



OPEN ACCESS

EDITED BY

Luz Pamela Blanco,
National Institutes of Health (NIH),
United States

REVIEWED BY

Bei Shi,
China Medical University, China
Xiaojin He,
First Affiliated Hospital of Anhui Medical
University, China

*CORRESPONDENCE

Qingling Yang
✉ qingling531@163.com
Yingpu Sun
✉ syp2008@vip.sina.com

†These authors have contributed equally to
this work

RECEIVED 27 February 2023

ACCEPTED 04 May 2023

PUBLISHED 16 May 2023

CITATION

Chen W, Yang Q, Hu L, Wang M, Yang Z,
Zeng X and Sun Y (2023) Shared diagnostic
genes and potential mechanism between
PCOS and recurrent implantation failure
revealed by integrated transcriptomic
analysis and machine learning.
Front. Immunol. 14:1175384.
doi: 10.3389/fimmu.2023.1175384

COPYRIGHT

© 2023 Chen, Yang, Hu, Wang, Yang, Zeng
and Sun. 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.

Shared diagnostic genes and potential mechanism between PCOS and recurrent implantation failure revealed by integrated transcriptomic analysis and machine learning

Wenhui Chen^{1,2,3†}, Qingling Yang^{1,2,3*†}, Linli Hu^{1,2,3},
Mengchen Wang^{1,2,3}, Ziyao Yang^{1,2,3}, Xinxin Zeng^{1,2,3}
and Yingpu Sun^{1,2,3*}

¹Center for Reproductive Medicine, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, ²Henan Key Laboratory of Reproduction and Genetics, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, ³Henan Provincial Obstetrical and Gynecological Diseases (Reproductive Medicine) Clinical Research Center, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Polycystic ovary syndrome (PCOS) is a complex endocrine metabolic disorder that affects 5–10% of women of reproductive age. The endometrium of women with PCOS has altered immune cells resulting in chronic low-grade inflammation, which attribute to recurrent implantation failure (RIF). In this study, we obtained three PCOS and RIF datasets respectively from the Gene Expression Omnibus (GEO) database. By analyzing differentially expressed genes (DEGs) and module genes using weighted gene co-expression networks (WGCNA), functional enrichment analysis, and three machine learning algorithms, we identified twelve diseases shared genes, and two diagnostic genes, including GLIPR1 and MAMLD1. PCOS and RIF validation datasets were assessed using the receiver operating characteristic (ROC) curve, and ideal area under the curve (AUC) values were obtained for each disease. Besides, we collected granulosa cells from healthy and PCOS infertile women, and endometrial tissues of healthy and RIF patients. RT-PCR was used to validate the reliability of GLIPR1 and MAMLD1. Furthermore, we performed gene set enrichment analysis (GSEA) and immune infiltration to explore the underlying mechanism of PCOS and RIF cooccurrence. Through the functional enrichment of twelve shared genes and two diagnostic genes, we found that both PCOS and RIF patients had disturbances in metabolites related to the TCA cycle, which eventually led to the massive activation of immune cells.

KEYWORDS

PCOS, RIF (Recurrent Implantation Failure), integrated transcriptomic analysis, machine learning, TCA cycle

1 Introduction

Polycystic ovarian syndrome (PCOS), one of the most common endocrine-reproductive-metabolic disorders in women, is characterized by chronic anovulation, hyperandrogenism, and polycystic ovarian morphology and is consistently associated with obesity, insulin resistance (IR), and low-grade chronic inflammation (1, 2). A variety of immune disorders often accompany PCOS patients, which are associated with infertility and consequently, impact ovarian function, oocyte quality, and endometrial receptivity (ER) (3). Especially in obese patients with high estrogen and androgen levels, dysfunction of immune cells may lead to continuous stimulation of the immune system, increasing pro-inflammatory cells (4). This pro-inflammatory condition can negatively affect critical physiological processes, such as proliferation, migration, and invasion of trophoblastic cells into the endometrium that ultimately lead to embryo implantation failure (5, 6). As a result, PCOS accounts for a nonnegligible part of the gynecological diseases that lead to recurrent implantation failure (RIF) during IVF/ICSI-ET (7).

RIF is defined as a lack of clinical pregnancy after at least four embryos have been transferred in a minimum of three fresh or frozen cycles in a couple under 40 years of age, with an incidence rate of about 15% (8). To date, many studies have shown that RIF etiology is largely attributed to three categories: decreased endometrial receptivity (ER), embryonic defects, and other combined effects (9). Among them, immune factors may play a crucial role in influencing ER and embryo implantation (10, 11). Normal immunological function of the maternal-fetal interface is critical for maintaining ER, which requires interactions with decidual cells, endothelial cells, and infiltrating immune cells (12, 13). These immune cells regulate maternal and fetal antigen responses, trophoblast invasion, and vascular remodeling (14). Recently, several transcriptomic studies of gene and protein expression profiles in the endometrium also identified some immune-related markers in RIF patients (15, 16). Significant progress has been made in search of the mechanism of RIF. However, RIF remains a common and insurmountable adverse event in assisted reproductive medicine.

Furthermore, the mechanism of impaired endometrial receptivity (ER) has attracted research interest in recent years (17, 18). Notably, in patients with PCOS, gene expression profiles are dysregulated (19, 20), with differentially expressed genes closely related to steroid hormone synthesis, inflammation, and oxidative stress (21). The establishment of ER depends on these biological processes. Moreover, some indicators are closely associated with ER (22), such as leukemia inhibitory factor (LIF) (23), homeobox genes A (HOXA), $\alpha v \beta 3$ -integrin, and intercellular junctions (24). As verified by clinical data analysis, damaged ER leads to an increased risk of repeated implant failure in PCOS patients (25). Poor endometrial receptivity is commonly identified as the primary cause of RIF, and abnormal gene expression contributes to ER deficiency (26). Therefore, the bioinformatics background of PCOS women may contain the underlying mechanism of recurrent implantation failure in this population.

In this study, we aimed to explore the potential biomarkers and underlying pathways involved in the development of PCOS and

RIF. To achieve this goal, we integrated transcriptomic data related to PCOS and RIF from the Gene Expression Omnibus (GEO) and applied “LIMMA” package and weighted gene co-expression network analysis (WGCNA) to identify differentially expressed genes and critical modules in each disease. Our analysis using the intersection and three machine learning methods led to the identification of two disease-shared diagnostic genes, GLIPR1 and MAMLD1, with good performance validated by external datasets. Further, we conducted gene set enrichment analysis (GSEA) on each of these two genes to identify common pathways associated with PCOS and RIF. Additionally, we investigated the role of immune cells in the co-pathogenesis of the two diseases through immune infiltration analysis. Our results suggested that the co-pathogenesis of PCOS and RIF might be linked to increased immune response levels arising from abnormal TCA cycle metabolism. In conclusion, this study provides valuable insights into the shared molecular mechanisms underlying PCOS and RIF and highlights the potential of GLIPR1 and MAMLD1 as diagnostic markers for these conditions.

2 Methods

2.1 Data collection and preparation

The data sets related to PCOS and RIF were screened in the Gene Expression Omnibus database (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>) since both diseases were designed in this study. For PCOS, we used the keyword “PCOS” or “granulosa cells” to search gene expression profiles. Inclusion criteria were as follows (1): PCOS patients and normal controls must be included in the profiles (2), The discovery profile should have at least ten samples to ensure accuracy, and (3) granulosa cells should be used for sequencing. Accordingly, we selected three datasets numbered GSE10946, GSE34526, and GSE80432. For RIF, the eligibility criteria were as follows (1): the profiles must include normal controls and the RIF patients, and (2) the sample source must be endometrial tissues. After the screening, the datasets GSE103465, GSE11974, and GSE26787 were included in this study. Among them, GSE10946 and GSE34526 of PCOS, and GSE103465 and GSE11974 of RIF were used as discovery cohorts for analyzing and screening. GSE80432 and GSE26787 were used for external validation for PCOS and RIF, respectively. The whole analytic workflow is shown in Figure 1.

Besides, when preparing the two datasets for each disease, the PCA plot showed a noticeable batch effect in the two disease groups. Thus, using the “sva” R package, which identified and built surrogate variables for high-dimensional data sets, the batch effect was eliminated. After removing the batch effect, the PCA plot was visualized by the “FactoMineR” and “Factoextra” R packages.

2.2 Differential gene expression analysis

After preparing the data for each disease, we compared PCOS and RIF datasets using the Linear Models for Microarray (LIMMA) package in R (version 4.1.2). Differentially expressed genes (DEGs)

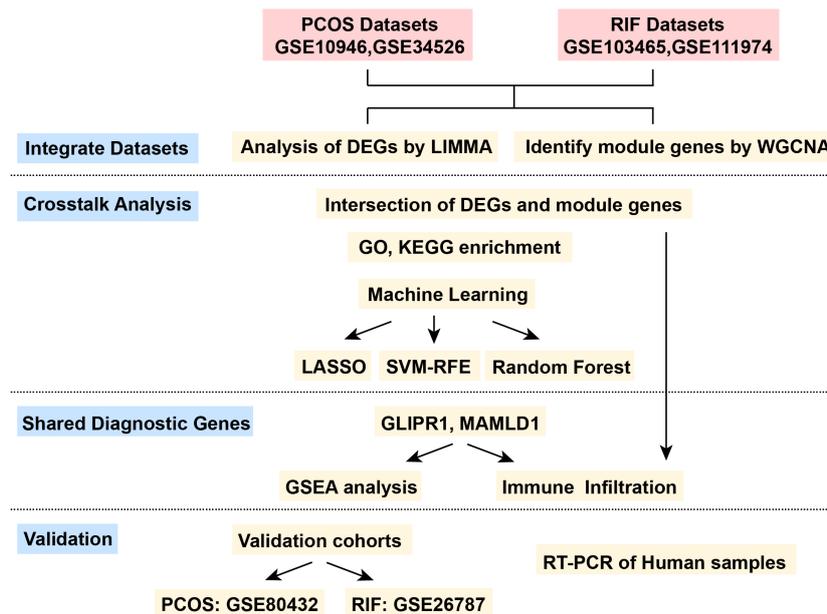


FIGURE 1

The flow chart for the whole design. PCOS, Polycystic Ovarian Syndrome; RIF, Recurrent Implantation Failure; GSE, Gene Expression Omnibus Series; LIMMA, Linear Models for Microarray Data; WGCNA, Weighted Gene Co-expression Network Analysis; DEGs, Differentially Expressed Genes; LASSO, Least Absolute Shrinkage and Selection Operator; SVM-RFE, Support Vector Machine- Recursive Feature Elimination; GSEA, Gene Set Enrichment Analysis.

were calculated between disease and control groups with it. For PCOS, the DEG threshold was set as P value <0.05 and $|\log_2\text{FC}(\text{fold change})| >0.585$. For RIF, the P value adjusted to 0.05 and $|\log_2\text{FC}| >1$ were used to identify the DEGs. Next, the difference analysis results for each group were presented using the heatmap and volcano plot. In both plots, blue indicated low expression, and red indicated high.

