt signaling and tumorigenesis by regulating the stability of  $\beta$ TrCP, COP1, and Smurf1 (Xu et al., 2014). In colorectal cancer, TRIB2 is observed to block cell senescence through ap4/p21 signaling (Hou et al., 2018).

Similar outcomes were observed with regard to MAN2A1 expression, which was found to be significantly associated with age, grade, and histology in EC. It has been demonstrated that MAN2A1 is markedly decreased in colorectal cancerous tissues in comparison to adjacent normal tissues. Furthermore, it has been observed that MAN2A1 is downregulated in metastatic colorectal cancer in contrast to non-metastatic colorectal cancer (CRC). A significant negative correlation was observed between increased MAN2A1 expression and the progression of CRC (Wang et al., 2022). The fusion of MAN2A1 with the tyrosine kinase FER results in the transformation of MAN2A1 into an oncogene, leading to the development of multiple cancers. MAN2A1-FER fusions have been identified in prostate cancer, liver cancer, esophageal cancer, and other types of malignant tumors (Yu et al., 2014; Chen et al., 2017; Yu et al., 2021). The MAN2A1-FER fusion facilitates the transport of the FER kinase to the Golgi apparatus, where it activates and promotes cancer through the epidermal growth factor receptor (EGFR) signaling pathway (Lee et al., 2014).

In conclusion, we employed multiple bioinformatics approaches to identify three core genes, MAN2A1, PAPSS1, and RIBC2, which were demonstrated to play a pivotal role in the occurrence and development of EMs and RPL. Subsequent analysis confirmed all these genes exhibited excellent performance in the diagnosis of EMs and RPL, which were also found to be involved in the regulation of immune cell infiltration. Interestingly, two core genes, RIBC2 and MAN2A1, were further identified as key players in the onset and progression of cervical cancer. These findings may prove beneficial in providing potential targets for the diagnosis and prognosis of EMs and RPL patients, as well as offering insights into novel treatment strategies.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Author contributions

JC: Formal Analysis, Methodology, Visualization, Writing–original draft. QL: Formal Analysis, Methodology, Visualization, Writing–original draft. XL: Conceptualization,

Methodology, Software, Writing–review and editing. FL: Formal Analysis, Methodology, Visualization, Writing–original draft. YJ: Investigation, Validation, Writing–original draft. JY: Investigation, Validation, Writing–original draft. LZ: Conceptualization, Methodology, Software, Writing–review and editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was financially supported by 2024 Sichuan Medical Research project of Sichuan Medical Association (S2024085) and Suining Health Science and Technology Program project (24CJDFB02).

## Acknowledgments

We acknowledge the Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases for data sharing.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2025.1529507/full#supplementary-material>

## References

- Anvarian, Z., Mykityn, K., Mukhopadhyay, S., Pedersen, L. B., and Christensen, S. T. (2019). Cellular signalling by primary cilia in development, organ function and disease. *Nat. Rev. Nephrol.* 15 (4), 199–219. doi:10.1038/s41581-019-0116-9
- Barnhart, K., Dunsmoor-Su, R., and Coutifaris, C. (2002). Effect of endometriosis on *in vitro* fertilization. *Fertil. Steril.* 77 (6), 1148–1155. doi:10.1016/s0015-0282(02)03112-6

