

Exploring Biomarkers Related to Cell Adhesion in Polycystic Ovary Syndrome Using Bioinformatics and Conducting Experimental Validation

Jiani Zhu ^{1,2,*}, Xinyue Qi ^{3,4,*}, Zhenyu Zhang ^{1,2,*}, Qun Zhou ⁵, Ran Gu ^{1,2}, Xiaorong Wu ^{6,7}, Lanping Zhong ^{6,7}

¹Department of Clinical Nutrition, The First People's Hospital of Yunnan Province, Kunming, Yunnan, People's Republic of China; ²Department of Clinical Nutrition, The Affiliated Hospital of Kunming University of Science and Technology, Kunming, Yunnan, People's Republic of China; ³Department of Preventive Health Care, The First People's Hospital of Yunnan Province, Kunming, Yunnan, People's Republic of China; ⁴Department of Preventive Health Care, The Affiliated Hospital of Kunming University of Science and Technology, Kunming, Yunnan, People's Republic of China; ⁵Department of Internal Medicine No.3, Shilin Yi Autonomous County People's Hospital, Kunming, Yunnan, People's Republic of China; ⁶Department of Reproductive Medicine, The First People's Hospital of Yunnan Province, Kunming, Yunnan, People's Republic of China; ⁷Department of Reproductive Medicine, The Affiliated Hospital of Kunming University of Science and Technology, Kunming, Yunnan, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xiaorong Wu; Lanping Zhong, The First People's Hospital of Yunnan Province (The Affiliated Hospital of Kunming University of Science and Technology), No. 157 Jinbi Road, Xishan District, Kunming City, Yunnan Province, People's Republic of China, Tel +8615087160943; +8618064820074, Fax +86-0871-63637735; +86-0871-63637735, Email jw021355@163.com; watxbb985428@163.com

Purpose: Related studies have pointed out that cell adhesion may play an important role for treating Polycystic Ovary Syndrome (PCOS). This study aimed to identify and analyze the biomarkers associated with cell adhesion-related genes (CRGs) for treating PCOS and their biological mechanisms.

Patients and Methods: In this study, GSE80432 was used to identify differentially expressed genes (DEGs) (PCOS vs control group) through differential expression analysis. Then, the DEGs were overlapped with 1531 CRGs to obtain the cross - genes. Subsequently, the Support Vector Machine-Recursive Feature Elimination combined with the least absolute shrinkage and selection operator was utilized to obtain candidate genes, and the genes with AUC greater than 0.7 and consistent expression trends in the two datasets were defined as biomarkers. Finally, a nomogram was constructed, and enrichment analysis, regulatory network, drug prediction, the association between biomarkers and PCOS, and reverse transcription quantitative PCR (RT-qPCR) were carried out respectively.

Results: A total of 10 cross-genes were identified, and 2 biomarkers (DSG2 and THBS1) were screened out from them. RT-qPCR analysis showed that the expression of THBS1 was increased in PCOS samples, while there was no significant difference in DSG2. In addition, enrichment analysis indicated that both DSG2 and THBS1 were enriched in the B-cell receptor signaling pathway. Then, based on these two biomarkers, lncRNA-miRNA-mRNA (81 nodes and 135 edges) and TFs biomarker networks (38 nodes and 38 edges), such as MIR17HG'-has-miR-7-5p'-THBS1, TFDPI-DSG2, were constructed respectively. By predicting drugs targeting biomarkers, 61 drugs were predicted to target DSG2, while 133 drugs were predicted to target THBS1. Moreover, a stronger association between THBS1 and PCOS was detected (inference score = 27.15).

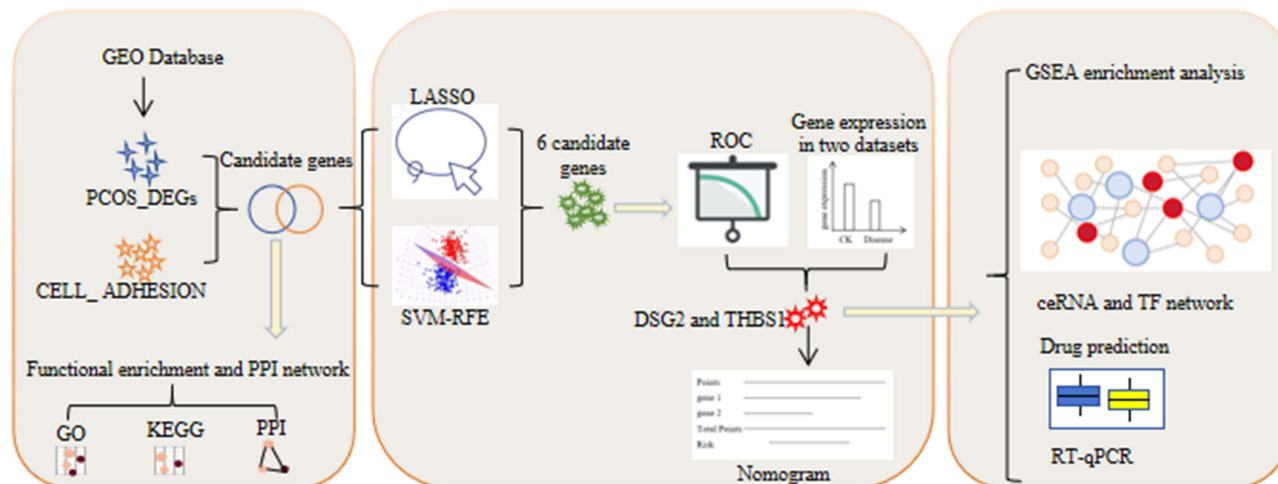
Conclusion: In this study, 2 biomarkers (DSG2 and THBS1) were identified, providing a potential theoretical basis for PCOS treatment.

Keywords: polycystic ovary syndrome, cell adhesion, nomogram, DSG2, THBS1

Introduction

Polycystic ovary syndrome (PCOS) is a common disorder of endocrine hormone regulation in women, with an incidence of 11%-13%, but the cause is not completely clear.¹ The basic pathological and physiological changes of its onset are hyperandrogenism (HA) and insulin resistance (IR).² Women with PCOS often present with irregular menstrual cycles, weight gain, noticeable increase in body hair, and acne. They are also at a higher risk of infertility and diabetes.³ At present, the etiology is not fully understood, and it is generally believed to be related to genetic factors, reproductive endocrine hormones,^{4,5} insulin

Graphical Abstract



resistance,⁶ inflammatory factors,⁷ gut microbiota dysbiosis,⁸ environmental endocrine disruptors,⁹ and other factors. Now, the main methods for treating PCOS are symptomatic treatment, including drug therapy,^{10,11} dietary intervention,¹² physical exercise,¹³ and surgical treatment,¹⁴ etc. however, which cannot effectively cure PCOS. In addition, the potential pathogenesis of PCOS still needs further research. Therefore, exploring new valuable biomarkers for effective treatment of PCOS is crucial.

Cell adhesion molecules are a collective term for numerous molecules that mediate the contact and binding between cells and extracellular matrix. They often function through receptor ligand interactions and participate in processes such as cell extension, adhesion, activation, signal transduction, and distant tumor metastasis. It connects cells in different ways and can participate in signal transduction for cell detection and response to changes in the surrounding environment,¹⁵ Changes in adhesion can disrupt important cellular processes and lead to various diseases.¹⁶ Related studies have shown that the secretion from the top of endometrial organoids can alter the adhesion of trophoblast cells, suggesting that cell adhesion may play an important role in the progression of uterine related diseases.¹⁷ In addition, it has been reported in the literature that the membrane protein- cadherin responsible for cell adhesion, can participate in cell signaling, regulating cell proliferation, apoptosis, survival and other biological processes. In adult ovaries, E- and N-cadherin ensure the integrity of ovarian follicles and the formation of luteum, indicating that cell adhesion also plays a key role in the ovarian organ.¹⁸ However, presently, there is no definitive theory on the possible biological mechanisms of cell adhesion-related genes (CRGs) in the progression of PCOS.

Currently, the potential biological mechanisms of cell adhesion-related genes in the progression of PCOS remain unclear. The objective of this study is to identify biomarkers and elucidate their biological mechanisms associated with cell adhesion in PCOS, thereby offering novel insights for PCOS treatment. By extracting transcriptome data and cell adhesion-related genes (CRGs) pertinent to PCOS from public databases, the study utilized a range of bioinformatics techniques to pinpoint biomarkers linked to PCOS cell adhesion. Furthermore, the study explored the biological function regulatory network, correlations, and drug prediction analyses of these biomarkers, with the goal of providing fresh references for the early diagnosis, prevention, and treatment of PCOS patients.

Material and Methods

Data Extraction

In this study, PCOS related data sets (GSE80432 and GSE34526) were derived from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/gds>). Among them, GSE80432 included 8 PCOS and 8 control granulosa cells samples, based on the GPL6244 platform. The GSE34526 included 7 PCOS and 3 control granulosa cells samples,

based on the GPL570 platform. Besides, we mined 1531 CRGs from the gobp cell adhesion.v2023.2.Hs.gmt downloaded from Molecular Signatures Database (MSigDB, <http://www.broadinstitute.org/gsea/msigdb/index.jsp>).

Differential Expression and Function Enrichment Analyses

Firstly, the differentially expressed genes (DEGs) (PCOS vs control) in GSE80432 were identified using “limma” (v3.56.2)¹⁹ ($P < 0.05, |\log_2\text{Fold Change (FC)}| > 0.50$). Then, the DEGs were visualized using the volcano plot and heatmap through “ggplot2” (v 3.4.2)²⁰ and “circlize” (v 0.4.15),²¹ respectively. Afterward, the intersection genes were identified by overlapping the DEGs and CRGs. Furthermore, to explore the biological functions of the intersection genes, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were carried out via “clusterProfiler” (v 4.8.2) ($P < 0.05$),²² and the results were displayed via “treemap” (v 2.4–4).²³ Moreover, to further understand the protein interaction of the intersection genes, a protein-protein interaction (PPI) network was constructed by STRING database (<https://string-db.org/>) (confidence > 0.4), and the results were presented through Cytoscape (v 3.9.1).²⁴

Machine Learning

In order to obtain the candidate genes, based on the intersection genes, the 2 machine learning algorithms were utilized to screen candidate genes, including Support Vector Machine-Recursive Feature Elimination (SVM-RFE) and least absolute shrinkage and selection operator (LASSO). The joint analysis of SVM-RFE and LASSO was enabled to more effectively handle complex datasets, and enhance the predictive ability and stability of the model through complementary feature selection methods, performance improvement of the model, dimension reduction, robustness enhancement, interpretability improvement, adaptation to multiple data types, and overfitting reduction. Concretely, based on the intersection genes, SVM-RFE was conducted using “e1071” (v 1.7–13).²⁵ The SVM-RFE method was utilized to obtain the importance and importance ranking of each gene in the intersection genes. Meanwhile, the error rate of each iterative combination was obtained. The combination with the lowest error rate was selected as the optimal combination, and the corresponding genes were obtained for subsequent analysis. LASSO analysis to intersection genes was performed using “glmnet” (v 4.1–7),²⁶ according to 3-fold cross-validation to screen genes. LASSO was a shrinkage estimation method. Its basic idea was that the residual sum of squares was minimized under the constraint that the sum of the absolute values of the regression coefficients was less than a constant, so that some regression coefficients could be strictly equal to 0, and a interpretable model could be further obtained. In this study, subsequent analysis was carried out based on the value with the minimum cross-validation error and the genes whose regression coefficients were not equal to 0. Importantly, the candidate genes were identified by overlapping the genes from 2 machine learning algorithms.