2.3 Weighted gene co-expression network analysis (WGCNA)

Using microarray specimens, WGCNA represents one of the most important and widely applied systems bioinformatics methods to describe the correlation patterns among genes. Genes can be grouped into modules based on their co-expression similarities across samples using the “WGCNA” R package. Additionally, the WGCNA method can be used to connect modules to clinical elements outside the genome. In this way, relevant functional networks can be used to identify biomarkers and new molecules. As input files, normalized mRNA expression data (calculated using the R package “LIMMA”) were used to perform WGCNA to identify gene coexpression and the correlation between gene modules and clinical characteristics (PCOS or RIF compared to control groups). For each disease group, the following steps are followed (1): by using the R package “gplots,” hierarchical clustering analysis was performed to identify outliers in the sample (2), the “pickSoftThreshold” package function was utilized to screen out soft-power parameters ranging from 1 to 20 (3), a topological overlap matrix (TOM) is created by converting the matrix of

correlations with the most appropriate β value to an adjacency matrix and then into a topological overlap matrix (4), based on the average linkage hierarchical clustering, an hierarchical clustering tree (linked gene best fit) was constructed, and then the dynamic tree cut algorithm ($\text{minModuleSize} = 30$) was used to find different gene modules. Similar modules were merged by a cutheight in each group, and (5) gene modules and clinical phenotypes (CTRL and PCOS or RIF) were correlated using the Pearson correlation coefficient. For PCOS, the highest correlation module antiquewhite4 was selected, and the genes associated with it were further analyzed. However, many modules were related strongly to RIF, so we further filtered the genes according to the gene significance (GS) and modular membership (MM). The disease was most closely related to genes with a high MM and GS, and we selected the genes in the hub modules with $|\text{MM}| > 0.8$ and $|\text{GS}| > 0.5$ for the RIF group.

2.4 Identification of shared genes and functional enrichment analysis

By combining DEGs and module genes identified by WGCNA, we were able to identify the shared key genes that contributed to the pathogenesis of both PCOS and RIF. For identifying genes’ biological functions and signaling pathways, we used the “clusterProfiler” package for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. The bar plots showed the significant enrichment results of the functional enrichment with $P < 0.05$.

2.5 Feature selection by three well-established machine learning algorithms

Further screening of co-existing genes between the two diseases was carried out using three well-established machine learning algorithms (LASSO: Least Absolute Shrinkage and Selection Operator; SVM-RFE: Support Vector Machine- Recursive Feature Elimination; RF: Random Forest). To ensure the repeatability of these algorithms, we set the seed at 123 in both disease groups.

Firstly, the 12 shared genes obtained previously were input into the LASSO algorithm in each disease group. We constructed a regression model using the R package “glmnet” with 10-fold cross-validation. In the “family” parameter, we set “binomial,” and we chose the best lambda value by “lambda.1min”. Logarithm (lambda) profiles of the LASSO coefficients were drawn for the 12 features. Then we drew the partial likelihood deviation (binomial deviation) curve and the logarithm (λ) curve. Next, we calculated the best value for 1se (1-SE standard) of minimum standard.

Following that, SVM-RFE was used to eliminate recursive features. Using the “e1071” and “MSVM-RFE” package for SVM modeling, SVM-RFE applied sequential backward feature elimination to determine the optimal hub gene. All of the 12 shared genes were used in our SVM model. The result of SVM-RFE was visualized, and by tenfold cross-validation, the red circle indicated maximum classification precision and the corresponding gene sets were the most accurate diagnostic markers at the lowest 5×CV error and the highest 5×CV accuracy.

Finally, we used Random Forest to classify the significant genes with the R package “randomForest”. Using a decision tree algorithm, random forest analysis identified which variables were most important. As a result of this algorithm, we were able to filter the shared genes to find disease signature genes. Our first step was to construct a random forest model using 500 trees on the discovery cohorts and determine the optimal number of trees using cross-validation errors. Then, we ranked genes by importance and plotted the 10 most significant genes. For each disease group, the significance threshold was set at 0.9 to decide the final result.

After screening the above three algorithms, we took the intersection of the results of each algorithm. There were five genes common to the PCOS group and seven common genes in the RIF group shown in a Venn plot. Furthermore, we took the intersection of the above common genes again and finally obtained two genes as the disease diagnostic target genes. The “pROC” package was used to construct ROC curves and displayed using “ggplot2” to assess the accuracy of the two diagnostic genes in discovery cohorts.

2.6 Prediction performance in validation cohorts

To further test the accuracy of the two diagnostic genes, we searched the GSE80432 for PCOS and GSE26787 of RIF for external validation. We first downloaded the raw data of these two data sets from the GEO database and normalized them with the “RMA” package. For GSE80432, we filtered out a deviation sample through

PCA and finally adopted four healthy control samples and three PCOS patients. For GSE26787, five healthy controls and five patients with RIF were used for validation. The diagnostic gene expression pattern in validation cohorts was shown by boxplot, and AUC (area under the ROC curve) was also calculated.

2.7 Implementation of GSEA for single diagnostic gene

After obtaining the diagnostic genes, we performed single-gene gene set enrichment analysis (GSEA) for each diagnostic gene in the two groups using the “clusterProfiler” package. Using GSEA, we compared the biological signaling pathways between the disease group and the healthy control group. MSigDB (c5.go.bp.v7.5.1.entrez.gmt) was used to download gene sets. Enrichplot was used to show the top 5 activating and inhibiting pathways for each gene in the two disease groups.

2.8 Immune cell abundance

Each sample of disease was subjected to CIBERSORT analysis to determine the relative levels of immune cells. The CIBERSORT algorithm resolves immune cell composition by deconvolution based on gene expression data. According to CIBERSORT’s website (<http://cibersort.stanford.edu/>), LM22 contains 22 annotated gene signatures. With CIBERSORT and 1000 iterations, we quantified 22 types of immune cells based on the LM22 gene signature. In the following analysis, we selected samples with a CIBERSORT *P* value less than 0.05 for analysis. For each sample, the output estimates from CIBERSORT were normalized to sum to one to facilitate comparisons across immune cell types and datasets. R packages “corplot,” “vioplot” and “ggplot2” were used to visualize the results. Further, the correlation between immune infiltrated cells and diagnostic target biomarkers was determined by nonparametric correlations (Spearman).

2.9 Human samples collection

Granulosa cells of PCOS and the control patients were collected from infertile women who underwent ART treatment in the Reproductive Medicine Center of the First Affiliated Hospital of Zhengzhou University in China. Inclusion criteria for PCOS patients were based on the 2003 Rotterdam Criteria (27). Mid-secretory endometrial samples from the RIF and control groups were also collected to verify our data analysis. This study was approved by the Ethics Committee of the First Affiliated Hospital of Zhengzhou University in China. RNA of these samples was extracted using TRIzol Reagent (InvitrogenTM, Japan) and cDNA was synthesized using HiScript III RT SuperMix for qPCR (Vazyme, Nanjing, China). Then real-time PCR (RT-PCR) was performed to quantitative the expression level of *GLIPR1* and *MAMLD1* in the two diseases. Primer sequences of the two genes were described in [Supplementary Table 1](#). Statistical analysis was performed to compare the groups, using Student’s *t*-test. The data were presented as mean

values accompanied by the standard error of the mean (SEM). A significance level of $P < 0.05$ was considered a significant difference.

3 Results

3.1 GEO information

A total of four data sets were selected for discovery analysis according to our inclusion criteria: GSE10946, GSE34526, GSE103465, and GSE111974. Table 1 summarized the detailed information on the four datasets. Among them, GSE10946 and GSE34526 were used as discovery cohorts for the PCOS, and GSE103465 and GSE111974 were regarded as discovery cohorts of RIF. Besides, GSE80432 and GSE26787 were the validation cohort of PCOS and RIF, respectively. In the present study, we called these two disease groups PCOS and RIF for short.

3.2 Identification of DEGs.

Before the biological information analysis, we tested the batch effects of the collected datasets and found that the batch effects of the two diseases were apparent (Figure 2A, E). Using the “sva” package, we removed the batch effects of the PCOS (Figure 2B) and RIF groups (Figure 2F) to obtain reliable analysis results. A LIMMA package was then used to characterize the DEGs between the two groups. 201 DEGs ($P < 0.05$, $|\log_2 FC| > 0.585$) with 101 up-regulated and 100 down-regulated genes were obtained for PCOS. For RIF, there were 253 DEGs (adj. $P < 0.05$, $|\log_2 FC| > 1$) with 167 up-regulated and 86 down-regulated genes. Volcano plots showed all DEGs of PCOS (Figure 2C) and RIF (Figure 2G) groups. Taken as a whole, the DEGs contained in the two groups were visualized by heatmaps (Figure 2D, H). DEGs associated with PCOS and RIF might play a role in their occurrence and development.