- Barreta, A., Sarian, L., Ferracini, A. C., Eloy, L., Brito, A. B. C., de Angelo Andrade, L., et al. (2018). Endometriosis-associated ovarian cancer: population characteristics and prognosis. *Int. J. Gynecol. Cancer* 28 (7), 1251–1257. doi:10.1097/JGC.0000000000001317
- Bohlmann, M. K., Luedders, D. W., Strowitzki, T., and von Wolff, M. (2010). Specific secretory phase endometrial leukocytes of women with two and more consecutive idiopathic abortions are not significantly different from healthy controls. *Arch. Gynecol. Obstet.* 281 (6), 983–990. doi:10.1007/s00404-009-1179-9
- Boudeau, J., Miranda-Saavedra, D., Barton, G. J., and Alessi, D. R. (2006). Emerging roles of pseudokinases. *Trends Cell Biol.* 16 (9), 443–452. doi:10.1016/j.tcb.2006.07.003
- Breit, S., Stanulla, M., Flohr, T., Schrappe, M., Ludwig, W. D., Tolle, G., et al. (2006). Activating NOTCH1 mutations predict favorable early treatment response and long-term outcome in childhood precursor T-cell lymphoblastic leukemia. *Blood* 108 (4), 1151–1157. doi:10.1182/blood-2005-12-4956
- Chen, Z. H., Yu, Y. P., Tao, J., Liu, S., Tseng, G., Nalesnik, M., et al. (2017). MAN2A1-FER fusion gene is expressed by human liver and other tumor types and has oncogenic activity in mice. *Gastroenterology* 153 (4), 1120–1132.e15. doi:10.1053/j.gastro.2016.12.036
- Chui, D., Oh-Eda, M., Liao, Y. F., Panneerselvam, K., Lal, A., Marek, K. W., et al. (1997). Alpha-mannosidase-II deficiency results in dyserythropoiesis and unveils an alternate pathway in oligosaccharide biosynthesis. *Cell* 90 (1), 157–167. doi:10.1016/s0092-8674(00)80322-0
- Chui, D., Sellakumar, G., Green, R., Sutton-Smith, M., McQuistan, T., Marek, K., et al. (2001). Genetic remodeling of protein glycosylation *in vivo* induces autoimmune disease. *Proc. Natl. Acad. Sci. U. S. A.* 98 (3), 1142–1147. doi:10.1073/pnas.98.3.1142
- Dimitriadis, E., Menkhorst, E., Saito, S., Kutteh, W. H., and Brosens, J. J. (2020). Recurrent pregnancy loss. *Nat. Rev. Dis. Prim.* 6 (1), 98. doi:10.1038/s41572-020-00228-z
- Eyers, P. A., Keeshan, K., and Kannan, N. (2017). Tribbles in the 21st century: the evolving roles of tribbles pseudokinases in biology and disease. *Trends Cell Biol.* 27 (4), 284–298. doi:10.1016/j.tcb.2016.11.002
- Fagerberg, L., Hallstrom, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell Proteomics* 13 (2), 397–406. doi:10.1074/mcp.M113.035600
- Foster, P. A. (2021). Steroid sulphatase and its inhibitors: past, present, and future. *Molecules* 26 (10), 2852. doi:10.3390/molecules26102852
- Guo, C., Tang, Y., Liu, Z., Chen, C., Hu, X., and Zhang, Y. (2024). Tumor immunological phenotype-derived gene classification predicts prognosis, treatment response, and drug candidates in ovarian cancer. *Genes Dis.* 11 (5), 101173. doi:10.1016/j.gendis.2023.101173
- Guo, C., Tang, Y., Yang, Z., Li, G., and Zhang, Y. (2022). Hallmark-guided subtypes of hepatocellular carcinoma for the identification of immune-related gene classifiers in the prediction of prognosis, treatment efficacy, and drug candidates. *Front. Immunol.* 13, 958161. doi:10.3389/fimmu.2022.958161
- Guo, C., Tang, Y., Zhang, Y., and Li, G. (2021). Mining TCGA data for key biomarkers related to immune microenvironment in endometrial cancer by immune score and weighted correlation network analysis. *Front. Mol. Biosci.* 8, 645388. doi:10.3389/fmolb.2021.645388
- Hannon, M. M., Lohan, F., Erbilgin, Y., Sayitoglu, M., O'Hagan, K., Mills, K., et al. (2012). Elevated TRIB2 with NOTCH1 activation in paediatric/adult T-ALL. *Br. J. Haematol.* 158 (5), 626–634. doi:10.1111/j.1365-2141.2012.09222.x