Biomarkers Were Identified and a Nomogram Was Constructed

To assess the diagnostic performance of candidate genes, receiver operating characteristic curve (ROC) curves of candidate genes were plotted (area under the curve (AUC) > 0.7) in GSE80432 using “pROC” (v 1.18.4).²⁷ Furthermore, gene expression analysis of candidate genes was performed in GSE80432 and GSE34526 datasets, respectively. The genes with significant differential expression between the case and control groups, as well as consistent expression trends in both GSE80432 and GSE34526 datasets, were defined as biomarkers. Based on the biomarkers, a nomogram was constructed using “rms” R package (v 6.7–0).²⁸ The decision curve analysis (DCA) curve and ROC curve were separately plotted to assess the prediction ability of nomogram.

Gene Set Enrichment Analysis (GSEA)

To understand the functions and signaling pathways of biomarkers, “psych” (v 2.3.6)²⁹ was used to conduct the Spearman correlation analysis, the correlation coefficient was calculated and ranked between biomarkers and other genes in GSE80432. Additionally, c2.cp.reactome.v7.0.symbols.gmt was selected as background gene sets from molecular signatures database (MSigDB) (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>). Subsequently, GSEA was conducted according to a screening condition of the adj.P < 0.05 .

Correlation Analysis of Biomarkers and Metabolic Pathways

To investigate the relationship between biomarkers and metabolic pathways, we referred to nine metabolism-related pathways from the literature.³⁰ In the training set, ssGSEA enrichment analysis was first performed using the “GSVA” (v1.48.2)³¹ to calculate the enrichment scores of the metabolic pathways. Subsequently, Spearman correlation analysis between the biomarkers and metabolic pathways was conducted using the “corrplot” (v 0.92).³², with $p < 0.05$ and $|\text{cor}| > 0.3$ considered statistically significant.

Regulation Networks Analysis

The miRNA targeting biomarkers was predicted using starbase database (<http://starbase.sysu.edu.cn/>), the screening conditions were clipExpNum > 3. Likewise, the lncRNA targeting predicted miRNA also was predicted utilizing starbase database (clipExpNum > 100). Subsequently, the Cytoscape was employed to present the lncRNA-miRNA-mRNA network. Furthermore, to investigate the upstream transcription factors (TFs) of biomarkers, the TF list targeting biomarkers was downloaded from the NetworkAnalyst (<https://www.networkanalyst.ca/NetworkAnalyst/>) website, and the TFs were selected from the ENCODE CHIP-seq database (<https://www.encodeproject.org/>).

Drug Prediction and Biomarkers-Disease Association Analysis

Firstly, the Comparative Toxicogenomics Database (CTD) (<http://ctdbase.org/>) was used to gain the biomarkers-related drugs and chemicals, and a biomarkers-drugs and chemicals network was constructed according to the drugs with a reference count greater than 2. In addition, using the CTD database to study the association between biomarkers and PCOS, the inference score represents the relevance score of biomarkers to PCOS, with higher score indicating a stronger correlation between biomarkers and PCOS.

The Reverse Transcription Quantitative PCR (RT-qPCR)

A total of 10 tissue samples (5 normal and 5 PCOS samples) were acquired from the clinic in the First People’s Hospital of Yunnan Province. All participants were given informed consent. The study had the approval of the First People’s Hospital of Yunnan Province ethics committee (approval number: KHLL2023-KY048).

Total RNA was extracted from 10 samples using the TRIzol reagent (Ambion, USA) following the manufacturer’s protocol. Briefly, cells were lysed and resuspended in centrifuge tubes, to which 1 mL of TRIzol was added. After a 10-minute incubation, 300 μL of chloroform was added, and the mixture was centrifuged at 12,000g, 4°C for 15 minutes to separate the RNA in the supernatant. Subsequently, an equal volume of ice-cold isopropanol was added to the RNA-containing supernatant, followed by a 10-minute incubation and centrifugation at 12,000g, 4°C for 10 minutes to precipitate the RNA. The RNA pellet was washed with 1 mL of 75% ethanol, incubated for 2 minutes, and then centrifuged at 7,500g, 4°C for 5 minutes. After air-drying for 20 minutes or blow-drying in an ultra-clean bench to remove residual ethanol and water, the RNA pellet became transparent. Finally, 20–50 μL of RNase-free water was added to dissolve the RNA completely, followed by a 15-minute incubation.

Next, the concentration of RNA was determined using the NanoPhotometer N50. Subsequently, cDNA was synthesized by reverse transcription using the SureScript First-strand cDNA synthesis kit, and the reverse transcription reaction was conducted using the S1000TM Thermal Cycler (Bio-Rad, USA). Quantitative PCR (qPCR) assays were performed using the CFX Connect Real-time Quantitative Fluorescence PCR Instrument (Bio-Rad, USA) (pre-denaturation at 95 °C for 1 min, denaturation at 95°C for 20s, annealing at 55°C for 20s, extension at 72°C for 30s, a total of 40 cycles). The relative quantification of mRNA levels was calculated using the $2^{-\Delta\Delta\text{CT}}$ method. Primer sequences can be found in Table 1.

Statistical Analysis

The R (v 4.2.2) was utilized to conduct statistical analysis. The differences between the two groups were tested using the Wilcoxon test. ($P < 0.05$), and the PCR results were analyzed using the *t*-test.

Table 1 Primer Sequence Lists

Primer	Sequence
DSG2 F	TCTTCTAGGCAGGCGCAAAA
DSG2 R	TGGTTGGTGGCATAAGTGGAC
THBS1 F	AAACAGGATGGTGGCTGGAG
THBS1 R	TGGACAGGCATCCATTGGG
Internal reference-GAPDH F	CGAAGGTGGAGTCAACGGATTT
Internal reference-GAPDH R	ATGGGTGGAATCATATTGGAAC

Abbreviations: PCOS, polycystic ovary syndrome; DEGs, differentially expressed genes; HA, hyperandrogenism; SVM-RFE, support vector machine recursive feature elimination; DCA, decision curve analysis; FC, fold change; IR, insulin resistance; LASSO, least absolute shrinkage and selection operator; MSigDBq, molecular signatures database; RT-qPCR, reverse transcription-quantitative polymerase chain reaction; PCA, principal component analysis; GEO, gene expression omnibus; GO, gene ontology; KEGG, Kyoto encyclopedia of genes and genomes; AUC, area under curve; BP, biological process; MF, molecular function; CC, cellular component; ROC, receiver operating characteristic; ssGSEA, single-sample gene set enrichment analysis; TFs, transcription factors; CTD, comparative toxicogenomics database; CRG, cell adhesion-related genes; F, forward; R, reverse.

Results

The 10 Intersection Genes Were Mainly Enriched in Various Pathway

To present the analysis results of DEGs and CRGs during the research process, functional enrichment analysis and PPI network construction were performed to reveal the potential roles and relationships of these genes in biological processes. A total of 105 DEGs were identified (23 up-regulation and 82 down-regulation) (Figure 1a and b). Then, 10 intersection genes were screened by overlapping the 105 DEGs and 1,531 CRGs (Figure 1c). Furthermore, GO and KEGG enrichment analysis showed that the 10 intersection genes were enriched for 771 GO terms, such as negative regulation of epithelial cell proliferation, platelet alpha granule lumen, protein binding involved in heterotypic cell-cell adhesion (Figure 1d). Meanwhile, intersection genes were enriched in 5 KEGG pathways, including malaria, microRNAs in cancer, TGF-beta signaling pathway, proteoglycans in cancer, bladder cancer (Figure 1e). Moreover, according to the confidence > 0.4, the PPI network of intersection genes included 9 nodes and 4 interaction relationships, including THBS1-FGG, THBS1-TGFB2, THBS1-DSG2, and THBS1-NRCAM (Figure 1f).

Six Candidate Genes Were Identified Through Machine Learning, Out of Which 2 Were Considered as Biomarkers

Based on 10 intersection genes, machine-learning algorithms and expression level validation were used to further screen for biomarkers. The 9 and 6 genes were separately identified by SVM-RFE (the model with the smallest error) and LASSO ($\lambda_{\min} = 0.021$) (Figure 2a–c). Subsequently, the 6 candidate genes (FGG, MIR222, DSG2, TGFB2, GLI3, and THBS1) were chosen by overlapping the genes from SVM-RFE and LASSO (Figure 2d). To assess the diagnostic performance of 6 candidate genes, the ROC curves in GSE80432 ($AUC_{FGG} = 0.859$, $AUC_{MIR222} = 0.844$, $AUC_{DSG2} = 0.875$; $AUC_{TGFB2} = 0.828$, $AUC_{GLI3} = 0.828$, $AUC_{THBS1} = 0.797$) showed the candidate genes could distinguish PCOS and control samples (Figure 2e). After, DSG2 and THBS1 had significant differential expression between the case and control groups, as well as consistent expression trends in both GSE80432 and GSE34526 datasets, thus DSG2 and THBS1 were defined as biomarkers. Specifically, DSG2 had a significantly low expression in case samples, while THBS1 was high expression in case samples (Figures 2f and g). Furthermore, RT-qPCR analysis revealed elevated expression of THBS1 in PCOS samples ($P = 0.0462$), while DSG2 showed no significant difference between the control and PCOS samples (Figures 2h and i).