3.3 Screening for key modules by WGCNA.

To investigate whether the diseases and key genes are correlated, we performed WGCNA in addition to analyzing the

differential expression between the two groups. Using the soft-thresholding approach, this study constructed a co-expression network. This parameter β was essential for co-expression networks to maintain a scale-free topology. Gene expression data-based biological networks were most likely to be scale-free. Accordingly, in the PCOS group, the fit index greater than 0.85 was considered scale-free topology, and β was set at 9 (Figure 3A). By using the adjacency function, the adjacency matrix was generated. As shown in Figure 3B, hierarchical clustering was constructed using the TOM dissimilarity measure. We have identified 20 co-expression modules in total. The modules that $P < 0.05$ were regarded as key modules. As shown in Figure 3C, the antiquewhite4 module had the strongest positive correlation, which contained 180 genes. Also, WGCNA was applied to the RIF group, and $\beta = 10$ was the optimal value for soft power (Figure 3D). We identified 19 modules in total, in which dark grey, dark green, and royal blue showed a strong positive correlation, and green-yellow, salmon, dark turquoise, and light yellow modules showed a strong negative correlation (Figures 3E, F). Among the genes in these 7 key modules in the RIF group, we further selected 334 genes with $|MM| > 0.8$ and $|GS| > 0.5$. These genes in key modules derived from the two groups could be potentially used as candidate cell-type-specific markers.

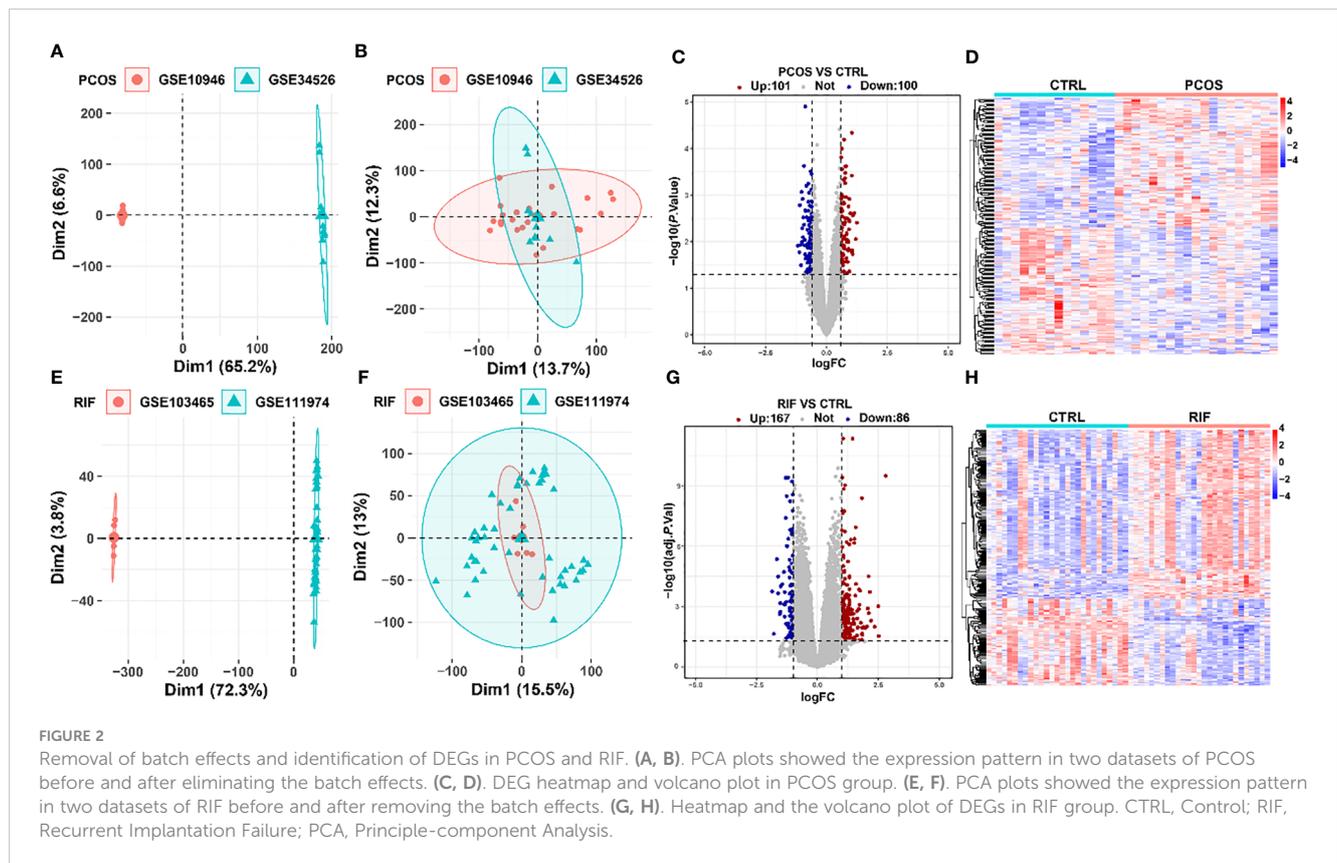
3.4 Analysis of the shared genes and functional enrichment

To explore the co-pathogenesis of PCOS and RIF, we took the intersection of the DEGs mentioned above and genes screened by WGCNA, respectively. Figure 4A showed an overlap between the DEGs of PCOS and RIF as a total of 11 genes (CHST11, FAM150B, GLIPR1, SLC16A6, MAMLD1, SLC46A2, ENPP3, HAPLN1, PLCXD3, FAM110C, GAS1). There was only one gene that overlapped the genes of WGCNA analysis (CCND2, Figure 4B). We speculated that these 12 genes might be related to the pathogenesis of PCOS and RIF and had a shared relationship (Figure 4C). Analyzing these genes for functional annotation and enrichment (Figures 4D, E), we sought to investigate the potential biological changes between PCOS and RIF. Not surprisingly, GO analysis of the shared genes revealed they were overrepresented in pathways associated with early embryonic

TABLE 1 Details of GEO datasets used in the study.

| Diseases | GEO Series | GPL Platform | Control Case | Sample Size | Group |
|----------|------------|--------------|--------------|-------------|-------------------|
| PCOS | GSE10946 | GPL570 | 11 | 12 | Discovery cohort |
| | GSE34526 | GPL570 | 3 | 7 | Discovery cohort |
| | GSE80432 | GPL6244 | 4 | 4 | Validation cohort |
| RIF | GSE103465 | GPL16043 | 3 | 6 | Discovery cohort |
| | GSE111974 | GPL17077 | 24 | 24 | Discovery cohort |
| | GSE26787 | GPL570 | 5 | 5 | Validation cohort |

PCOS: Polycystic Ovarian Syndrome, RIF: Recurrent Implantation Failure, GEO: Gene Expression Omnibus.



organ morphogenesis. Notably, among all the GO terms enriched, we noticed that several significantly enriched biological processes pathways, such as regulation of T cell apoptotic process, mast cell activation, and leukocyte homeostasis, suggesting that the activation and apoptosis of immune cells might make a considerable contribution to the co-pathogenesis of PCOS and RIF. Besides, KEGG enrichment was consistent with the GO analysis. In addition, we noted that several pathways related to follicular development are enriched, such as the p53 signaling pathway, FOXO signaling pathway, hippo signaling pathway, and PI3K-Akt signaling pathway.

3.5 Identify potential shared diagnostic genes based on machine learning algorithms.

For a further selection of the most candidate diagnostic gene targets with a significantly characteristic value of classifying the disease groups and control groups, three different algorithms (LASSO, SVM-RFE, and Random Forest) were applied based on the above 12 shared genes. In the PCOS group, based on the LASSO coefficient profiles and the optimal tuning parameter selection map, λ was set at 0.06851194 (PCOS) (Figure 5A). Afterward, eight genes with non-zero coefficients were found. Then we inputted the above 12 genes into the RF classifier, and the top 10 genes were shown on the importance scale. We selected 0.9 as the screening threshold of importance, and a set of 9 genes was identified (Figure 5B). Also, the SVM algorithm identified 5 genes with the lowest 5-point CV error and best 5-point CV accuracy (Figure 5C).

Furthermore, by overlapping these three algorithms, we developed 5 shared biomarkers (CHST11, GLIPR1, SLC16A6, MAMLD1, HAPLN1, GAS1) for the PCOS group (Figures 5D, I).

Similarly, 7 featured genes were obtained for the RIF group when the λ was set at 0.03001025 by the LASSO algorithm (Figure 5E). Figure 5F showed the top 10 genes on the importance scale and we chose 9 genes (importance>0.9) as the RM result. Next, a subset of 10 hub genes were identified using the SVM-REF algorithm (Figure 5G). Then, 7 common gene biomarkers obtained by the three algorithms overlapped (Figures 5H, I).

3.6 Diagnostic value and validation of diagnostic hub biomarkers.

For a more precise understanding of the relationship between PCOS and RIF, we took the intersection of machine learning results in the PCOS and RIF group and got 2 shared diagnostic genes, GLIPR1 and MAMLD1 (Figure 6A). And the prediction and the discriminatory ability of the shared diagnostic genes were assessed by analyzing the expression pattern of the two genes. Also, an analysis of the receiver operating characteristic curves (ROC curves) was conducted.

Firstly, we analyzed the PCOS and RIF expression levels of the two discovery cohorts. Figure 6B showed that *GLIPR1* was lower in the RIF groups ($P<0.01$) and higher in the PCOS groups ($P<0.01$). *MAMLD1* expressed lower both in the PCOS ($P<0.01$) and RIF groups ($P<0.0001$) (Figure 6F).

