Xu et al. Journal of Ovarian Research (2026) 19:46
https://doi.org/10.1186/s13048-025-01956-0

Journal of Ovarian Research

RESEARCH Open Access

Interpreting the molecular and cellular landscape of PCOS through bulk transcriptomics, single-cell transcriptomics and machine learning

Kangjie Xu[1,2*†], Shuyun Zhang[3†], Lijuan Guo[4], Tongtong Liu[4], Ying Li[4] and Yanhua Zhang[4*]

Abstract

Background Polycystic ovary syndrome (PCOS) is a prevalent endocrine-metabolic disorder in women, hallmarked by hyperandrogenism, anovulation, and polycystic ovarian morphology. This study integrates multi-omics and machine-learning analyses to elucidate the molecular mechanisms and cellular constituents underlying PCOS, aiming to uncover potential therapeutic targets and enhance diagnostic precision.

Methods Bulk and single-cell RNA sequencing identified key granulosa cell subpopulations and gene expression patterns in PCOS; subsequently, machine-learning algorithms were applied to construct a diagnostic model and to screen for key gene signatures. Consequently, the identified signatures were validated at both mRNA and protein levels in independent clinical samples using qPCR and western blotting.

Results Compared with controls, PCOS patients exhibited a markedly increased proportion of the GC9 granulosa cell subset, which displayed an active proliferative phenotype. Up-regulated genes in PCOS were closely associated with immune function, responsiveness to stimuli, and diverse cellular biological processes. Machine-learning analysis further pinpointed a three-gene signature-comprising HLA-DRA, SRM, and CTSL-and yielded a diagnostic model with superior accuracy and specificity. Moreover, validation in clinical samples confirmed significant up-regulation of HLA-DRA, SRM, and CTSL at both mRNA and protein levels in follicular cells of PCOS patients.

Conclusions Our findings delineate a previously unrecognized cellular landscape and gene signature associated with PCOS, thereby proposing novel diagnostic and therapeutic targets.

Keywords Polycystic ovary syndrome, Single-cell transcriptomics, Bulk transcriptomics, Machine learning

†Kangjie Xu and Shuyun Zhang contributed equally to this work.

*Correspondence:
Kangjie Xu
839974461@qq.com
Yanhua Zhang
18066132759@163.com

[1]Medical College, Institute of Translational Medicine, Yangzhou University, Yangzhou, Jiangsu Province 225009, PR China
[2]Department of Central Laboratory, Binhai County People's Hospital, Yancheng, Jiangsu Province 224000, PR China
[3]Department of Obstetrics and Gynecology, Second Affiliated Hospital of Soochow University, Suzhou, Jiangsu Province 215000, PR China
[4]Department of Obstetrics and Gynecology, Binhai County People's Hospital, Yancheng, Jiangsu Province 224000, PR China

Introduction

Polycystic ovary syndrome (PCOS) is a prevalent endocrine disorder that affects approximately 6–10% of reproductive-aged women worldwide. Its cardinal features encompass menstrual irregularities, hyperandrogenism, and polycystic ovarian morphology, thereby predisposing patients to infertility and metabolic syndrome [1]. Although the pathogenesis of PCOS remains intricate, it is generally accepted that the interplay among genetic, environmental, and endocrine factors collectively drives disease progression [2]. PCOS is characterized by dysfunction across multiple cell types-including granulosa cells, endometrium endothelial cells, and immune cells-highlighting the necessity of dissecting cellular heterogeneity and gene-expression profiles [3]. In recent years, advances in transcriptomics and machine learning have therefore provided powerful tools to elucidate the molecular and cellular underpinnings of PCOS, thus offering novel insights for early diagnosis and individualized therapy.