The Nomogram Had an Outstanding Predicted Ability

A nomogram was constructed based on two biomarkers, and its predictive performance and diagnostic utility for PCOS were evaluated. The results showed its superior performance in prediction and diagnosis. Based on 2 biomarkers,

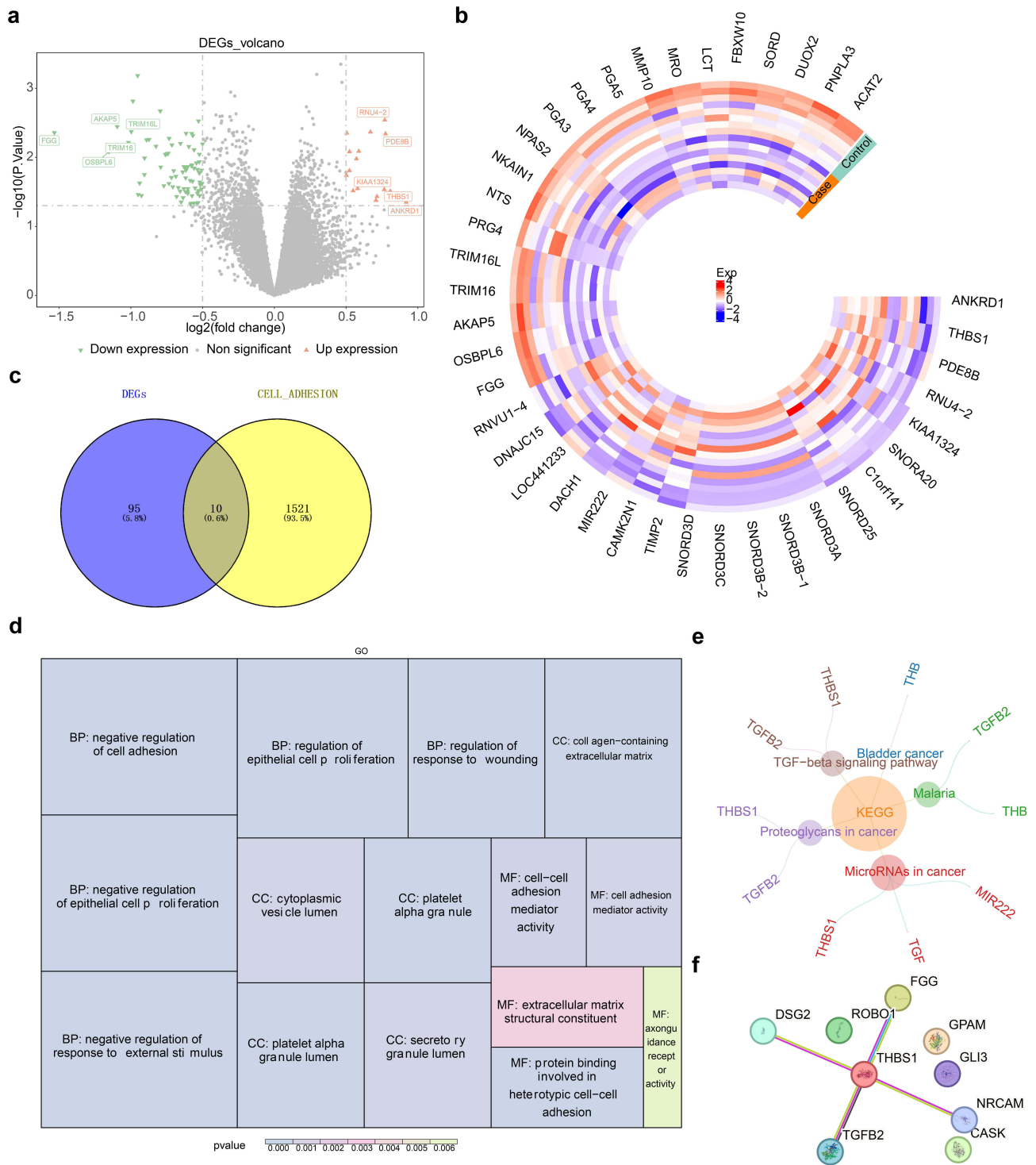


Figure 1 The results of differential expression analysis and enrichment analysis. (a) The volcano map based on the results of differential expression analysis, with Orange indicating genes with up-regulated expression and green indicating genes with down-regulated expression. (b) The expression heat map of differentially expressed genes. Red colour indicated up-regulated genes and blue colour represented down-regulated genes. (c) The ven plot of the intersection of differentially expressed genes (DEGs) and cell adhesion-related genes (CRGs). (d) The gene ontology (GO) enrichment analysis results. The block size represented the number of enriched genes; Color represented significance. (e) The Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis results. The size of the circle represented the number of genes contained. (f) The protein-protein interaction (PPI) network plot.

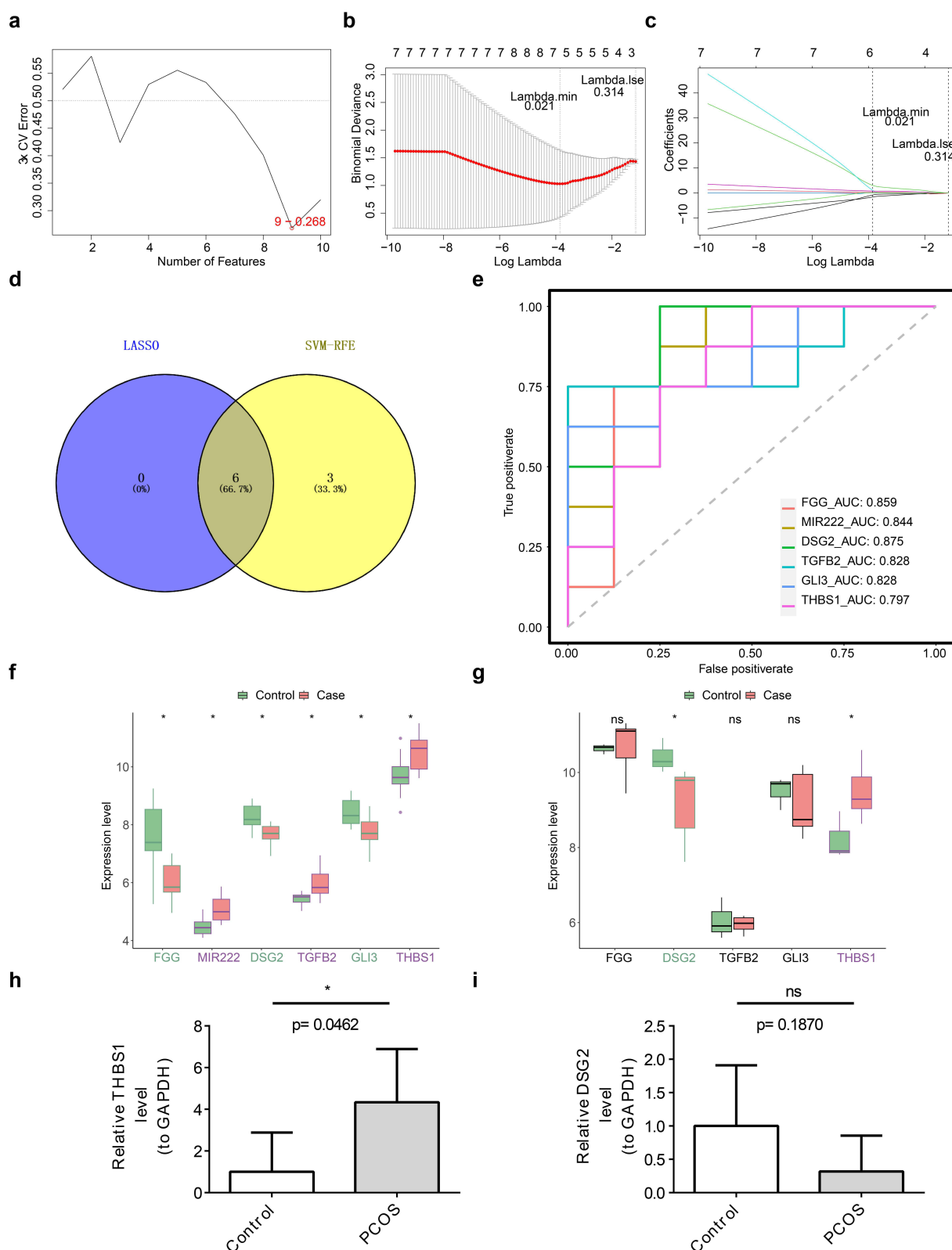


Figure 2 The results of the Machine learning. (a) The results of support vector machine recursive feature elimination (SVM-RFE) analysis. (b) The coefficients of the optimal (λ) corresponding genes. The left dashed line represented the position with the minimum cross validation error (λ_{\min}), with the number of feature genes displayed on top. The right dashed line represented the optimal log (λ) value. (c) Regression coefficient plot of least absolute shrinkage and selection operator (LASSO). (d) The intersection of LASSO and SVM-RFE analysis results taken to obtain candidate genes. (e) The receiver operating characteristic (ROC) curves of candidate genes in GSE80432. (f) The expression validation results of candidate genes in dataset GSE80432. “*” represented $P < 0.05$. (g) The expression validation results of candidate genes in dataset GSE34526. “ns” represented no significance, “*” represented $P < 0.05$. (h) Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) analysis revealed elevated expression of THBS1 in polycystic ovary syndrome (PCOS) samples. “*” represented $P < 0.05$. (i) RT-qPCR analysis revealed elevated expression of DSG2 in PCOS samples. “ns” represented no significance.

a nomogram was constructed, the higher the total points, the greater the probability of PCOS diagnosis (Figure 3a). Additionally, the DCA values indicated that the net income of the nomogram surpasses that of the single factor, suggesting a superior predictive effect of the nomogram (Figure 3b). After that, the ROC curve (AUC=0.812) indicating that the nomogram was a good predictor (Figure 3c).