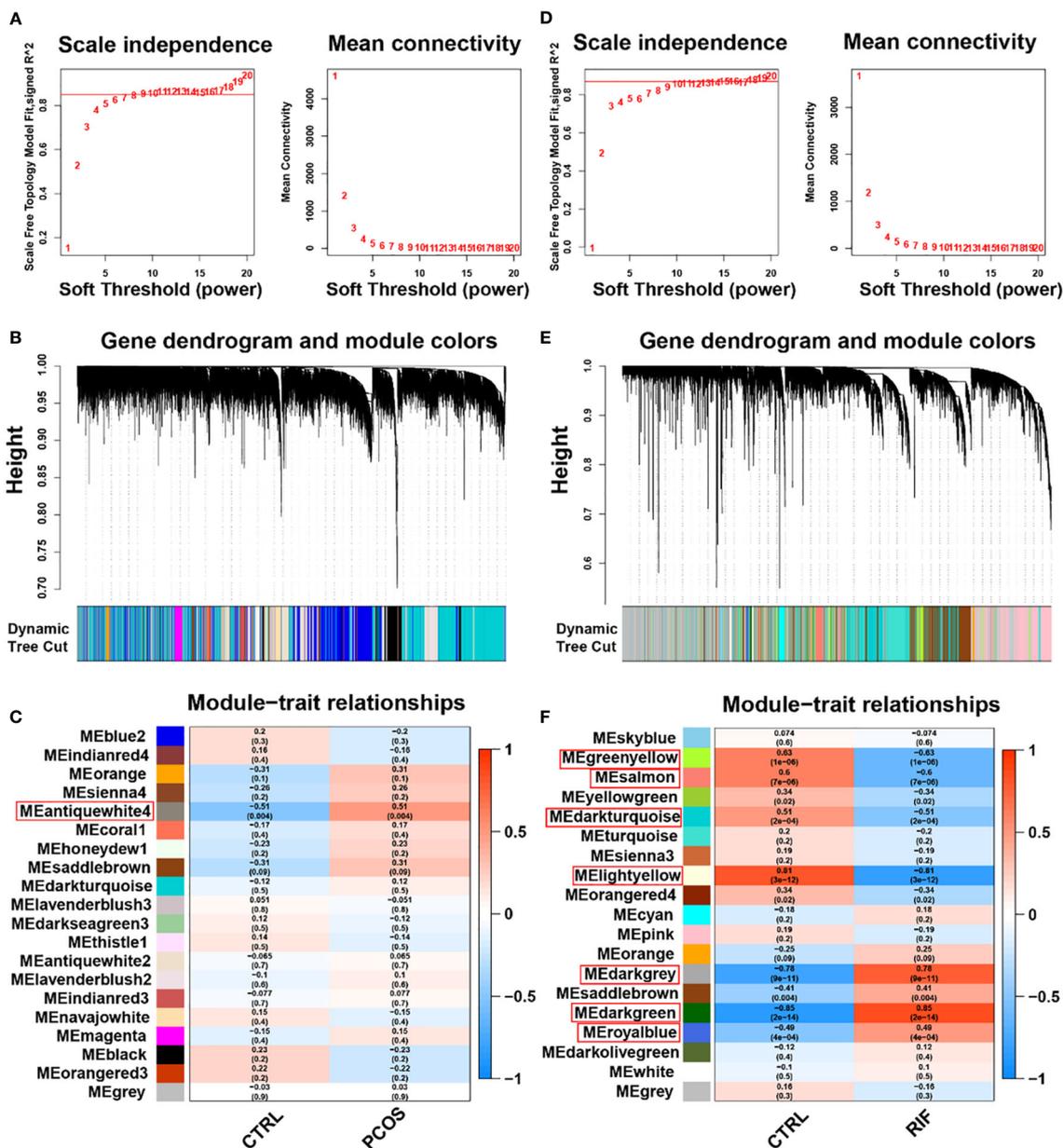
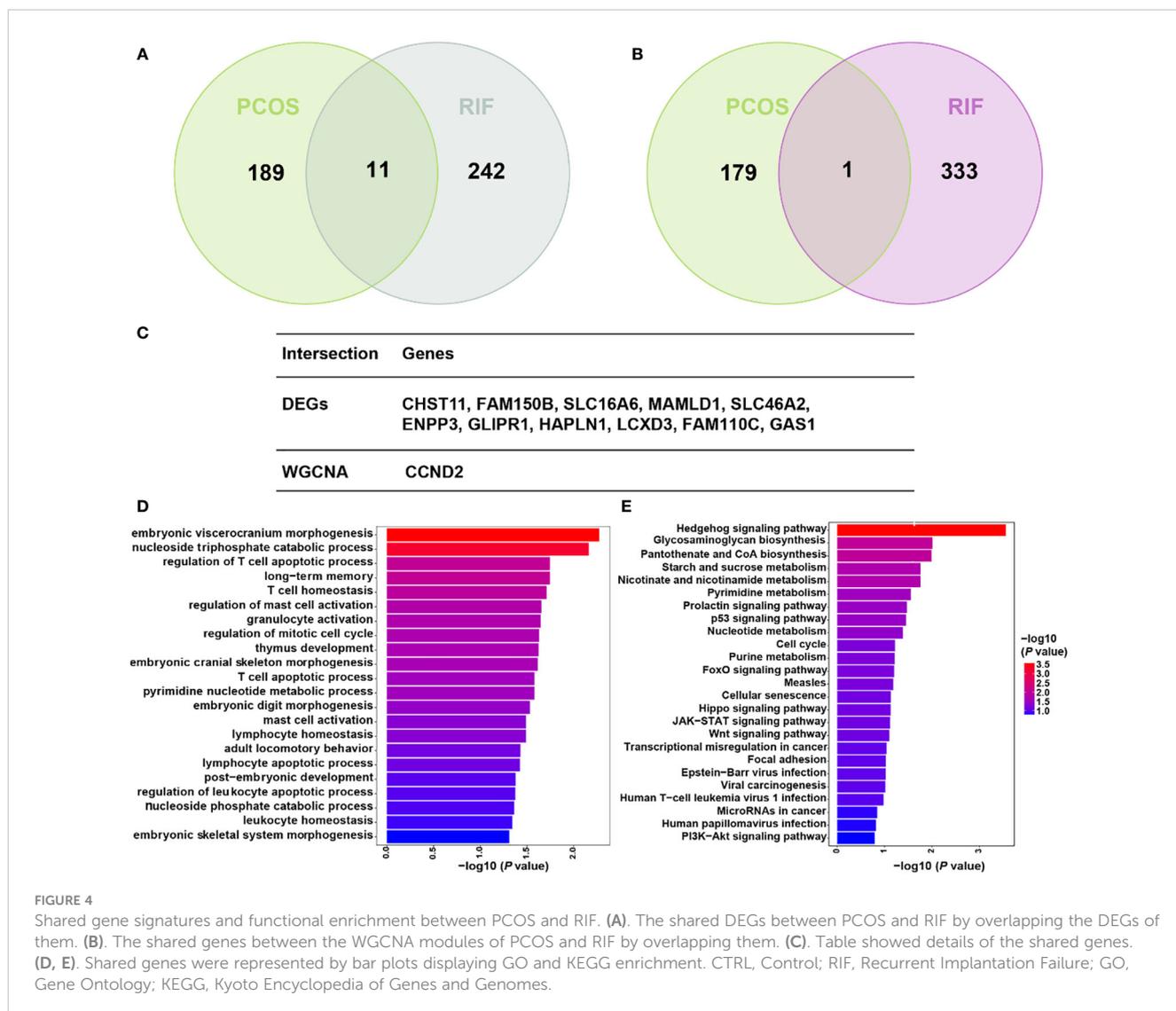


FIGURE 3 Weighted gene co-expression network analysis (WGCNA) of PCOS and RIF. (A). Determination of soft-threshold power for PCOS. (B). Cluster dendrogram of PCOS highly connected genes in key modules. (C). Relationships between modules and traits in PCOS. Correlations and *P* values are included in each cell. (D). Calculation of soft-threshold power for RIF. (E). Cluster dendrogram of RIF modules with highly connected genes. (F). Module-trait relationships in RIF. A correlation and *P* value are included in each cell. CTRL, Control; RIF, Recurrent Implantation Failure.

Next, to test the specificity and sensitivity of the two target genes for the diagnosis of two diseases, ROC analysis was applied. In terms of PCOS biomarkers, these two genes had favorable results: *GLIPR1* (AUC=0.812), and *MAMLD1* (AUC=0.774). The RIF group was also subjected to the same ROC analysis. Predictive performance was robust for each biomarker: *GLIPR1* (AUC=0.723), and *MAMLD1* (AUC=0.879) (Figures 6C, G).

Moreover, we confirmed the reliability of *GLIPR1* and *MAMLD1* as core diagnostic genes for PCOS and RIF by conducting external validation. In the two validation groups, the

expression levels of the two hub genes matched those in the discovery cohorts. *GLIPR1* was decreased ($P < 0.05$) in the RIF group and increased in the PCOS groups ($P < 0.05$) (Figure 6D). *MAMLD1* was reduced both in the PCOS groups ($P < 0.01$) and the RIF groups ($P < 0.05$) (Figure 6H). Figure 6E showed that *GLIPR1* had excellent diagnostic accuracy in the validation cohort of PCOS (AUC=1.000) and RIF (AUC=0.920). Similarly, *MAMLD1* also properly diagnosed PCOS (AUC=1.000) and RIF (AUC=0.840) (Figure 6I). As a result, the results confirmed their ability to serve as key discriminatory molecules for PCOS and RIF, respectively.



3.7 Single-Gene GSEA of diagnostic genes.

Subsequently, we employed single-gene GSEA analysis of the two biomarkers in PCOS and RIF datasets, respectively, and the top 5 up and down-regulated pathways were visualized by the “GSEA” package. Figure 7 showed that in both disease groups, these two genes were both involved in metabolic pathways such as glycine, serine, threonine metabolism, alpha-linolenic acid metabolism, and propanoate metabolism. Besides, both genes enriched in inflammation-related pathways which linked the PCOS and RIF.

3.8 Immune infiltration analysis of shared diagnostic genes.

Considering that PCOS and RIF are characterized by a high immune response. The abundances of immune cells in different groups were analyzed with CIBERSORT. In each group, the proportion of 22 immune cells was shown as a bar plot.

Generally, the bar graphs clearly illustrate the significant differences between the percentages of T cells, macrophages, and NK cell populations between PCOS (Figure 8A) and RIF (Figure 8E). Compared with the control samples, the dendritic cells activated were increased in the PCOS samples (Figure 8B). While, in the RIF samples, NK cells resting, Macrophages M0 were increased, and NK cells gamma delta, dendritic cells activated were decreased (Figure 8F).

Moreover, the relationship between biomarkers and immune cell contents was investigated. In PCOS samples, CD8 T cells were significantly positively correlated with GLIPR1 (Figure 8C). In contrast, macrophage M0 correlated negatively. MAMLD1 was significantly positively correlated with B cells memory and negative for macrophages M2, dendritic cells resting, and monocytes (Figure 8D). In RIF samples, GLIPR1 had a significant negative correlation with NK cells resting (Figure 8G). While Neutrophils and NK cells rested negatively with MAMLD1 (Figure 8H). It appears that immune function is crucial to the development of PCOS and RIF.