Single-cell transcriptomics, uniquely resolves intercellular heterogeneity and cell-type-specific expression patterns, thereby furnishing an unprecedented perspective on disease mechanisms [4]. Through single-cell RNA sequencing, the transcriptional landscape of PCOS ovarian tissue can be interrogated at single-cell resolution, thereby enabling the identification of disease-relevant cell subsets and their functional states [5]. For example, studies have reported that granulosa cells derived from PCOS patients exhibit pronounced alterations in inflammatory responses and lipid metabolism, changes that may therefore be intimately linked to the pathophysiology of PCOS [6]. Moreover, by delineating intercellular communication networks via single-cell analyses, researchers can gain a more comprehensive understanding of the complex biology underlying PCOS [7]. Hence, the integration of single-cell transcriptomic data with conventional bulk RNA-seq datasets allows for a holistic interpretation of PCOS molecular mechanisms and consequently provides a robust foundation for future therapeutic strategies [8].

Machine-learning algorithms-including support-vector machines (SVM), random forests (RF), and least absolute shrinkage and selection operator (LASSO) regression-are capable of processing heterogeneous disease datasets and thereby revealing latent biological patterns. These methods extract informative features from high-dimensional gene-expression data, identify key genes, and construct predictive models. For instance, SVM has been effectively applied to the classification of PCOS, consequently reducing the under-diagnosis rate of hyperandrogenic PCOS [9]. RF, on the other hand, accommodates non-linear relationships and interaction effects among features and has therefore been widely adopted for both classification and regression tasks; indeed, Wang et al. employed RF to uncover associations between follicular-fluid metal exposure and PCOS risk [10]. Moreover, LASSO regression has been utilized to reduce dimensionality and to pinpoint gene signatures relevant to PCOS [11]. Hence, the integration of such algorithms enables the extraction of biologically meaningful information from complex transcriptomic datasets, thereby advancing personalized therapeutic strategies and deepening our understanding of PCOS pathophysiology [12].

In the present study, single-cell sequencing revealed marked cellular heterogeneity in PCOS and specifically identified the GC9 granulosa-cell subset as exhibiting an actively proliferative phenotype. Immune-infiltration analyses further demonstrated distinct immune-cell compositions between normal controls and PCOS patients, as well as among individual PCOS patients. Additionally, by employing multiple machine-learning approaches, we derived a three-gene diagnostic signature comprising HLA-DRA, SRM, and CTSL. Collectively, these findings delineate the molecular mechanisms and cellular landscape of PCOS, highlight several pivotal genes and cell subpopulations, and therefore provide a robust theoretical basis for early diagnosis and targeted intervention. Future investigations should, however, validate the clinical utility of these discoveries.

Method

Data acquisition

To comprehensively delineate the transcriptional landscape of PCOS, bulk RNA-seq datasets were retrieved from the following repositories: GSE34526 contains microarray data from granulosa cells of 3 healthy controls and 7 PCOS women; GSE293353 comprises granulosa-cell profiles from 9 controls and 9 PCOS patients. Consequently, three additional bulk datasets were employed as external validation cohorts to confirm both the expression of key genes and the diagnostic efficacy of the derived signatures-namely, GSE138518 (3 controls vs. 3 PCOS), GSE193123 (3 controls vs. 3 PCOS), and GSE155489 (10 controls vs. 10 PCOS). Gene-expression matrices were normalized using the normalizeBetweenArrays function, thereby mitigating inter-sample technical variability. However, residual batch effects across datasets were further attenuated by applying the ComBat algorithm. Single-cell RNA-seq data were obtained from GSE240688, encompassing 3 normal and 3 PCOS samples; these data were therefore integrated to resolve cell-type-specific transcriptional alterations in PCOS.

Data processing

Single-cell datasets were processed with the Seurat package. After initial filtering-retaining genes expressed in ≥ 3 cells and cells expressing ≥ 200 genes-stringent quality control was applied. Specifically, cells were excluded if

their gene counts were outside 500-4,000 or if their mitochondrial proportion exceeded 15%. Data were then normalized via NormalizeData, and the 2,000 most variable features were identified with FindVariableFeatures. These features were scaled using ScaleData, subjected to principal component analysis, and corrected for batch effects with Harmony. Consequently, the first nine harmonized principal components were used to construct a K-nearest-neighbor graph, followed by clustering and uniform manifold approximation and projection (UMAP) visualization. This pipeline therefore enables delineation of cellular heterogeneity and latent cell states.