THBS1 and DSG2 Were Enriched in B-Cell Receptor-Signaling Pathway, and There Was a Correlation Between Biomarkers and Metabolic Pathways

Functional pathways and metabolic correlations associated with the biomarkers THBS1 and DSG2 were explored. Based on the 2 biomarkers, the top 5 KEGG enrichment results of GSEA were separately presented in Figure 4a and b. Specifically, THBS1 was mainly enriched in B-cell receptor-signaling pathway, chemokine signaling pathway, graft versus host disease, etc. DSG2 was involved in B-cell receptor-signaling pathway, natural killer cell mediated cytotoxicity, and so on. Furthermore, the correlation between biomarkers and metabolic pathway showed DSG2 had a strongly

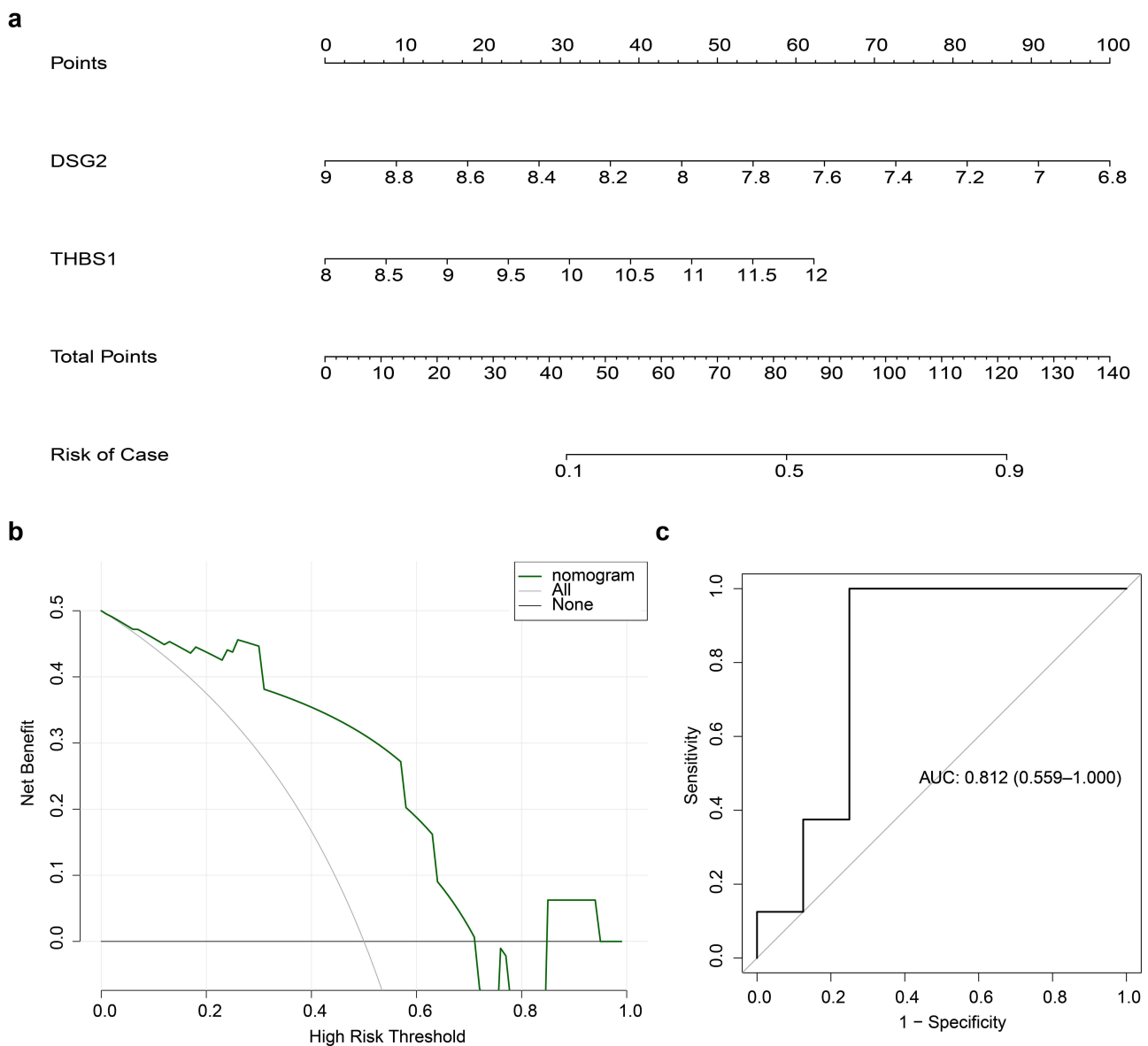


Figure 3 Constructed nomogram model. (a) The construction of a nomogram based on 2 biomarkers. (b) The decision curve analysis (DCA) values. (c) The receiver operating characteristic (ROC) curve.

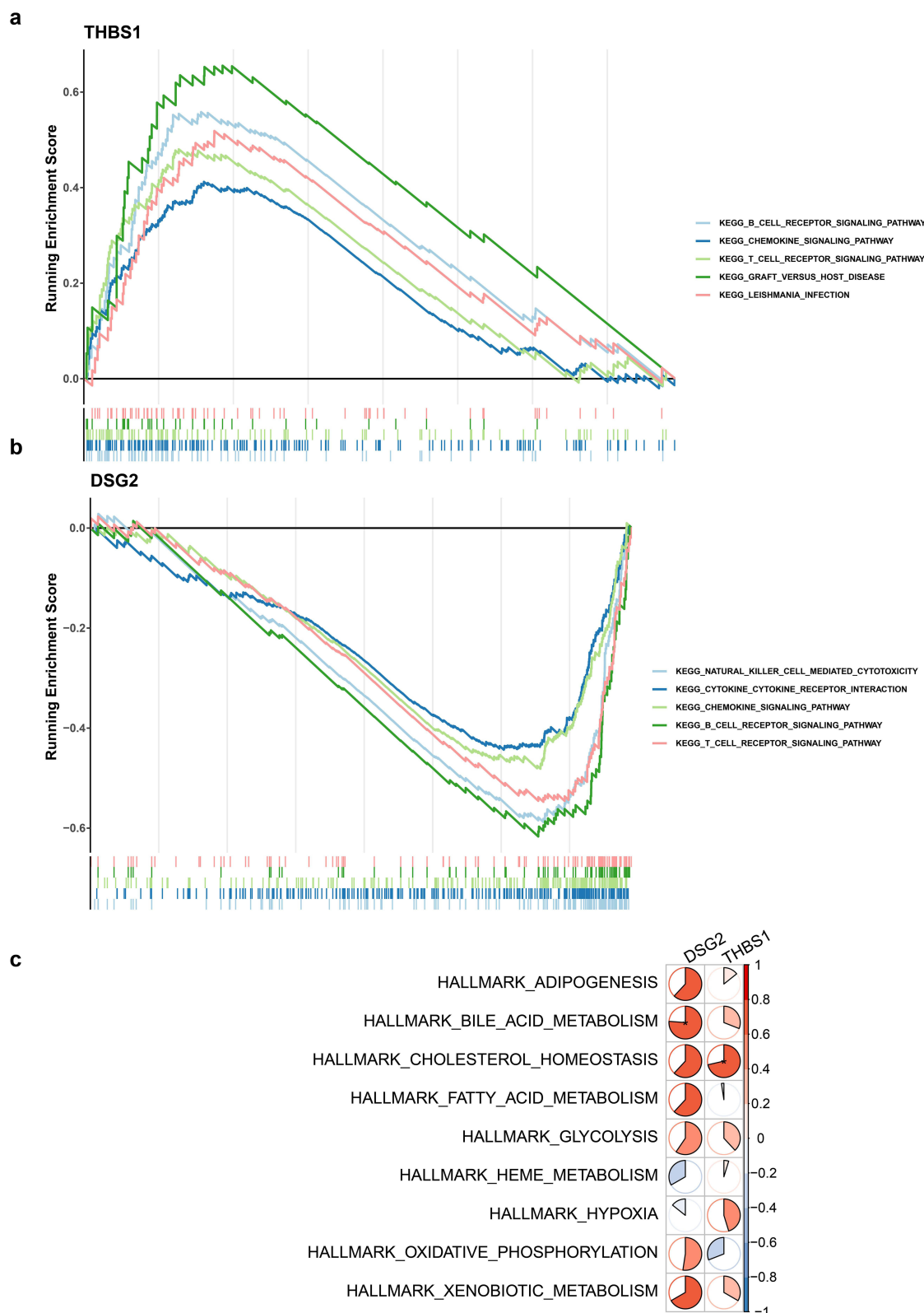


Figure 4 The functional analysis of biomarkers. (a) The results of GSEA analysis of THBS1D. The above section showed the process of calculating enrichment score (ES) values. For each gene from left to right, an ES value was calculated and connected into a line. The peak was the ES of this pathway gene set, and the lower line marked the genes located under this gene set. (b) The results of gene set enrichment analysis (GSEA) analysis of DSG2. (c) the correlation between biomarkers and metabolic pathway. Blue represented negative correlation, red represented positive correlation.

positive correlation with bile acid metabolism ($r=0.7619$, $p<0.05$), THBS1 and cholesterol homeostasis existed strongly correlation ($r=0.7143$, $p<0.05$) (Figure 4c).

Regulation Network Could Aid in the Exploration of Potential Mechanisms for PCOS

lncRNA-miRNA-mRNA and TF-biomarker networks were constructed to explore regulatory interactions involving DSG2 and THBS1. By selecting the predicted miRNAs targeting biomarkers, the 4 miRNAs targeting DSG2 and 189 miRNAs targeting THBS1 were separately acquired. Then, 193 lncRNAs were predicted by targeting miRNAs, however, only 20 lncRNAs targeting 59 miRNAs were selected according to the screening standard clipExpNum > 100. Ultimately, a lncRNA-miRNA-mRNA network, including 81 nodes and 135 edges, was built, for example, MIR17HG'-has-miR-7-5p'-THBS1, SNHG29'-has-miR-223-3p-DSG2' and so on (Figure 5a). The 36 TFs targeting 2 biomarkers were predicted, and a TFs-biomarkers network (38 nodes and 38 edges) was constructed, such as TFDP1-DSG2, ELF3-THBS1, and so on (Figure 5b).

Both THBS1 and DSG2 Provided a Potential Theoretical Basis for the Treatment of PCOS