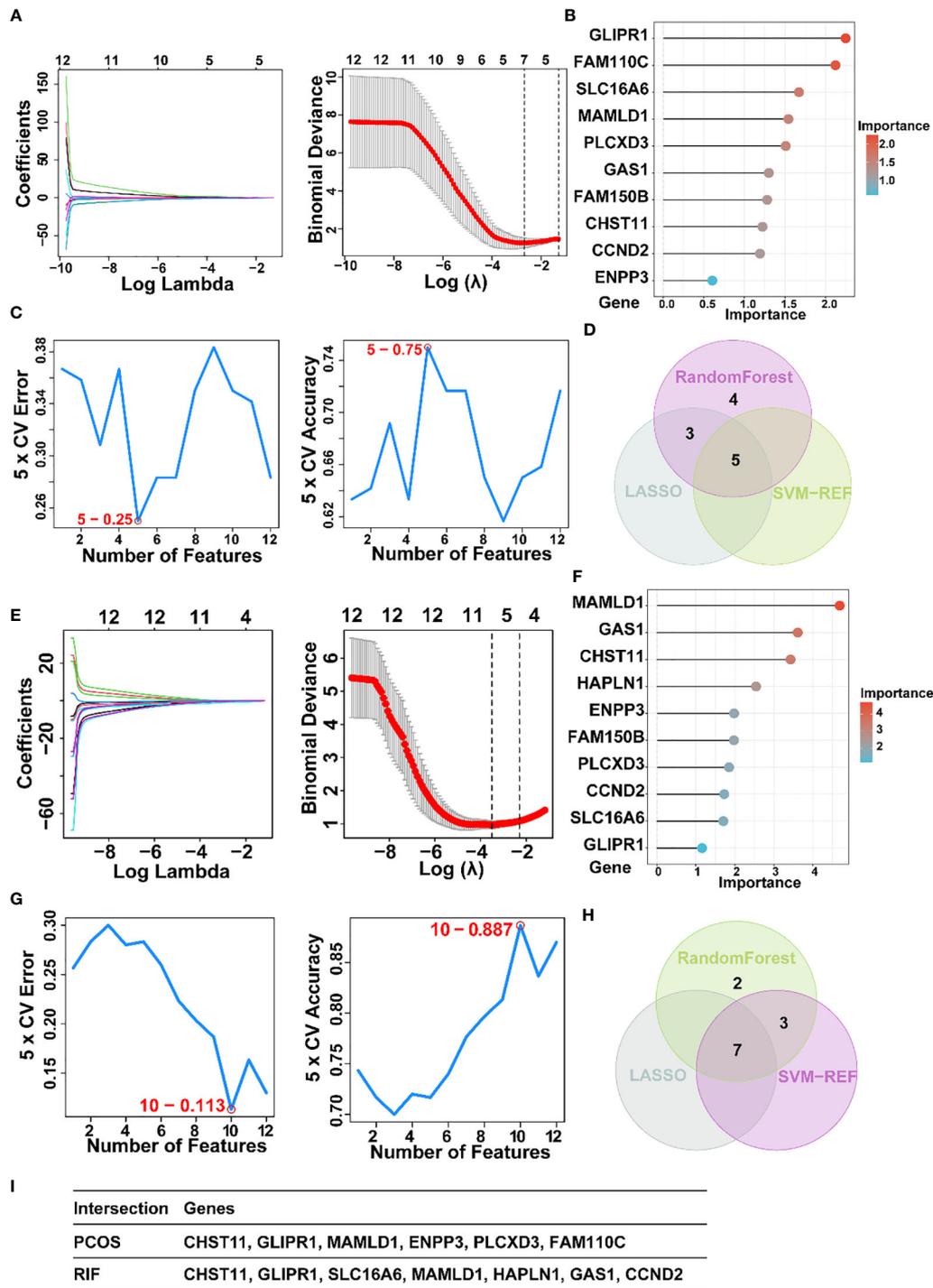


FIGURE 5

The screening of candidate PCOS and RIF diagnostic genes using three machine learning algorithms. (A). Coefficient profile plot of the LASSO model for PCOS showed the final parameter selection λ (lambda). (B). PCOS top-10 genes according to their discriminant ability in the RF algorithm. (C). Five crosstalk genes were selected by using the SVM-RFE algorithm for PCOS. (D). The Venn diagram showed five candidate diagnostic genes in PCOS by intersecting the results of three algorithms. (E). Coefficient RIF profile plot of the LASSO model showed the selection of the optimal parameter λ (lambda). (F). Top-10 RF algorithm step discriminant ability genes for RIF. (G). Ten crosstalk genes were selected by using the SVM-RFE algorithm for RIF. (H). The Venn diagram showed seven candidate diagnostic genes in RIF by intersecting the results of three algorithms (I). Table showed the details of candidate diagnostic genes in PCOS and RIF. CTRL, Control; RIF, Recurrent implantation failure; LASSO, Least Absolute Shrinkage and Selection Operator; SVM-RFE, Support Vector Machine-Recursive Feature Elimination; RF, Random Forest.

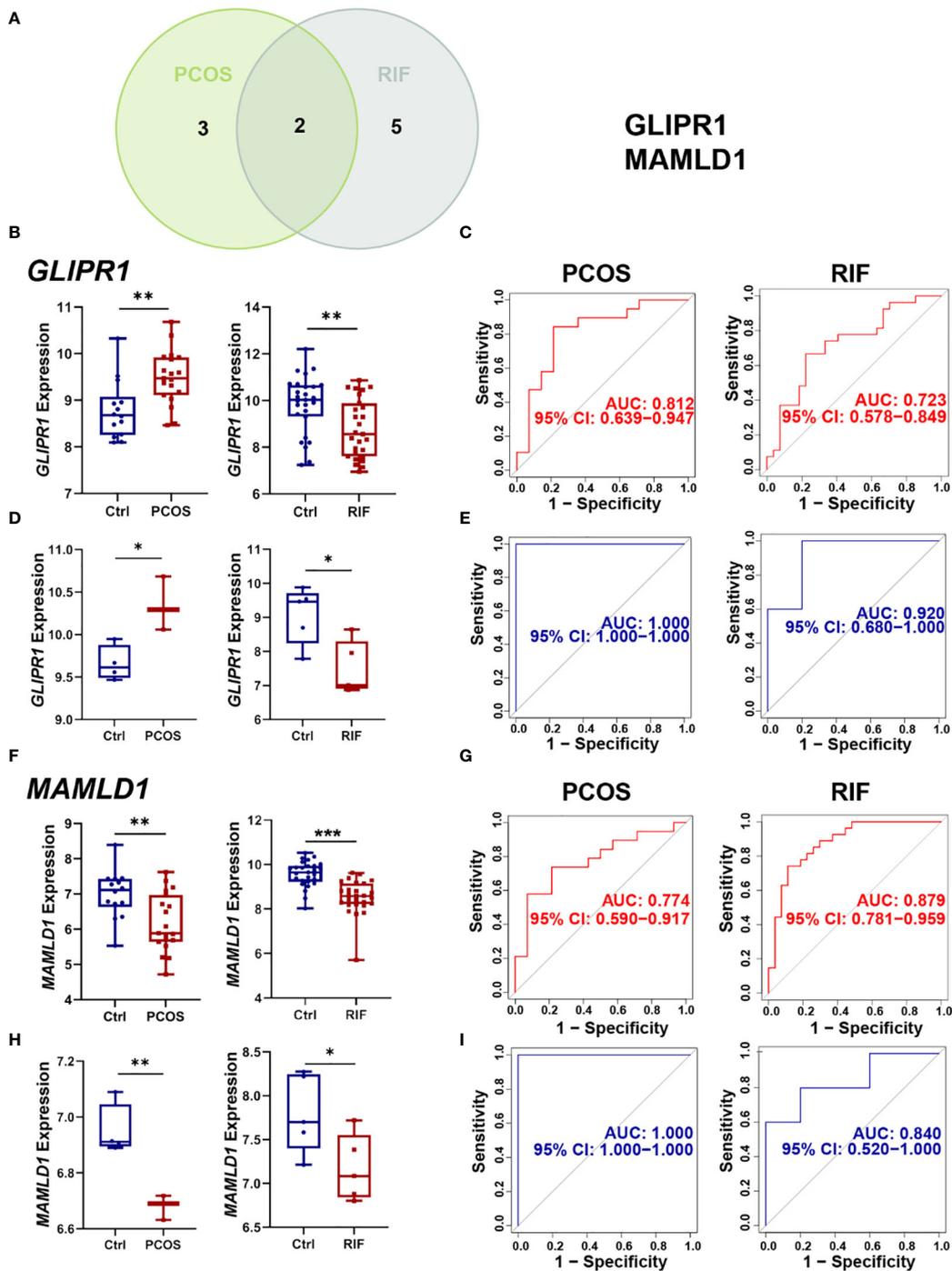


FIGURE 6

Selection and validation of the two shared diagnostic genes. (A). The Venn plot showed the two shared diagnostic genes. (B). Differential expression of GLIPR1 in the training group for PCOS and RIF. (C). ROC curve of GLIPR1 in the training group for PCOS and RIF. (D). Differential expression of GLIPR1 in the validation group for PCOS and RIF. (E). ROC curve of GLIPR1 in the validation group for PCOS and RIF. (F). Differential expression of MAMLD1 in the training group for PCOS and RIF. (G). ROC curve of MAMLD1 in the training group for PCOS and RIF. (H). Differential expression of MAMLD1 in the validation group for PCOS and RIF. (I). ROC curve of MAMLD1 in the validation group for PCOS and RIF. ROC, receiver operating characteristic. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