Differential expression and functional enrichment

For bulk RNA-seq data, differential expression was determined with the limma package: lmFit fitted the linear model, eBayes performed empirical Bayes moderation, and resulting statistics were extracted for all genes. For single-cell data, FindMarkers identified genes differentially expressed between control and PCOS cells, requiring detection in ≥ 10% of cells and a log-fold change ≥ 0.15. Shared up-regulated genes were visualized with ggvenn. Pathway and process enrichment were performed with Metascape (https://metascape.org/), and enrichment networks were visualized in Cytoscape. Protein-protein interaction (PPI) networks were subsequently analyzed, and densely connected modules were extracted via the MCODE algorithm.

Diagnostic model construction and validation

Five machine-learning algorithms-LASSO, SVM-RFE, Random Forest, XGBoost, and Boruta-were employed to select signature genes and build diagnostic models. Nomogram was constructed to visualize predicted disease status from signature-gene expression. Model calibration, assessed with the calibrate function, evaluated concordance between predicted and observed outcomes. Classification performance was quantified by receiver operating characteristic (ROC) curves and area under the curve (AUC) values computed with the pROC package.

Immune-infiltration analysis

Single-sample gene set enrichment analysis (ssGSEA) quantified immune-cell infiltration; ssGSEA scores were subsequently normalized for cross-sample comparison. Spearman correlation coefficients were calculated between signature-gene expression and the relative abundance of each immune subset. Differences in immune-cell proportions across groups were depicted with ggboxplot, whereas "lollipop" plots illustrated correlations between immune-cell infiltration and signature-gene expression.

Isolation of human granulosa cells

Human granulosa cells were obtained from three PCOS patients and three women with tubal-factor infertility undergoing IVF/ICSI-ET. All participants provided written informed consent, and every procedure strictly adhered to the ethical principles of the Declaration of Helsinki. The study protocol was approved by the Ethics Committee of Binhai County People's Hospital (No. 2025-BYKYLL-027). PCOS inclusion criteria were as follows: (1) age 18–35 years; (2) scheduled for IVF/ICSI-ET at the Reproductive Medicine Department of Binhai County People's Hospital; (3) fulfillment of both the 2018 Chinese and the 2003 Rotterdam PCOS diagnostic criteria, including polycystic ovarian morphology, oligo- or anovulation, and biochemical/clinical hyperandrogenism. Controls were recruited from women undergoing IVF/ICSI-ET solely because of tubal or male-factor infertility, who exhibited normal ovarian reserve (FSH < 10 IU/L, AMH > 1.1 ng/mL) under either GnRH-agonist or antagonist protocols. Common exclusion criteria for both groups comprised additional reproductive disorders, systemic diseases affecting gestation, and any relevant chronic medical history. Granulosa cells were isolated from discarded follicular fluid collected after transvaginal ultrasound-guided oocyte retrieval. The fluid was centrifuged at 2,000 rpm for 10 min, the pellet was resuspended in culture medium, layered onto Ficoll solution (Solarbio, China), and centrifuged again; the resulting interphase containing granulosa cells was harvested for subsequent RNA extraction.