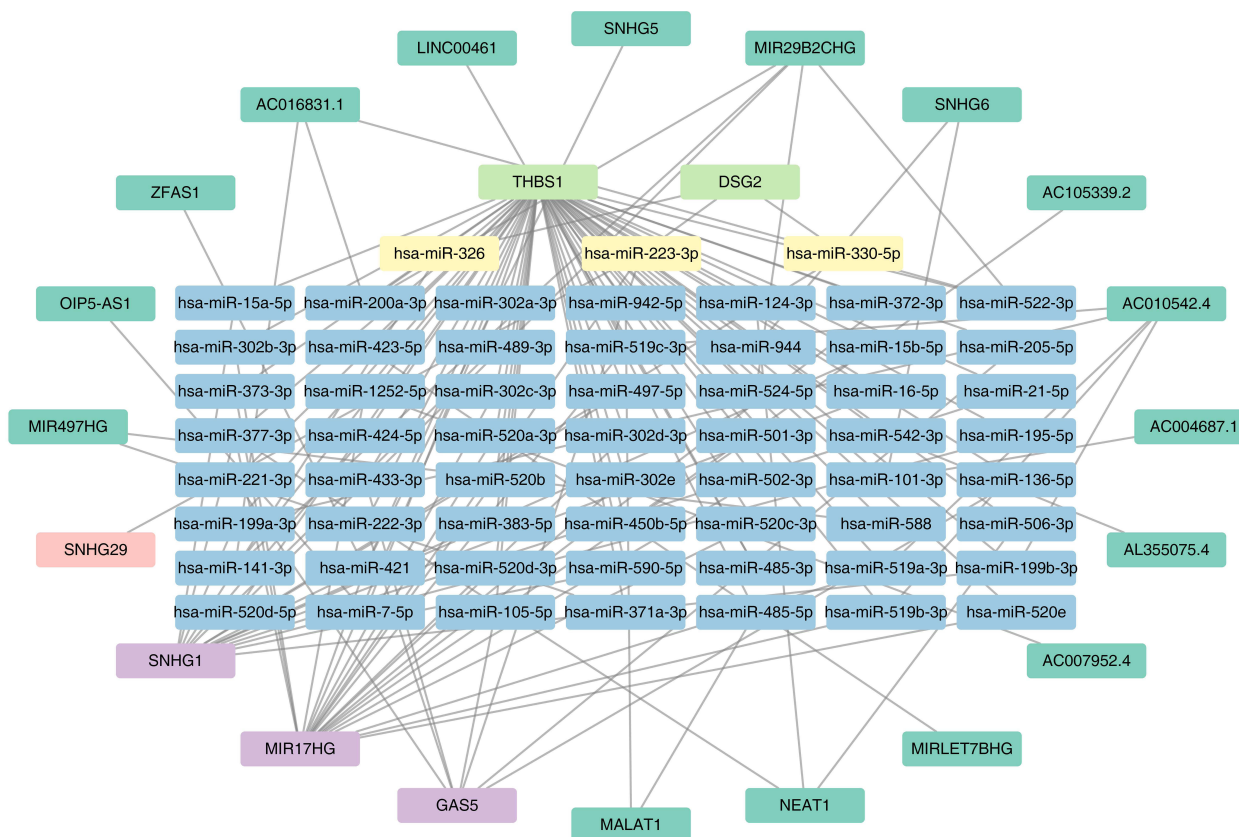
Drugs targeting DSG2 and THBS1 were predicted and analyzed, and their associations with diseases, particularly PCOS, were explored. According to the CTD database, 61 drugs were predicted to target DSG2, while 133 drugs were predicted to target THBS1. Importantly, there were 23 drugs were acquired by overlapping the 61 drugs and 133 drugs (Figure 6a). Subsequently, based on the predicted drugs, a biomarkers-drugs and chemicals network, including 27 nodes and 27 relationships, was constructed using the screening condition Reference Count>2 (Figure 6b). Such as THBS1-Tobacco Smoke Pollution, DSG2-Valproic Acid, and so on. Moreover, the association between genes and diseases indicated that the inference score between THBS1 and PCOS was 27.15, suggesting a stronger association between THBS1 and PCOS compared to DSG2 and PCOS (Figure 6c).

Discussion

In multicellular organisms, each cell needs to interact with surrounding cells and the environment to achieve its function. This interaction is mainly completed by cell adhesion. Problems in the formation or regulation of adhesive structures can lead to various diseases including immune disorders, developmental disorders, hematological disorders, and cancer.^{16,33–35} PCOS is a common endocrine and metabolic disease, and cell adhesion molecules may play multiple key roles in its pathophysiological processes. Revealing the role of these molecular mechanisms in PCOS provides potential targets for early diagnosis and clinical intervention of the disease.

DSG2 gene is a gene located on human chromosome 18q12.1, encoding the protein desmoglein 2, which is a member of the calcium binding protein family. As a transmembrane glycoprotein, it plays an important role in intercellular adhesion and cell signal transduction, extensively expressed in the basal cell layer of monolayer epithelium and stratified epithelium, and cardiac muscle, etc.^{36,37} DSG2 has been extensively studied in oncology, with recent findings revealing its non-classical roles in regulating cancer cell proliferation, migration, and other processes.³⁸ The dysregulated expression and prognostic significance of DSG2 vary across cancer types. For instance, elevated DSG2 expression correlates with poor prognosis in cervical cancer,^{39,40} lung cancer,^{41,42} and pancreatic cancer,⁴³ whereas reduced DSG2 expression is associated with unfavorable outcomes in colon cancer,⁴⁴ gastric cancer,⁴⁵ prostate cancer,⁴⁶ and high-grade serous ovarian cancer.⁴⁷ However, conflicting findings exist. For example, one study demonstrated that DSG2 expression positively correlates with ovarian cancer grade, showing significantly higher levels in high-grade tissues compared to low-grade tissues. Patients with elevated serum DSG2 levels exhibited markedly shorter progression-free survival (PFS) and overall survival (OS), with median PFS of 16 months in the high-DSG2 group versus 26 months in the low-DSG2 group.⁴⁸ Mechanistically, DSG2 interacts with the epidermal growth factor receptor (EGFR) to activate c-Src and STAT3 signaling pathways, thereby promoting cancer cell proliferation and migration.⁴⁹ Additionally, DSG2 undergoes proteolytic cleavage by multiple proteases, including matrix metalloproteinase (MMP)-9, ADAM10,^{50,51} and ADAM17.^{52–54} This cleavage is linked to weakened intercellular adhesion and reduced cell proliferation. DSG2 may also disrupt ovarian endothelial cell adhesion mechanisms, potentially influencing follicular development and hormonal secretion, thereby contributing to the clinical manifestations of polycystic ovary syndrome (PCOS). Further investigation into its underlying mechanisms is urgently needed.

a



b

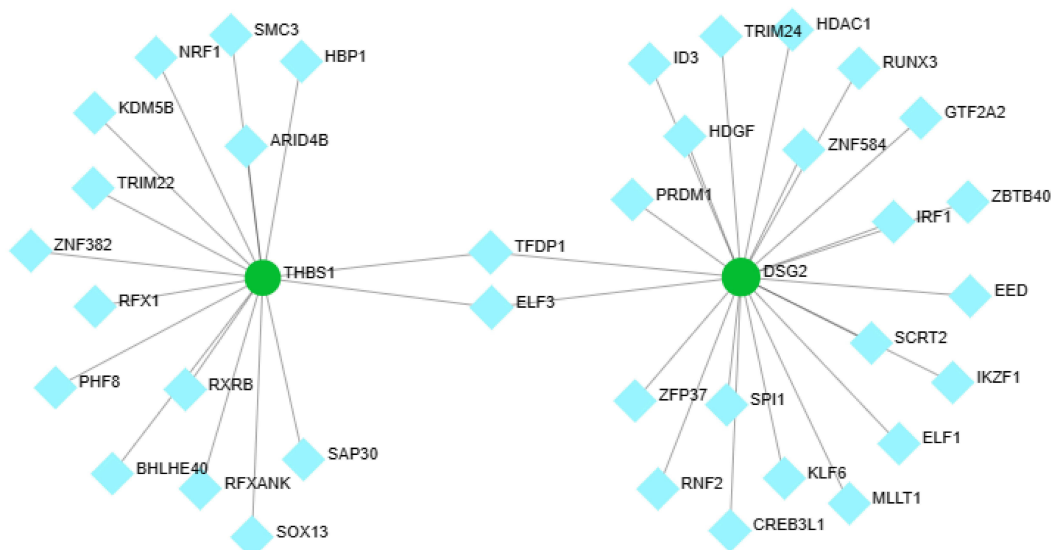


Figure 5 Network construction. (a) A lncRNA-miRNA-mRNA network. Green represented mRNA, blue represented miRNA predicted by THBS1, yellow represented miRNA predicted by DSG2, and cyan represented lncRNA predicted by THBS1, pink represented lncRNA predicted by DSG2, and purple represented shared lncRNA. (b) A TF-biomarker network was constructed based on 2 biomarkers. Green represented mRNA, and blue squares represented transcription factors.

Studies indicate that inflammatory mediators such as IL-1 β and TNF- α can activate these proteases, inducing extracellular and intracellular cleavage of DSG2. These mediators also regulate extracellular vesicle release, potentially modulating the tumor microenvironment.^{50,51,53,55}

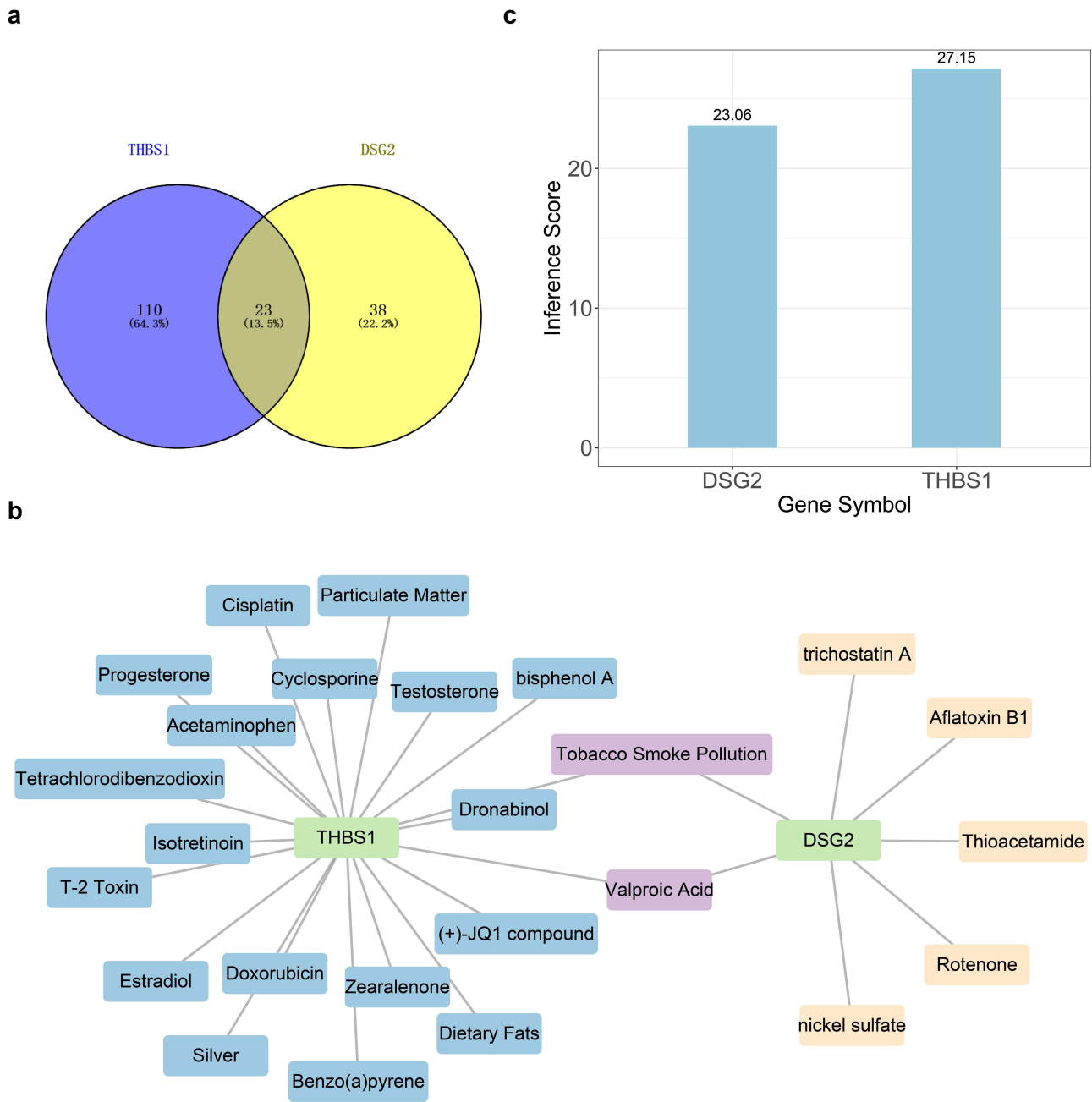


Figure 6 Biomarker-based targeted drug prediction, network construction and gene-disease correlation analysis. (a) Biological biomarker-based targeted drug prediction. (b) A biomarker-drug and chemical network was constructed. Green represented mRNA, blue represented drugs predicted by THBS1, Orange represented drugs predicted by DSG2, and purple represented drugs predicted jointly. (c) Correlation analysis between genes and diseases.