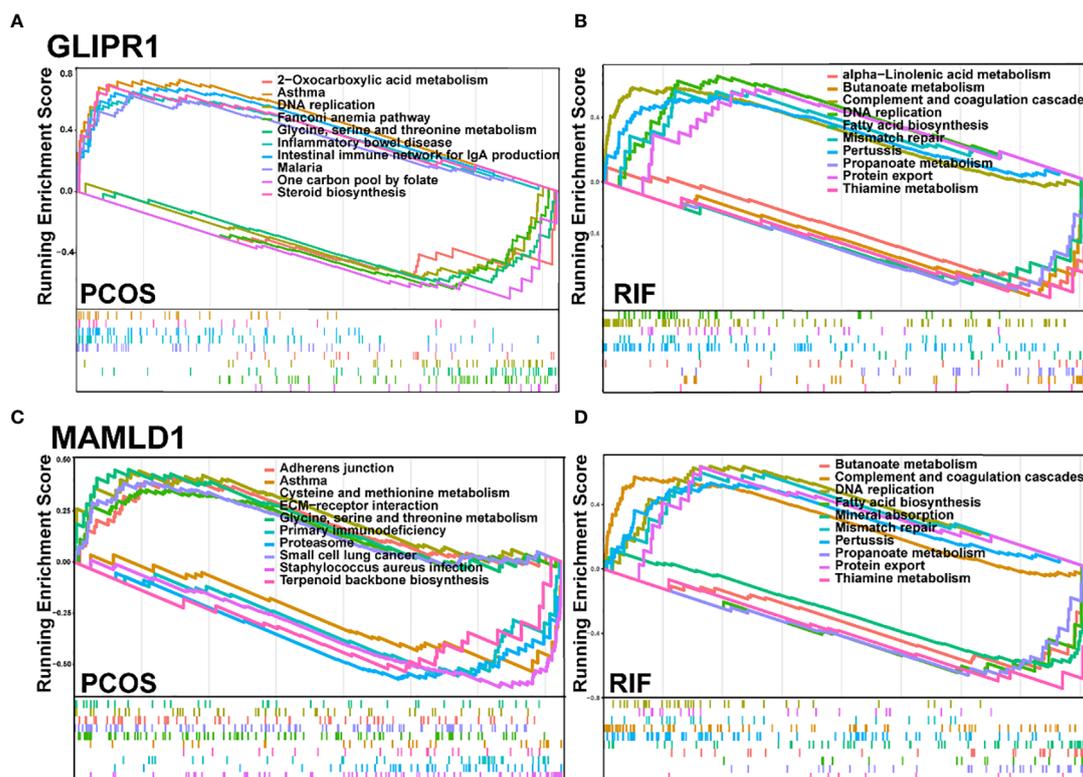


FIGURE 7

GSEA for the single diagnostic gene. (A, B). GSEA analysis for *GLIPR1* in PCOS and RIF group. (C, D). GSEA analysis for *MAMLD1* in PCOS and RIF group. GSEA, Gene Set Enrichment Analysis.

3.9 Validation of *GLIPR1* and *MAMLD1* by RT-PCR in human tissues.

RT-PCR was performed on follicular fluid-derived granulosa cells from normal women and PCOS patients and endometrial tissues from healthy and RIF women. This confirmed the gene expression levels of the two diagnostic biomarkers, *GLIPR1* and *MAMLD1*. Consistent with the data analysis, our results showed that *GLIPR1* expression was upregulated and *MAMLD1* expression was decreased in the granulosa cells of PCOS patients, (Figure 9A) while both *GLIPR1* and *MAMLD1* expression was reduced in endometrial tissues of RIF patients (Figure 9B).

4 Discussion

Due to the high occurrence of recurrent implantation failure for PCOS patients in ART cycles (8, 25), it has become essential to explore the common pathogenesis of the two diseases. In the present study, we applied the WGCNA (28) and three machine learning approaches to identify the common gene population and 2 diagnostic genes of *GLIPR1* and *MAMLD1* which were both significantly related to PCOS and RIF. Further, by GSEA (29) analysis, we uncovered that the co-pathogenesis of the two diseases lay in the abnormal metabolism of many metabolites associated with the TCA cycle, which led to the abnormal

activation of immune cells and immune response in the disease group. Taken together, the newly discovered diagnostic genes and potential molecular mechanism in this study provided new clinical insights and guidance for diagnosing and treating PCOS and RIF patients.

GLI pathogenesis-related 1 (*GLIPR1*) is a gene that encodes a protein with diverse biological functions, including cell apoptosis, cell cycle regulation, and DNA damage response. While the direct relationship between *GLIPR1* and PCOS and RIF has not been fully established, several studies have suggested a potential involvement of *GLIPR1* in the pathogenesis of these diseases. Notably, altered expression levels of *GLIPR1* have been observed in PCOS patients, indicating a possible association between *GLIPR1* and PCOS (30). Studies have also identified that *GLIPR1* encodes proteins that regulate sperm-oocyte binding and mature male germ cells, which have been linked to PCOS risk modification and metabolic mechanisms (31). Moreover, RNA sequencing of adipose tissue in PCOS patients has identified *GLIPR1* to be differentially expressed according to genotype near PCOS risk loci (32), suggesting a possible relationship between insulin resistance and fertility. Single-cell mRNA sequencing of proliferative phase endometrial cells also identified *GLIPR1* as having an important role in the perivascular environment (33). Additionally, *GLIPR1* has been implicated in RIF. A study reported differential expression of *GLIPR1* in RIF patients compared to those with successful pregnancies, suggesting its involvement in the immune response

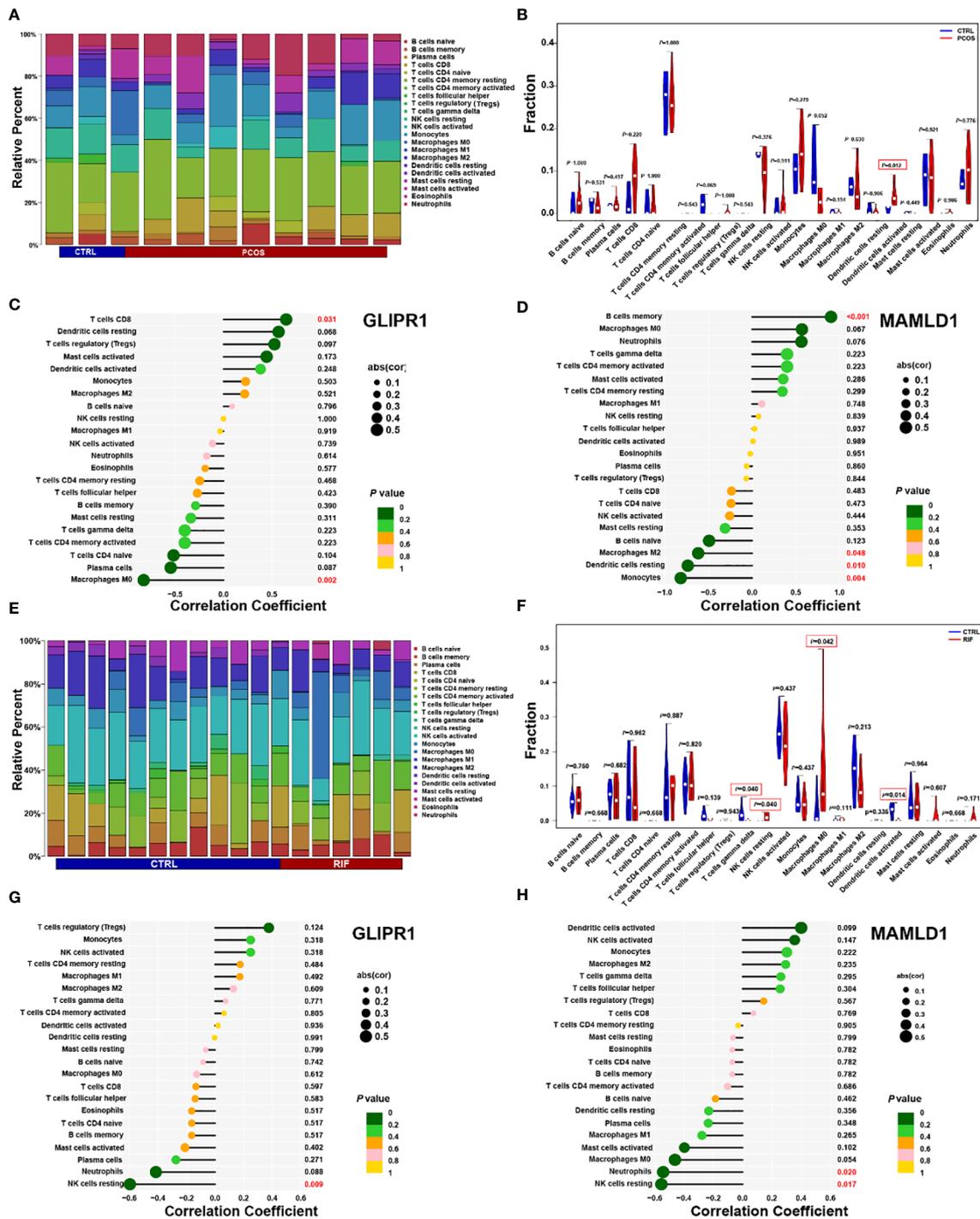
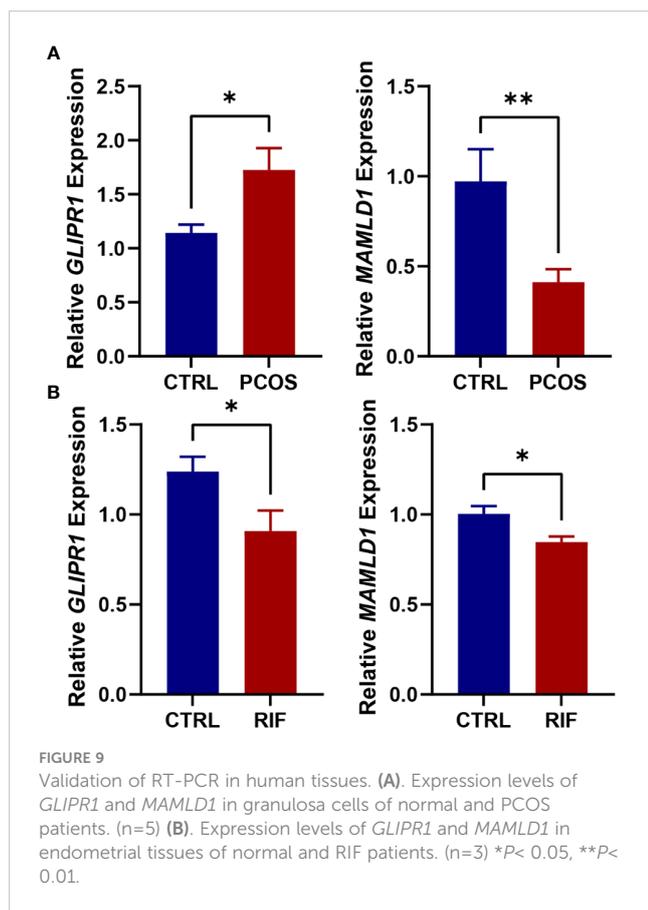