- Herrera, E. A., Cifuentes-Zuniga, F., Figueroa, E., Villanueva, C., Hernandez, C., Alegria, R., et al. (2017). N-Acetylcysteine, a glutathione precursor, reverts vascular dysfunction and endothelial epigenetic programming in intrauterine growth restricted Guinea pigs. *J. Physiol.* 595 (4), 1077–1092. doi:10.1113/JP273396
- Hill, R., Kalathur, R. K., Colaco, L., Brandao, R., Ugurel, S., Futschik, M., et al. (2015). TRIB2 as a biomarker for diagnosis and progression of melanoma. *Carcinogenesis* 36 (4), 469–477. doi:10.1093/carcin/bgv002
- Homer, H. A. (2019). Modern management of recurrent miscarriage. *Aust. N. Z. J. Obstet. Gynaecol.* 59 (1), 36–44. doi:10.1111/ajo.12920
- Hou, Z., Guo, K., Sun, X., Hu, F., Chen, Q., Luo, X., et al. (2018). TRIB2 functions as novel oncogene in colorectal cancer by blocking cellular senescence through AP4/p21 signaling. *Mol. Cancer* 17 (1), 172. doi:10.1186/s12943-018-0922-x
- Lee, C. C., Shiao, H. Y., Wang, W. C., and Hsieh, H. P. (2014). Small-molecule EGFR tyrosine kinase inhibitors for the treatment of cancer. *Expert Opin. Investig. Drugs* 23 (10), 1333–1348. doi:10.1517/13543784.2014.928283
- Leung, A. W., Dragowska, W. H., Ricaurte, D., Kwok, B., Mathew, V., Roosendaal, J., et al. (2015). 3'-Phosphoadenosine 5'-phosphosulfate synthase 1 (PAPSS1) knockdown sensitizes non-small cell lung cancer cells to DNA damaging agents. *Oncotarget* 6 (19), 17161–17177. doi:10.18632/oncotarget.3635
- Li, C. W., Lim, S. O., Chung, E. M., Kim, Y. S., Park, A. H., Yao, J., et al. (2018). Eradication of triple-negative breast cancer cells by targeting glycosylated PD-L1. *Cancer Cell* 33 (2), 187–201.e10. doi:10.1016/j.ccell.2018.01.009
- Li, Q., Chen, S., Dong, X., Fu, S., Zhang, T., Zheng, W., et al. (2023). The progress of research on genetic factors of recurrent pregnancy loss. *Genet. Res. (Camb)* 2023, 9164374. doi:10.1155/2023/9164374
- Liu, X., Chen, Y., Ye, C., Xing, D., Wu, R., Li, F., et al. (2021). Hereditary thrombophilia and recurrent pregnancy loss: a systematic review and meta-analysis. *Hum. Reprod.* 36 (5), 1213–1229. doi:10.1093/humrep/deab010
- Lyons, J. J., Milner, J. D., and Rosenzweig, S. D. (2015). Glycans instructing immunity: the emerging role of altered glycosylation in clinical immunology. *Front. Pediatr.* 3, 54. doi:10.3389/fped.2015.00054
- Mikhaleva, L. M., Davydov, A. I., Patsap, O. I., Mikhaylenko, E. V., Nikolenko, V. N., Neganova, M. E., et al. (2020). Malignant transformation and associated biomarkers of ovarian endometriosis: a narrative review. *Adv. Ther.* 37 (6), 2580–2603. doi:10.1007/s12325-020-01363-5
- Misago, M., Liao, Y. F., Kudo, S., Eto, S., Mattei, M. G., Moremen, K. W., et al. (1995). Molecular cloning and expression of cDNAs encoding human alpha-mannosidase II and a previously unrecognized alpha-mannosidase II isozyme. *Proc. Natl. Acad. Sci. U. S. A.* 92 (25), 11766–11770. doi:10.1073/pnas.92.25.11766
- Monga, J., Adrianto, I., Rogers, C., Gadgeel, S., Chitale, D., Alumkal, J. J., et al. (2022). Tribbles 2 pseudokinase confers enzalutamide resistance in prostate cancer by promoting lineage plasticity. *J. Biol. Chem.* 298 (2), 101556. doi:10.1016/j.jbc.2021.101556
- Moremen, K. W., and Robbins, P. W. (1991). Isolation, characterization, and expression of cDNAs encoding murine alpha-mannosidase II, a Golgi enzyme that controls conversion of high mannose to complex N-glycans. *J. Cell Biol.* 115 (6), 1521–1534. doi:10.1083/jcb.115.6.1521
- Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37 (7), 773–782. doi:10.1038/s41587-019-0114-2
- Ningappa, M., Adenuga, M., Ngo, K. A., Mohamed, N., Narayanan, T., Prasad, K., et al. (2021). Mechanisms of impaired lung development and ciliation in mannosidase-1-alpha-2 (Man1a2) mutants. *Front. Physiol.* 12, 658518. doi:10.3389/fphys.2021.658518
- Pouyani, T., and Seed, B. (1995). PSGL-1 recognition of P-selectin is controlled by a tyrosine sulfation consensus at the PSGL-1 amino terminus. *Cell* 83 (2), 333–343. doi:10.1016/0092-8674(95)90174-4
- Practice Committee of the American Society for Reproductive Medicine (2020). Definitions of infertility and recurrent pregnancy loss: a committee opinion. *Fertil. Steril.* 113 (3), 533–535. doi:10.1016/j.fertnstert.2019.11.025
- Quenby, S., Gallos, I. D., Dhillon-Smith, R. K., Podsek, M., Stephenson, M. D., Fisher, J., et al. (2021). Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. *Lancet* 397 (10285), 1658–1667. doi:10.1016/S0140-6736(21)00682-6
- Salome, M., Magee, A., Yalla, K., Chaudhuri, S., Sarrou, E., Carmody, R. J., et al. (2018). A Trib2-p38 axis controls myeloid leukaemia cell cycle and stress response signalling. *Cell Death Dis.* 9 (5), 443. doi:10.1038/s41419-018-0467-3
- Shi, S., Gu, S., Han, T., Zhang, W., Huang, L., Li, Z., et al. (2020). Inhibition of MAN2A1 enhances the immune response to anti-PD-L1 in human tumors. *Clin. Cancer Res.* 26 (22), 5990–6002. doi:10.1158/1078-0432.CCR-20-0778
- Shih, W. L., Yu, M. W., Chen, P. J., Wu, T. W., Lin, C. L., Liu, C. J., et al. (2009). Evidence for association with hepatocellular carcinoma at the PAPSS1 locus on chromosome 4q25 in a family-based study. *Eur. J. Hum. Genet.* 17 (10), 1250–1259. doi:10.1038/ejhg.2009.48
- Sun, L., Ji, W. X., Li, Y., Li, Z. L., Duan, C. C., Xia, B. R., et al. (2023a). The PAPSS1 gene is a modulator of response to cisplatin by regulating estrogen receptor alpha signaling activity in ovarian cancer cells. *J. Ovarian Res.* 16 (1), 187. doi:10.1186/s13048-023-01262-7
- Sun, W., Wang, J., Wang, Z., Xu, M., Lin, Q., Sun, P., et al. (2023b). Combining WGCNA and machine learning to construct basement membrane-related gene index helps to predict the prognosis and tumor microenvironment of HCC patients and verifies the carcinogenesis of key gene CTSA. *Front. Immunol.* 14, 1185916. doi:10.3389/fimmu.2023.1185916
- Tang, Y., Guo, C., Chen, C., and Zhang, Y. (2022a). Characterization of cellular senescence patterns predicts the prognosis and therapeutic response of hepatocellular carcinoma. *Front. Mol. Biosci.* 9, 1100285. doi:10.3389/fmolb.2022.1100285
- Tang, Y., Guo, C., Yang, Z., Wang, Y., Zhang, Y., and Wang, D. (2022b). Identification of a tumor immunological phenotype-related gene signature for predicting prognosis, immunotherapy efficacy, and drug candidates in hepatocellular carcinoma. *Front. Immunol.* 13, 862527. doi:10.3389/fimmu.2022.862527
- Wang, J., Park, J. S., Wei, Y., Rajurkar, M., Cotton, J. L., Fan, Q., et al. (2013). TRIB2 acts downstream of Wnt/TCF in liver cancer cells to regulate YAP and C/EBP $\alpha$  function. *Mol. Cell* 51 (2), 211–225. doi:10.1016/j.molcel.2013.05.013
- Wang, Y., Zhao, J., Fu, G., Sheng, C., Zhu, J., Zhong, T., et al. (2022). MAN2A1 predicts prognosis and progression through cancer-related pathways in colorectal cancer. *Transl. Cancer Res.* 11 (10), 3686–3697. doi:10.21037/tcr-22-629