To date, no direct studies have explored the role of DSG2 in PCOS, and its specific mechanisms remain unclear. PCOS, a reproductive endocrine disorder involving multiple systems, is characterized by hyperandrogenism, insulin resistance, and abnormal follicular development. Building on DSG2’s established roles in cell adhesion, signaling, and reproductive system function, we hypothesize that aberrant DSG2 expression, as a core desmosomal protein, may disrupt adhesive junctions between ovarian granulosa cells and oocytes, altering follicular structure, hormonal secretion patterns, and follicular maturation/ovulation. Subsequent research can be validated through single-cell sequencing technology and animal models.

THBS1 gene, located on human chromosome 15q14, encodes thrombospondin-1, an extracellular matrix glycoprotein. THBS1 protein plays an important role in physiological and pathological processes such as cell adhesion, angiogenesis, inflammation, and tumorigenesis. It regulates cell adhesion, migration, and proliferation by interacting with cell surface receptors such as integrin, CD36, and heparan sulfate proteoglycans.⁵⁶ The dynamic changes of the ovary are significant, and follicular development, ovulation, and corpus luteum formation involve complex cell adhesion and extracellular matrix remodeling. As a key cell adhesion molecule, THBS1 may affect the normal development and ovulation of follicles by regulating the adhesion and communication between ovarian granulosa cells and follicular membrane cells. Studies have shown that abnormal expression of THBS1 in PCOS patients may lead to follicular maturation disorders and ovulation failure.⁵⁷

Insulin resistance is a hallmark of PCOS. Proteomic studies reveal a positive correlation between THBS1 and prediabetes.⁵⁸ Moreover, elevated THBS1 expression in adipose tissue of PCOS patients significantly correlates with the insulin resistance index (HOMA-IR).⁵⁹

THBS1 is pivotal in PCOS pathogenesis, and therapeutic strategies targeting THBS1 may offer novel approaches. Inhibiting THBS1 expression or blocking its signaling pathways could improve ovarian function, mitigate inflammation and insulin resistance, and alleviate PCOS symptoms.

GSEA enrichment analysis identified both biomarkers (DSG2 and THBS1) in the B-cell receptor and T-cell receptor signaling pathways. PCOS, a multifactorial endocrine and reproductive disorder, involves diverse molecular mechanisms and signaling pathways.

Current research on PCOS and T-cell pathways primarily focuses on immune dysregulation and chronic inflammation. PCOS patients exhibit altered local and systemic immune environments, with immune dysfunction and chronic inflammation closely linked to disease progression.⁶⁰ Studies report significantly reduced percentages of CD3+ and CD8+ T cells in PCOS patients compared to controls, while CD4+ T cell alterations may promote PCOS via cytokine dysregulation.⁶¹ Proteomic analysis of CD4+ T lymphocytes from PCOS patients using two-dimensional gel electrophoresis and mass spectrometry revealed distinct protein expression patterns compared to healthy women, suggesting immune system involvement in PCOS pathogenesis.⁶² As a bridging protein, DSG2 may play a role in immune synapse formation, affecting the contact between T cells and antigen-presenting cells. THBS1 and DSG2 may affect immune responses in B cell and T cell receptor signaling pathways by regulating intercellular interactions and signal transduction. Although no direct studies link B-cell signaling to PCOS, indirect evidence highlights associations. PCOS is characterized by chronic low-grade inflammation, with ovarian granulosa cells exhibiting a proinflammatory state marked by overexpression of cytokines like IL-6 and TNF- α , which impair follicular development and oocyte quality.⁶³ Immune dysregulation is prevalent in PCOS, with abnormal macrophage and T-cell function exacerbated by obesity and insulin resistance, amplifying chronic inflammation.^{7,64}

Immune mechanisms are critical in PCOS development, and deeper exploration may yield novel diagnostic and therapeutic targets. Drug prediction analysis identified two shared biomarkers: Tobacco Smoke Pollution and Valproic Acid.

China is a major producer and consumer of tobacco, with 40.7% of women exposed daily to secondhand smoke (SHS).⁶⁵ SHS contains over 9,000 chemicals, including 80 toxic carcinogens.⁶⁶ Nicotine, a key endocrine disruptor in SHS, inhibits aromatase activity, which converts androgens to estrogens in the ovary.^{67,68} SHS also disrupts estrogen/progesterone synthesis, impairing follicular development and ovulation. Women with high SHS exposure exhibit elevated follicle-stimulating hormone (FSH) and reduced estradiol levels, leading to follicular arrest, polycystic ovarian morphology, and hyperandrogenism. SHS further damages female germ cells and reproductive function, increasing infertility risk.^{69–71}

Acute SHS exposure impairs glucose tolerance, while chronic exposure induces insulin resistance and metabolic syndrome—core features of PCOS^{72,73}. Despite widespread awareness of SHS-induced lung cancer, research on its impact on PCOS-related reproductive endocrinology and pregnancy outcomes remains limited.

Valproic acid (VPA) is a widely used antiepileptic drug and mood stabilizer. However, a large number of studies have shown that the use of VPA is closely related to female reproductive endocrine dysfunction, especially the occurrence of PCOS. For example, a clinical case study showed that a 15-year-old woman who successfully controlled her seizures after receiving VPA treatment experienced weight gain and amenorrhea symptoms, ultimately diagnosed with PCOS. After discontinuing the medication, his weight decreased and his menstrual cycle returned to normal.⁷⁴ Another study used VPA to treat human ovarian granulosa cell line KGN and found that cell activity and progesterone production were

both inhibited. VPA interferes with steroid metabolism pathways related to ovarian function by suppressing the expression of steroidogenic genes, thereby inhibiting the biosynthesis of progesterone.⁷⁵ Some studies analyze the relationship between PCOS and the use of VPA through large-scale medical data analysis. For example, a meta-analysis evaluating the relationship between VPA and the incidence rate of PCOS showed that women receiving VPA had a significantly increased risk of PCOS.⁷⁶ Although VPA has significant therapeutic effects in treating neurological and emotional disorders, its impact on female reproductive health, especially PCOS, cannot be ignored. Therefore, it is still necessary to explore treatment options that reduce these side effects. Clinical doctors should carefully assess the individual risks of patients before treatment and continuously monitor them during the treatment process.

Valproic acid (VPA), a widely prescribed antiepileptic drug and mood stabilizer, has been increasingly associated with female reproductive endocrine dysfunction, particularly the development of polycystic ovary syndrome (PCOS). Substantial clinical and experimental evidence supports this association. For instance, a case report documented a 15-year-old female who developed PCOS manifestations—including significant weight gain and amenorrhea—during VPA therapy for seizure control. Notably, these symptoms resolved upon drug discontinuation, with subsequent normalization of body weight and menstrual cyclicity.⁷⁴ Mechanistic insights from *in vitro* studies utilizing the human ovarian granulosa cell line KGN revealed that VPA treatment suppressed both cellular viability and progesterone production. This effect was attributed to VPA-mediated downregulation of steroidogenic gene expression, which disrupts ovarian steroid metabolism pathways critical for progesterone biosynthesis.⁷⁵ Large-scale epidemiological analyses further corroborate these findings. A meta-analysis evaluating neuropsychiatric drug-related adverse effects demonstrated that women undergoing VPA therapy faced an 80% increased risk of PCOS diagnosis compared to non-exposed populations.⁷⁶ While VPA remains clinically indispensable for managing neurological and psychiatric disorders, its iatrogenic impact on reproductive health necessitates heightened vigilance. To mitigate these effects, future research should prioritize the development of targeted therapeutic strategies that preserve VPA's neuroprotective benefits while minimizing endocrine disruption. Clinically, rigorous pretreatment risk stratification and continuous monitoring of metabolic and reproductive parameters are imperative for patients requiring long-term VPA administration.

This study faces several limitations, primarily the small sample size, which could compromise the universality and representativeness of the findings. Additionally, the constraints of data sources may result in partial unreliability of certain results. The research methodology may not be universally applicable to all data types, posing technical constraints. Furthermore, time and resource limitations impact the depth and breadth of the experimental scope. Future research should aim to expand the sample size, utilize multi-center data to enhance the representativeness of the outcomes, refine research methods to accommodate diverse data types, and secure additional resource support to conduct more thorough and comprehensive experimental investigations.

Conclusion

This study identified two PCOS-related biomarkers (DSG2 and THBS1) via differential analysis, machine learning, and gene expression assessment, and developed a nomogram. This achievement bears substantial importance for theoretical progress in the fields of reproductive medicine and endocrinology. The evaluation and validation of the nomogram demonstrated that the model exhibits robust predictive performance for diseases. Furthermore, a GSEA enrichment analysis was performed, revealing that two biomarkers were predominantly enriched in pathways such as the B cell receptor pathway and the T cell receptor pathway. These pathways play a pivotal role in lymphocyte activation and immune response, underscoring the significance of B cell and T cell signaling in autoimmune and inflammatory diseases. We conducted drug prediction studies and identified that 23 common drugs and chemicals target two specific biomarkers. Through a detailed biomarkers-disease association analysis, we observed a stronger association between THBS1 and PCOS. These insights potentially offer fresh perspectives for the clinical diagnosis and treatment of PCOS. In practical applications, these findings are anticipated to propel the advancement of personalized medicine, thereby enhancing the early diagnosis rate and treatment efficacy of PCOS, ultimately improving the quality of life for patients. In addition, it is also necessary to explain the shortcomings of our research: for example, as we belong to bioinformatics analysis, we have already performed RT-qPCR, and in the future, we can study its mechanism action through immunohistochemistry and animal experiments.