FIGURE 8

PCOS and RIF immune cell composition. (A). Infiltrating immune cells were plotted in the stacked bar chart for the PCOS group. (B). Violin diagram indicated that the PCOS group exhibited a significantly different type of immune cell. (C). Correlation between GLIPR1 expression and immune cells in the PCOS group. (D). Relationship between the expression of MAMLD1 and immunity in PCOS patients. (E). Stacked bar chart showed the RIF group's characteristics of infiltrating immune cells. (F). Evident difference in immune cell types shown by violin diagram of the RIF group. (G). Detection of GLIPR1 expression in immune cells in the RIF group. (H). Relationship between MAMLD1 expression and immune cells in the RIF group. $P < 0.05$ was highlighted.

associated with embryo implantation (34). Our study suggested that GLIPR1 represented a potential diagnostic target for RIF in PCOS patients which could be supported by previous research partly.

MAMLD1 (NM_001177465) is a gene associated with disorders of sexual development in 46, XY individuals, and it has been reported that mutations in MAMLD1 resulted in male fetal sexual



development (35) (36, 37). Furthermore, *MAMLD1* hypomethylation and upregulation of its transcript in granulosa cells have been suggested as contributing to PCOS ovaries with excess androgen and hormone imbalance (38). Many studies have proved that the excess of maternal androgen leads to abnormal placental morphogenesis and impaired endometrial receptivity (39), resulting in granulosa cell apoptosis and inhibition of proliferation. Consequently, decreased ER ability and maternal-fetal interface eventually led to embryo implantation failure. Additionally, knock-out mice models have shown that lack of *MAMLD1* causes parturition failure, high neonatal mortality rates, and an adverse effect on functional luteolysis (40). Consistent with previous studies, our analysis revealed that *MAMLD1* appeared to play a role in embryo development and was associated with recurrent implantation failure in women with PCOS.

To further explore the underlying pathogenic association and mechanisms between PCOS and RIF, we performed GSEA for the two diagnostic genes in two disease groups. As a result of GSEA analysis, both PCOS and RIF genes were enriched in metabolism pathways. In the PCOS group, both two genes were enriched in glycine, serine, and threonine metabolism. Similarly, *GLIPR1* and *MAMLD1* were found to involve in the metabolism of butanoate, propanoate, and thiamine in the RIF group. Interestingly, metabolites like glycine (41), serine (42), and threonine (43) have been reported to contribute to increasing protein motive force, which is attributed to the activation of the TCA cycle and

mitochondrial respiration chain. A previous study found that *GLIPR1* influenced the energetic potential of mitochondria which was in accordance with these enrichment analyses (44). Specifically, butyrate has been reported to provide energy substrates for the host, which enter into the TCA cycle as acetyl-CoA to produce glucose (45). Propanoate can also increase the level of mitochondrial CoA and promote energy production through the TCA cycle (46). Known as vitamin B1, thiamine was used in the TCA cycle to form thiamine diphosphate (ThDP) (47). These metabolism changes all suggest us metabolites involved in the TCA cycle are disrupted in patients with PCOS complicated by repeated implantation failure and finally lead to the imbalance of energy metabolic homeostasis.

Consistent with previous studies (48, 49), our data showed that the common gene population mainly enriched in the activation pathways of immune cells like T cells, NK cells, macrophages, and neutrophils. These results hinted us the high level of immunity in PCOS and RIF patients and prompted us to further investigate which immune cells are abnormal in the two diseases. As anticipated, by immune infiltration (50), we detected Dendritic cells (DCs) activated in PCOS samples and four immune cell types in RIF samples, including T cells gamma delta, NK cells resting, macrophages M0 and Dendritic cells activated with significantly different abundance compared to healthy samples. The obesity status of PCOS patients will lead to changes in the number of DCs, and high androgen and low progesterone status will reduce the ability to recruit NK cells (4, 51). So that the secretion of cytokines is not enough to maintain maternal immune tolerance, and eventually leads to endometrial receptivity impairment, abortion, implantation failure, and other adverse events (52). While normal human pregnancy requires complex coordination between maternal immune tolerance and homeostasis (53). Besides, we uncovered that the prevalence of DCs in each disease was significantly different from the normal group. This suggests that the abnormal function of NCs may be the main cause of repeated implantation failure in PCOS patients.

Based on the above findings, we hypothesize that activation of a large number of immune cells in cellular immunity caused by disturbances of metabolites associated with the TCA cycle accounts for the co-pathogenesis of the PCOS and RIF. As previously mentioned, activated DCs and macrophages showed a truncated TCA cycle that resulted in an accumulation of citrate. Of note, the TCA cycle intermediates may leak from impaired mitochondria, and increasing evidence indicated that these metabolites of the TCA cycle played a substantial role in immune regulation (54). In PCOS patients with coexisting metabolic disorders, disruptions to their mitochondrial membranes may result in the release of TCA cycle intermediates into the cytosol, thereby impairing the cellular immune system. The dysregulated immune system is characterized by a large number of activated immune cells, particularly DCs, as observed in the present study. These processes ultimately contribute to the failure of embryo implantation, which is a hallmark feature of RIF.

Some limitations need to be mentioned in the present study. First, only 3 datasets were selected for each disease, including 18

cases and 14 control samples in the PCOS group, and 27 cases in both the RIF disease and normal groups for discovery. For WGCNA and immune infiltration analysis, more sufficient samples are needed to ensure test accuracy. Second, we know little about the processing of the raw data, and the datasets used in our study were profiled by microarray which was far behind today's advanced sequencing technology. These intrinsic traits of these public datasets pose a challenge to the accuracy and advancement of our analysis.

In conclusion, diagnostic biomarkers GLIPR1 and MAMLD1 were identified as critical biomarkers responsible for regulating immune cell activation caused by an imbalance of TCA cycle metabolites. Our analysis reinforces the theoretical basis for the co-pathogenesis of recurrent pregnancy failure in PCOS patients.

Data availability statement

The original contributions presented in the study are publicly available. Publicly available datasets used in this study can be downloaded from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>). The PCOS group includes three datasets with GSE10946, GSE34526, and GSE580432. GSE103465, GSE111974, and GSE26787 are involved in the RIF group. Further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Scientific Research and Clinical Trial of the First Affiliated Hospital of Zhengzhou University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

References

- Teede H, Deeks A, Moran L. Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC Med* (2010) 8:41. doi: 10.1186/1741-7015-8-41
- Joham AE, Norman RJ, Stener-Victorin E, Legro RS, Franks S, Moran LJ, et al. Polycystic ovary syndrome. *Lancet Diabetes Endocrinol* (2022) 10(9):668–80. doi: 10.1016/S2213-8587(22)00163-2
- Jabbour HN, Sales KJ, Catalano RD, Norman JE. Inflammatory pathways in female reproductive health and disease. *Reproduction* (2009) 138(6):903–19. doi: 10.1530/REP-09-0247
- Hu C, Pang B, Ma Z, Yi H. Immunophenotypic profiles in polycystic ovary syndrome. *Mediators Inflammation* (2020) 2020:5894768. doi: 10.1155/2020/5894768
- Fessler DMT, Natterson-Horowitz B, Azziz R. Evolutionary determinants of polycystic ovary syndrome: part 2. *Fertil Steril* (2016) 106(1):42–7. doi: 10.1016/j.fertnstert.2016.05.016
- Velez LM, Seldin M, Motta AB. Inflammation and reproductive function in women with polycystic ovary syndrome. *Biol Reprod* (2021) 104(6):1205–17. doi: 10.1093/biolre/iaob050
- Zhang D, Yu Y, Duan T, Zhou Q. The role of macrophages in reproductive-related diseases. *Heliyon* (2022) 8(11):e11686. doi: 10.1016/j.heliyon.2022.e11686
- Coughlan C, Ledger W, Wang Q, Liu F, Demirel A, Gurgan T, et al. Recurrent implantation failure: definition and management. *Reprod BioMed Online* (2014) 28(1):14–38. doi: 10.1016/j.rbmo.2013.08.011
- Simon A, Laufer N. Assessment and treatment of repeated implantation failure (Rif). *J Assist Reprod Genet* (2012) 29(11):1227–39. doi: 10.1007/s10815-012-9861-4
- Robertson SA, Moldenhauer LM. Immunological determinants of implantation success. *Int J Dev Biol* (2014) 58(2-4):205–17. doi: 10.1387/ijdb.140096sr
- Guzeloglu-Kayisli O, Kayisli UA, Taylor HS. The role of growth factors and cytokines during implantation: endocrine and paracrine interactions. *Semin Reprod Med* (2009) 27(1):62–79. doi: 10.1055/s-0028-1108011
- Prabhudas M, Bonney E, Caron K. Et al. *Immune Mech at Maternal-Fetal Interface: Perspect Challenges. Nat Immunol* (2015) 16(4):328–34. doi: 10.1038/Ni.3131
- Xu L, Li Y, Sang Y, Li DJ, Du M. Crosstalk between trophoblasts and decidual immune cells: the cornerstone of maternal-fetal immunotolerance. *Front Immunol* (2021) 12:642392. doi: 10.3389/fimmu.2021.642392
- Ander Se, Diamond Ms, Coyne Cb. Immune responses at the maternal-fetal interface. *Sci Immunol* (2019) 4(31):Eaat6114. doi: 10.1126/Sciimmunol.Aaat6114
- Lai ZZ, Wang Y, Zhou WJ, Liang Z, Shi JW, Yang HL, et al. Single-cell transcriptome profiling of the human endometrium of patients with recurrent implantation failure. *Theranostics* (2022) 12(15):6527–47. doi: 10.7150/thno.74053

Author contributions

WC and QY conceived the study idea. WC performed the data analysis and produced the results. YS, QY, and WC wrote and revised the manuscript. LH, WC, MW, ZY, and XZ were responsible for sample collection and experiment. All authors contributed to the article and approved the submitted version.