- Wilkins, P. P., Moore, K. L., McEver, R. P., and Cummings, R. D. (1995). Tyrosine sulfation of P-selectin glycoprotein ligand-1 is required for high affinity binding to P-selectin. *J. Biol. Chem.* 270 (39), 22677–22680. doi:10.1074/jbc.270.39.22677
- Xu, S., Tong, M., Huang, J., Zhang, Y., Qiao, Y., Weng, W., et al. (2014). TRIB2 inhibits Wnt/ $\beta$ -Catenin/TCF4 signaling through its associated ubiquitin E3 ligases,  $\beta$ -TrCP, COP1 and Smurf1, in liver cancer cells. *FEBS Lett.* 588 (23), 4334–4341. doi:10.1016/j.febslet.2014.09.042
- Xu, Y., Liu, X., Guo, F., Ning, Y., Zhi, X., Wang, X., et al. (2012). Effect of estrogen sulfation by SULF1E1 and PAPSS on the development of estrogen-dependent cancers. *Cancer Sci.* 103 (6), 1000–1009. doi:10.1111/j.1349-7006.2012.02258.x
- Xu, Z. H., Otterness, D. M., Freimuth, R. R., Carlini, E. J., Wood, T. C., Mitchell, S., et al. (2000). Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and PAPSS2: gene cloning, characterization and chromosomal localization. *Biochem. Biophys. Res. Commun.* 268 (2), 437–444. doi:10.1006/bbrc.2000.2123
- Yoshihara, K., Shahmoradgoli, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612
- Yu, Y. P., Ding, Y., Chen, Z., Liu, S., Michalopoulos, A., Chen, R., et al. (2014). Novel fusion transcripts associate with progressive prostate cancer. *Am. J. Pathol.* 184 (10), 2840–2849. doi:10.1016/j.ajpath.2014.06.025
- Yu, Y. P., Liu, S., Nelson, J., and Luo, J. H. (2021). Detection of fusion gene transcripts in the blood samples of prostate cancer patients. *Sci. Rep.* 11 (1), 16995. doi:10.1038/s41598-021-96528-9
- Zhang, Y., Tang, Y., Guo, C., and Li, G. (2021). Integrative analysis identifies key mRNA biomarkers for diagnosis, prognosis, and therapeutic targets of HCV-associated hepatocellular carcinoma. *Aging (Albany NY)* 13 (9), 12865–12895. doi:10.18632/aging.202957
- Zhang, Y., Yang, Z., Tang, Y., Guo, C., Lin, D., Cheng, L., et al. (2023). Hallmark guided identification and characterization of a novel immune-relevant signature for prognostication of recurrence in stage I-III lung adenocarcinoma. *Genes Dis.* 10 (4), 1657–1674. doi:10.1016/j.gendis.2022.07.005
- Zhao, L., Tang, Y., Yang, J., Lin, F., Liu, X., Zhang, Y., et al. (2023). Integrative analysis of circadian clock with prognostic and immunological biomarker identification in ovarian cancer. *Front. Mol. Biosci.* 10, 1208132. doi:10.3389/fmolb.2023.1208132