Data Sharing Statement

The datasets analyzed for this study can be found in the [Gene Expression Omnibus (GEO) Database] [GSE80432 and GSE34526, <https://www.ncbi.nlm.nih.gov/geo/>], [Molecular Signatures Database (MSigDB)] [CRGs, <http://www.broadinstitute.org/gsea/msigdb/index.jsp>], [NetworkAnalyst] [<https://www.networkanalyst.ca/NetworkAnalyst/>], [ENCODE ChIP-seq database] [<https://www.encodeproject.org/>], [Comparative Toxicogenomics Database (CTD) Database] [<http://ctdbase.org/>].

Ethics Approval and Informed Consent

The study had the approval of the First People's Hospital of Yunnan Province ethics committee (approval number: KHLL2023-KY048) and conformed to national and international human research guidelines. All studies were carried out in accordance with the principles stipulated in the Declaration of Helsinki, and all participants were given informed consent.

Consent to Participate

All participants signed an informed written consent form to participate in the study.

Acknowledgments

We would like to express our sincere gratitude to all individuals and organizations who supported and assisted us throughout this research. Special thanks to the following funding institutions: Kunming University of Science and Technology Medical Joint Special Project and Technology, the National Health Commission Key Laboratory of Preconception Health Birth in Western China and the Yunnan Provincial Key Laboratory of Clinical Virology. In conclusion, we extend our thanks to everyone who has supported and assisted us along the way. Without your support, this research would not have been possible.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

The research reported in this project was generously supported by Kunming University of Science and Technology Medical Joint Special Project and Technology, the National Health Commission Key Laboratory of Preconception Health Birth in Western China and the Yunnan Provincial Key Laboratory of Clinical Virology under grant agreement number KUST-KH2022029Y, 2024XBYSKF015 and 2024XBYSKF016, 202205AG070061-BD-03. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Stener-Victorin E, Teede H, Norman RJ. et al. Polycystic ovary syndrome. *Nat Rev Dis Primers*. 2024;10(1):27. doi:10.1038/s41572-024-00511-3
2. Ge YJ, Xu W, Guan SM, Wang LN. Research progress in etiology and pathogenesis of polycystic ovary syndrome. *J Jilin Univ Med Ed*. 2024;50(01):288–294.
3. Lentscher JA, Decherney AH. Clinical Presentation and Diagnosis of Polycystic Ovarian Syndrome. *Clin Obstet Gynecol*. 2021;64(1):3–11. doi:10.1097/GRF.0000000000000563
4. Abbott DH, Hutcherson BA, Dumesic DA. Anti-Müllerian Hormone: a Molecular Key to Unlocking Polycystic Ovary Syndrome? *Semin Reprod Med*. 2024;42(1):41–48. doi:10.1055/s-0044-1787525
5. Xu Y, Zhang Z, Wang R, Xue S, Ying Q, Jin L. Roles of estrogen and its receptors in polycystic ovary syndrome. *Front Cell Dev Biol*. 2024;12:1395331. doi:10.3389/fcell.2024.1395331

6. Zhao H, Zhang J, Cheng X, Nie X, He B. Insulin resistance in polycystic ovary syndrome across various tissues: an updated review of pathogenesis, evaluation, and treatment. *J Ovarian Res.* 2023;16(1):9. doi:10.1186/s13048-022-01091-0
7. Liu L, Liu S, Bai F, Deng Y, Zhang X, Wang L. Investigating the Role of Inflammatory Response in Polycystic Ovary Syndrome Using Integrated RNA-Seq Analysis. *J Inflamm Res.* 2024;17:4701–4719. doi:10.2147/JIR.S460437
8. Xu X, Zhang X, Chen J, et al. Exploring the molecular mechanisms by which per- and polyfluoroalkyl substances induce polycystic ovary syndrome through in silico toxicogenomic data mining. *Ecotoxicol Environ Saf.* 2024;275:116251. doi:10.1016/j.ecoenv.2024.116251
9. Sun Y, Gao S, Ye C, Zhao W. Gut microbiota dysbiosis in polycystic ovary syndrome: mechanisms of progression and clinical applications. *Front Cell Infect Microbiol.* 2023;13:1142041. doi:10.3389/fcimb.2023.1142041
10. Petrie JR. Metformin beyond type 2 diabetes: emerging and potential new indications. *Diabetes Obes Metab.* 2024;26(3):31–41. doi:10.1111/dom.15756
11. de Athayde De Hollanda Morais B A, Martins Prizão V, de Moura de Souza M, et al. The efficacy and safety of GLP-1 agonists in PCOS women living with obesity in promoting weight loss and hormonal regulation: a meta-analysis of randomized controlled trials. *J Diabetes Complications.* 2024;38(10):108834. doi:10.1016/j.jdiacomp.2024.108834
12. Li X, Jiang B, Gao T, et al. Effects of inulin on intestinal flora and metabolism-related indicators in obese polycystic ovary syndrome patients. *Eur J Med Res.* 2024;29(1):443. doi:10.1186/s40001-024-02034-9
13. Hoeger KM. Exercise therapy in polycystic ovary syndrome. *Semin Reprod Med.* 2008;26(1):93–100. doi:10.1055/s-2007-992929
14. Benham JL, Corbett KS, Yamamoto JM, et al. Impact of bariatric surgery on anthropometric, metabolic, and reproductive outcomes in polycystic ovary syndrome: a systematic review and meta-analysis. *Obes Rev.* 2024;25(6):e13737. doi:10.1111/obr.13737
15. Taylor L, Wankell M, Saxena P, McFarlane C, Hebbard L. Cell adhesion an important determinant of myogenesis and satellite cell activity. *Biochim Biophys Acta mol Cell Res.* 2022;1869(2):119170. doi:10.1016/j.bbamcr.2021.119170
16. Lin W, Fang J, Wei S, et al. Extracellular vesicle-cell adhesion molecules in tumours: biofunctions and clinical applications. *Cell Commun Signal.* 2023;21(1):246. doi:10.1186/s12964-023-01236-8
17. Zhou W, Barton S, Cui J, et al. Infertile human endometrial organoid apical protein secretions are dysregulated and impair trophoblast progenitor cell adhesion. *Front Endocrinol.* 2022;13:1067648. doi:10.3389/fendo.2022.1067648
18. Pipek RP, Kloc M, Mizia P, Kubiak JZ. The Central Role of Cadherins in Gonad Development, Reproduction, and Fertility. *Int J mol Sci.* 2020;21(21):8264. doi:10.3390/ijms21218264
19. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47. doi:10.1093/nar/gkv007
20. Gustavsson EK, Zhang D, Reynolds RH, Garcia-Ruiz S, Ryten M. ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. *Bioinformatics.* 2022;38(15):3844–3846. doi:10.1093/bioinformatics/btac409
21. Liu T, Su X, Kong X, et al. Whole transcriptome sequencing identifies key lncRNAs, circRNAs, and mRNAs for exploring the pathogenesis and therapeutic target of mouse pneumoconiosis. *Gene.* 2024;901:148169. doi:10.1016/j.gene.2024.148169
22. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. *Omics.* 2012;16(5):284–287. doi:10.1089/omi.2011.0118
23. Liu L, Chandrashekar P, Zeng B, Sanderford MD, Kumar S, Gibson G. TreeMap: a structured approach to fine mapping of eQTL variants. *Bioinformatics.* 2021;37(8):1125–1134. doi:10.1093/bioinformatics/btaa927
24. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: network Analysis and Visualization of Proteomics Data. *J Proteome Res.* 2019;18(2):623–632. doi:10.1021/acs.jproteome.8b00702
25. Chen H, Zhang J, Sun X, Wang Y, Qian Y. Mitophagy-mediated molecular subtypes depict the hallmarks of the tumour metabolism and guide precision chemotherapy in pancreatic adenocarcinoma. *Front Cell Dev Biol.* 2022;10:901207. doi:10.3389/fcell.2022.901207
26. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenet.* 2019;11(1):123. doi:10.1186/s13148-019-0730-1
27. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf.* 2011;12(1):77. doi:10.1186/1471-2105-12-77
28. Liu TT, Li R, Huo C, et al. Identification of CDK2-Related Immune Forecast Model and ceRNA in Lung Adenocarcinoma, a Pan-Cancer Analysis. *Front Cell Dev Biol.* 2021;9:682002. doi:10.3389/fcell.2021.682002
29. Kasyanov ED, Yakovleva YV, Mudrakova TA, Kasyanova AA, Mazo GE. Comorbidity patterns and structure of depressive episodes in patients with bipolar disorder and major depressive disorder. *Zh Nevrol Psikhiatr Im S S Korsakova.* 2023;123(11. Vyp. 2):108–114. doi:10.17116/jnevro2023123112108
30. Jw L, Zhang Y, Ti L, et al. Identification of diagnostic genes for both Alzheimer's disease and Metabolic syndrome by the machine learning algorithm. *Front Immunol.* 2022;13:1037318. doi:10.3389/fimmu.2022.1037318
31. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* 2013;14(1):7. doi:10.1186/1471-2105-14-7
32. Xu P, Li S, Liu K, et al. Downregulation of dermatopontin in cholangiocarcinoma cells suppresses CCL19 secretion of macrophages and immune infiltration. *J Cancer Res Clin Oncol.* 2024;150(2):66. doi:10.1007/s00432-023-05532-1
33. Tseng WY, Stacey M, Lin HH. Role of Adhesion G Protein-Coupled Receptors in Immune Dysfunction and Disorder. *Int J mol Sci.* 2023;24(6):5499. doi:10.3390/ijms24065499
34. Sisto M, Ribatti D, Lisi S. Cadherin Signaling in Cancer and Autoimmune Diseases. *Int J mol Sci.* 2021;22(24):13358. doi:10.3390/ijms222413358
35. Janiszewska M, Primi MC, Izzard T. Cell adhesion in cancer: beyond the migration of single cells. *J Biol Chem.* 2020;295(8):2495–2505. doi:10.1074/jbc.REV119.007759
36. Chen J, Nekrasova OE, Patel DM, et al. The C-terminal unique region of desmoglein 2 inhibits its internalization via tail-tail interactions. *J Cell Biol.* 2012;199(4):699–711. doi:10.1083/jcb.201202105
37. Ishida-Yamamoto A, Igawa S, Kishibe M, Honma M. Clinical and molecular implications of structural changes to desmosomes and corneodesmosomes. *J Dermatol.* 2018;45(4):385–389. doi:10.1111/1346-8138.14202
38. Min KKM, Ffrench CB, McClure BJ, McClure BJ. Desmoglein-2 as a cancer modulator: friend or foe? *Front Oncol.* 2023;13:1327478. doi:10.3389/fonc.2023.1327478