Funding

National Key R&D Program of China to Yingpu Sun (grant number: 2019YFA0110900) and National Natural Science Foundation of China to Qingling Yang (grant number: 31970800).

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.

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/fimmu.2023.1175384/full#supplementary-material>

16. Huang J, Jin N, Qin H, Shi X, Liu Y, Cheung W, et al. Transcriptomic profiles in peripheral blood between women with unexplained recurrent implantation failure and recurrent miscarriage and the correlation with endometrium: a pilot study. *PLoS One* (2017) 12(12):e0189159. doi: 10.1371/journal.pone.0189159
17. Lessey BA, Young SL. What exactly is endometrial receptivity? *Fertility Sterility* (2019) 111(4):611–7. doi: 10.1016/j.fertnstert.2019.02.009
18. Revel A. Defective endometrial receptivity. *Fertil Steril* (2012) 97(5):1028–32. doi: 10.1016/j.fertnstert.2012.03.039
19. Jansen E, Laven JS, Dommerholt HB, Polman J, van Rijt C, van den Hurk C, et al. Abnormal gene expression profiles in human ovaries from polycystic ovary syndrome patients. *Mol Endocrinol* (2004) 18(12):3050–63. doi: 10.1210/me.2004-0074
20. Su NJ, Ma J, Feng DF, Zhou S, Li ZT, Zhou WP, et al. The peripheral blood transcriptome identifies dysregulation of inflammatory response genes in polycystic ovary syndrome. *Gynecol Endocrinol* (2018) 34(7):584–8. doi: 10.1080/09513590.2017.1418851
21. Iudica-Thomas S, Gawde U, Bhaye S, Pokar K, Bader GD. Meta-analysis of gene expression profiles of lean and obese pcos to identify differentially regulated pathways and risk of comorbidities. *Comput Struct Biotechnol J* (2020) 18:1735–45. doi: 10.1016/j.csbj.2020.06.023
22. Kara M, Ozcan SS, Aran T, Kara O, Yilmaz N. Evaluation of endometrial receptivity by measuring hoxa-10, hoxa-11, and leukemia inhibitory factor expression in patients with polycystic ovary syndrome. *Gynecol Minim Invasive Ther* (2019) 8(3):118–22. doi: 10.4103/GMIT.GMIT_112_18
23. Nicola NA, Babon JJ. Leukemia inhibitory factor (Lif). *Cytokine Growth Factor Rev* (2015) 26(5):533–44. doi: 10.1016/j.cytogfr.2015.07.001
24. Jiang NX, Li XL. The complicated effects of extracellular vesicles and their cargos on embryo implantation. *Front Endocrinol (Lausanne)* (2021) 12:681266. doi: 10.3389/fendo.2021.681266
25. Bai X, Zheng L, Li D, Xu Y. Research progress of endometrial receptivity in patients with polycystic ovary syndrome: a systematic review. *Reprod Biol Endocrinol* (2021) 19(1):122. doi: 10.1186/s12958-021-00802-4
26. Achaache H, Revel A. Endometrial receptivity markers, the journey to successful embryo implantation. *Hum Reprod Update* (2006) 12(6):731–46. doi: 10.1093/humupd/dml004
27. Rotterdam EA-SPCWG. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril* (2004) 81(1):19–25. doi: 10.1016/j.fertnstert.2003.10.004
28. Langfelder P, Horvath S. Wgcna: an R package for weighted correlation network analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559
29. 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
30. Day F, Karaderi T, Jones MR, Meun C, He C, Drong A, et al. . doi: 10.1101/290502
31. Gibbs GM, Lo JC, Nixon B, Jamsai D, O'Connor AE, Rijal S, et al. Glioma pathogenesis-related 1-like 1 is testis enriched, dynamically modified, and redistributed during male germ cell maturation and has a potential role in sperm-oocyte binding. *Endocrinology* (2010) 151(5):2331–42. doi: 10.1210/en.2009-1255
32. Pau CT, Mosbrugger T, Saxena R, Welt CK. Phenotype and tissue expression as a function of genetic risk in polycystic ovary syndrome. *PLoS One* (2017) 12(1):e0168870. doi: 10.1371/journal.pone.0168870
33. Queckbörner S, von Grothusen C, Boggavarapu NR, Francis RM, Davies LC, Gemzell-Danielsson K. Stromal heterogeneity in the human proliferative endometrium—a single-cell RNA sequencing study. *J Personalized Med* (2021) 11(6):448. doi: 10.3390/jpm11060448
34. Diaz-Gimeno P, Horcujadas JA, Martinez-Conejero JA, Esteban FJ, Alama P, Pellicer A, et al. A genomic diagnostic tool for human endometrial receptivity based on the transcriptomic signature. *Fertil Steril* (2011) 95(1):50–60. doi: 10.1016/j.fertnstert.2010.04.063
35. Ogata T, Sano S, Nagata E, Kato F, Fukami M. Maml1 and 46,Xy disorders of sex development. *Semin Reprod Med* (2012) 30(5):410–6. doi: 10.1055/s-0032-1324725
36. Ahmed SF, Bashamboo A, Lucas-Herald A, McElreavey K. Understanding the genetic aetiology in patients with xy dsd. *Br Med Bull* (2013) 106:67–89. doi: 10.1093/bmb/ldt008
37. Chen Y, Thai HT, Lundin J, Lagerstedt-Robinson K, Zhao S, Markljung E, et al. Mutational study of the Maml1-gene in hypospadias. *Eur J Med Genet* (2010) 53(3):122–6. doi: 10.1016/j.ejmg.2010.03.005
38. Sagvekar P, Kumar P, Mangoli V, Desai S, Mukherjee S. DNA Methylation profiling of granulosa cells reveals altered methylation in genes regulating vital ovarian functions in polycystic ovary syndrome. *Clin Epigenet* (2019) 11(1):61. doi: 10.1186/s13148-019-0657-6
39. Vannuccini S, Clifton VL, Fraser IS, Taylor HS, Critchley H, Giudice LC, et al. Infertility and reproductive disorders: impact of hormonal and inflammatory mechanisms on pregnancy outcome. *Hum Reprod Update* (2016) 22(1):104–15. doi: 10.1093/humupd/dmv044
40. Miyado M, Miyado K, Katsumi M, Saito K, Nakamura A, Shihara D, et al. (2015) 5:14705. doi: 10.1038/srep14705
41. Wang W, Wu Z, Dai Z, Yang Y, Wang J, Wu G. Glycine metabolism in animals and humans: implications for nutrition and health. *Amino Acids* (2013) 45(3):463–77. doi: 10.1007/s00726-013-1493-1
42. Holecsek M. *Serine metabolism in health and disease and as a conditionally essential amino acid*. *nutrients* (2022) 14(9). Epub 2022/05/15. doi: 10.3390/nu14091987
43. Cheng ZX, Guo C, Chen ZG, Yang TC, Zhang JY, Wang J, et al. Glycine, serine and threonine metabolism confounds efficacy of complement-mediated killing. *Nat Commun* (2019) 10(1):3325. doi: 10.1038/s41467-019-11129-5
44. Gong X, Liu J, Zhang D, Yang D, Min Z, Wen X, et al. Glipr1 modulates the response of cisplatin-resistant human lung cancer cells to cisplatin. *PLoS One* (2017) 12(8):e0182410. doi: 10.1371/journal.pone.0182410
45. Zhang L, Liu C, Jiang Q, Yin Y. Butyrate in energy metabolism: there is still more to learn. *Trends Endocrinol Metab* (2021) 32(3):159–69. doi: 10.1016/j.tem.2020.12.003
46. Wang Y, Christopher BA, Wilson KA, Muoio D, McGarrah RW, Brunengraber H, et al. Propionate-induced changes in cardiac metabolism, notably coa trapping, are not altered by l-carnitine. *Am J Physiol Endocrinol Metab* (2018) 315(4):E622–E33. doi: 10.1152/ajpendo.00081.2018
47. Ageyi-Owusu K, Leeper FJ. Thiamin diphosphate in biological chemistry: analogues of thiamin diphosphate in studies of enzymes and riboswitches. *FEBS J* (2009) 276(11):2905–16. doi: 10.1111/j.1742-4658.2009.07018.x
48. Gao Q, Ma C, Meng S, Wang G, Xing Q, Xu Y, et al. Exploration of molecular features of pcos with different androgen levels and immune-related prognostic biomarkers associated with implantation failure. *Front Endocrinol (Lausanne)* (2022) 13:946504. doi: 10.3389/fendo.2022.946504
49. Erlebacher A. Immunology of the maternal-fetal interface. *Annu Rev Immunol* (2013) 31:387–411. doi: 10.1146/annurev-immunol-032712-100003
50. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with cibersort. *Methods Mol Biol* (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1_12
51. Pusalkar M, Meherji P, Gokral J, Savardekar L, Chinnaraj S, Maitra A. Obesity and polycystic ovary syndrome: association with androgens, leptin and its genotypes. *Gynecol Endocrinol* (2010) 26(12):874–82. doi: 10.3109/09513590.2010.487586
52. Mor G, Cardenas I, Abrahams V, Guller S. Inflammation and pregnancy: the role of the immune system at the implantation site. *Ann N Y Acad Sci* (2011) 1221(1):80–7. doi: 10.1111/j.1749-6632.2010.05938.x
53. PrabhuDas M, Bonney E, Caron K, Dey S, Erlebacher A, Fazleabas A, et al. Immune mechanisms at the maternal-fetal interface: perspectives and challenges. *Nat Immunol* (2015) 16(4):328–34. doi: 10.1038/ni.3131
54. Choi I, Son H, Baek JH. *Tricarboxylic acid (Tca) cycle intermediates: regulators of immune responses*. *life (Basel)* (2021) 11(1). Epub 2021/01/23. doi: 10.3390/life11010069