39. Qin SH, Liao YD, Du QQ, et al. DSG2 expression is correlated with poor prognosis and promotes early-stage cervical cancer. *Cancer Cell Int.* 2020;20(1):206. doi:10.1186/s12935-020-01292-x
40. Meng HY, Liu JH, Qiu JN, et al. Identification of Key Genes in Association with Progression and Prognosis in Cervical Squamous Cell Carcinoma. *DNA Cell Biol.* 2020;39(5):848–863. doi:10.1089/dna.2019.5202
41. Jin RS, Wang XF, Zang RC, et al. Desmoglein-2 modulates tumor progression and osimertinib drug resistance through the EGFR/Src/PAK₁ pathway in lung adenocarcinoma. *Cancer Lett.* 2020;483:46–58. doi:10.1016/j.canlet.2020.04.001
42. Sun RY, Ma C, Wang W, Yang SY. Upregulation of desmoglein 2 and its clinical value in lung adenocarcinoma: a comprehensive analysis by multiple bioinformatics methods. *Peer J.* 2020;8:e8420. doi:10.7717/peerj.8420
43. Ormanns S, Altendorf-Hofmann A, Jackstadt R, et al. Desmogleins as prognostic biomarkers in resected pancreatic ductal adenocarcinoma. *Br J Cancer.* 2015;113(10):1460–1466. doi:10.1038/bjc.2015.362
44. Kamekura R, Kolegraff KN, Nava P, et al. Loss of the desmosomal cadherin desmoglein-2 suppresses colon cancer cell proliferation through EGFR signaling. *Oncogene.* 2014;33(36):4531–4536. doi:10.1038/onc.2013.442
45. Biedermann K, Vogelsang H, Becker I, et al. Desmoglein 2 is expressed abnormally rather than mutated in familial and sporadic gastric cancer. *J Pathol.* 2005;207(2):199–206. doi:10.1002/path.1821
46. Barber AG, Castillo-Martin M, Bonal DM, Rybicki BA, Christiano AM, Cordon-Cardo C. Characterization of desmoglein expression in the normal prostatic gland. Desmoglein 2 is an independent prognostic factor for aggressive prostate cancer. *PLoS One.* 2014;9(6):e98786. doi:10.1371/journal.pone.0098786
47. Chen L, Liu X, Zhang J, et al. Characterization of desmoglein 2 expression in ovarian serous tumors and its prognostic significance in high-grade serous carcinoma. *Int. J Clin Exp Pathol.* 2018;11(10):4977–4986.
48. Kim J, Beidler P, Wang H, et al. Desmoglein-2 as a prognostic and biomarker in ovarian cancer. *Cancer Biol Ther.* 2020;21(12):1154–1162. doi:10.1080/15384047.2020.1843323
49. Overmiller AM, McGuinn KP, Roberts BJ, et al. c-Src/Cav1-dependent activation of the EGFR by Dsg2. *Oncotarget.* 2016;7(25):37536–37555. doi:10.18632/oncotarget.7675
50. Lorch JH, Klessner J, Park JK, et al. Epidermal growth factor receptor inhibition promotes desmosome assembly and strengthens intercellular adhesion in squamous cell carcinoma cells. *J Biol Chem.* 2004;279(35):37191–37200. doi:10.1074/jbc.M405123200
51. Kamekura R, Nava P, Feng M, et al. Inflammation-induced desmoglein-2 ectodomain shedding compromises the mucosal barrier. *mol Biol Cell.* 2015;26(18):3165–3177. doi:10.1091/mbc.e15-03-0147
52. Klessner JL, Desai BV, Amargo EV, Getsios S, Green KJ. EGFR and ADAMs cooperate to regulate shedding and endocytic trafficking of the desmosomal cadherin desmoglein 2. *mol Biol Cell.* 2009;20(1):328–337. doi:10.1091/mbc.e08-04-0356
53. Bech-Serra JJ, Santiago-Josefat B, Esselens C, et al. Proteomic identification of desmoglein 2 and activated leukocyte cell adhesion molecule as substrates of ADAM17 and ADAM10 by difference gel electrophoresis. *mol Cell Biol.* 2006;26(13):5086–5095.
54. Santiago-Josefat B, Esselens C, Bech-Serra JJ, Arribas J. Post-transcriptional up-regulation of ADAM17 upon epidermal growth factor receptor activation and in breast tumors. *J Biol Chem.* 2007;282(11):8325–8331. doi:10.1074/jbc.M608826200
55. Nava P, Laukoetter MG, Hopkins AM, et al. Desmoglein-2: a novel regulator of apoptosis in the intestinal epithelium. *mol Biol Cell.* 2007;18(11):4565–4578. doi:10.1091/mbc.e07-05-0426
56. Deng LY, Zeng XF, Tang D, Deng W, Liu HF, Xie YK. Expression and prognostic significance of thrombospondin gene family in gastric cancer. *J Gastrointest Oncol.* 2021;12(2):355–364. doi:10.21037/jgo-21-54
57. Bender HR, Campbell GE, Aytoda P, Mathiesen AH, Duffy DM. Thrombospondin 1 (THBS1) promotes follicular angiogenesis, luteinization, and ovulation in primates. *Front Endocrinol.* 2019;10:727. doi:10.3389/fendo.2019.00727
58. von Toerne C, Huth C, de Las Heras Gala T, et al. MASPL1, THBS1, GPLD1 and ApoA-IV are novel biomarkers associated with prediabetes: the KORA F4 study. *Diabetologia.* 2016;59(9):1882–1892. doi:10.1007/s00125-016-4024-2
59. Zanjirband M, Hodayi R, Safaiejad Z, Nasr-Esfahani MH, Ghaedi-Heydari R. Evaluation of the p53 pathway in polycystic ovarian syndrome pathogenesis and apoptosis enhancement in human granulosa cells through transcriptome data analysis. *Sci Rep.* 2023;13(1):11648. doi:10.1038/s41598-023-38340-1
60. Nasri F, Doroudchi M, Jahromi BN, Ghahesi-Fard B. T helper cells profile and CD4+CD25+Foxp3+ regulatory T cells in polycystic ovary syndrome. *Iran J Immunol.* 2018;15(3):165–174. doi:10.22034/IJI.2018.39386
61. Li Z, Peng A, Feng Y, et al. Detection of T lymphocyte subsets and related functional molecules in follicular fluid of patients with polycystic ovary syndrome. *Sci Rep.* 2019;9(1):6040. doi:10.1038/s41598-019-42631-x
62. Nasri F, Zare M, Doroudchi M, Ghahesi-Fard B. Proteome analysis of CD4+ T cells reveals differentially expressed proteins in infertile polycystic ovary syndrome patients. *Endocr Metab Immune Disord Drug Targets.* 2021;21(11):1998–2004. doi:10.2174/1871530320666201119152323
63. Dai M, Hong L, Yin T, Liu S. Disturbed follicular microenvironment in polycystic ovary syndrome: relationship to oocyte quality and infertility. *Endocrinology.* 2024;165(4):bqae023. doi:10.1210/endo/bqae023
64. Wang T, Xiong X, Xiao N, et al. The therapeutic effect of anti-CD19 antibody on DHEA-induced PCOS mice. *Int Immunopharmacol.* 2024;130:111711. doi:10.1016/j.intimp.2024.111711
65. Nan X, Lu H, Wu J, Xue M, Guo W, Wang X. Prevalence, knowledge and education level associated with secondhand smoke exposure among never-smoking women in Inner Mongolia, Northern China. *Tob Induc Dis.* 2020;18:35. doi:10.18332/tid/119162
66. Li Y, Hecht SS. Carcinogenic components of tobacco and tobacco smoke: a 2022 update. *Food Chem Toxicol.* 2022;165:113179. doi:10.1016/j.fct.2022.113179
67. Biegon A, Alia-Klein N, Fowler JS. Potential contribution of aromatase inhibition to the effects of nicotine and related compounds on the brain. *Front Pharmacol.* 2012;3:185. doi:10.3389/fphar.2012.00185
68. Boon WC, Chow JDY, Simpson ER. The multiple roles of estrogens and the enzyme aromatase. *Prog Brain Res.* 2010;181:209–232.
69. Li F, Wang Y, Xu M, et al. Single-nucleus RNA Sequencing reveals the mechanism of cigarette smoke exposure on diminished ovarian reserve in mice. *Ecotoxicol Environ Saf.* 2022;245:114093. doi:10.1016/j.ecoenv.2022.114093
70. Li S, Feng A, Peng Y, et al. Association between secondhand smoke exposure and serum sex hormone concentrations among US female adults: a cross-sectional analysis using data from the National Health and Nutrition Examination Survey, 2013–2016. *BMJ Open.* 2024;14(5):e073527. doi:10.1136/bmjopen-2023-073527

71. Jiang Z, He R, Wu H, et al. The causal association between smoking initiation, alcohol and coffee consumption, and women's reproductive health: a two-sample Mendelian randomization analysis. *Front Genet.* 2023;14:1098616. doi:10.3389/fgene.2023.1098616
72. Piatti PM, Setola E, Galluccio E, et al. Smoking is associated with impaired glucose regulation and a decrease in insulin sensitivity and the disposition index in first-degree relatives of type 2 diabetes subjects independently of the presence of metabolic syndrome. *Acta Diabetol.* 2014;51(5):793–799. doi:10.1007/s00592-014-0599-6
73. Jia WP. The impact of cigarette smoking on metabolic syndrome. *Biomed Environ Sci.* 2013;26(12):947–952. doi:10.3967/bes2013.029
74. Kawakami Y, Fujii S, Ishikawa G, Sekiguchi A, Nakai A, Takase M. Valproate-Induced Polycystic Ovary Syndrome in a Girl with Epilepsy: a Case Study. *J Nippon Med Sch.* 2018;85(5):287–290. doi:10.1272/jnms.JNMS.2018_85-46
75. Li S, Qi J, Sun Y, Gao X, Ma J, Zhao S. An integrated RNA-Seq and network study reveals that valproate inhibited progesterone production in human granulosa cells. *J Steroid Biochem mol Biol.* 2021;214:105991. doi:10.1016/j.jsbmb.2021.105991
76. Hu X, Wang J, Dong W, Fang Q, Hu L, Liu C. A meta-analysis of polycystic ovary syndrome in women taking valproate for epilepsy. *Epilepsy Res.* 2011;97(1–2):73–82. doi:10.1016/j.eplepsyres.2011.07.006

International Journal of General Medicine

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>

Dovepress

Taylor & Francis Group