Real-time quantitative PCR (RT-qPCR)

Total RNA was extracted from granulosa cells with TRIzol reagent (Invitrogen, USA) according to the manufacturer's instructions. RNA concentration and purity were subsequently assessed with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA); only samples exhibiting an A260/A280 ratio of 1.8–2.0 were accepted. Thereafter, 1 μg of total RNA was reverse-transcribed into cDNA with the SuperScript VILO cDNA Synthesis Kit (Invitrogen, USA) in a 20-μL reaction (25 °C, 10 min; 42 °C, 60 min; 85 °C, 5 min to terminate). Quantitative PCR was performed on an ABI 7500 Real-Time PCR system (Applied Biosystems, USA) with UltraSYBR Mixture (Cwbio, China). Each 20-μL reaction contained 2 μL cDNA, 0.4 μL of each 10 μM primer, 10 μL UltraSYBR Mixture, and nuclease-free water to volume. The thermal profile consisted of an initial denaturation at 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. All samples were run in triplicate, and the entire experiment was independently repeated three times. Relative expression of the target genes (HLA-DRA, SRM, CTSL) was calculated by the

2^(−ΔΔCt) method with GAPDH as the endogenous control. Primer sequences are listed below:

GAPDH forward
5′-GTCTCCTCTGACTTCAACAGCG-3′, reverse
5′-ACCACCCTGTTGCTGTAGCCAA-3′;
HLA-DRA forward
5′-GATGATCTTGGGTTGGAA-3′, reverse
5′-TGGGAGTTTGATGCTCCA-3′;
SRM forward
5′-ATGGAGCCGGCCCGGCCGACGG-3′, reverse
5′-GCTGAAGTACAACTCCGACGTGC-3′;
CTSL forward
5′-TTGCCTCAGCTACTCTAACA-3′, reverse
5′-CTGAGTGAGCAGAATCTGGT-3′.

Western blot (WB)

Granulosa-cell total protein was extracted with RIPA lysis buffer containing 1% PMSF and 1% protease inhibitor cocktail. After 30 min on ice, lysates were centrifuged at 12 000 rpm for 15 min at 4 °C, and the supernatants were collected. Protein concentrations were determined with a BCA assay kit (Beyotime, China); samples were equalized, mixed with 5× loading buffer, and denatured at 95 °C for 5 min. Subsequently, 30 μg of protein per lane was resolved on 10% SDS-PAGE gels (80 V, 30 min; 120 V, 1 h) and electro-transferred to PVDF membranes (Millipore, USA) at 250 mA for 90 min. Membranes were blocked with 5% non-fat milk in TBST for 1 h at room temperature, then incubated overnight at 4 °C with primary antibodies against HLA-DRA (1:1000, Invitrogen, USA), SRM (1:1000, Invitrogen, USA), CTSL (1:1000, Abcam, UK), GAPDH (1:1000, CST, USA), or β-Actin (1:1000, CST, USA). After three 10-min washes with TBST, membranes were probed with the corresponding HRP-conjugated secondary antibodies (1:5000, CST, USA) for 1 h at room temperature, followed by three additional TBST washes. Immunoreactive bands were visualized with ECL reagent (Millipore, USA) and captured with a ChemiDoc XRS + system (Bio-Rad, USA). Band intensities were quantified with ImageJ and normalized to GAPDH or β-Actin. Each experiment was independently performed three times.

Statistical analysis

All statistical analyses were conducted in R 4.2.1; a two-sided P value < 0.05 was considered statistically significant.

Results

Single-cell transcriptomic landscape of PCOS

In this study, single-cell RNA sequencing was employed to delineate the cellular heterogeneity underlying PCOS. Consequently, 6 samples-3 PCOS and 3 matched controls-were retrieved from the GSE240688 dataset. Preliminary quality control revealed comparable numbers of detected genes across all samples (Fig. 1A). Unsupervised UMAP clustering resolved nine transcriptionally distinct cell populations (Fig. 1B). Clusters 0–8 exhibited high expression of canonical granulosa-cell markers (CYP19A1, SERPINE2, FST) [13], whereas cluster 9 uniquely expressed monocyte-specific genes (S100A9, CXCR2, CSF3R) [14] (Fig. 1C). Hence, clusters 0–8 were annotated as granulosa cells and cluster 9 as monocytes (Fig. 1D).

Next, we quantified the relative abundance of each cell type across conditions. Notably, the proportions of granulosa sub-clusters GC6 and GC9 were markedly elevated in PCOS, whereas GC3 and GC8 were significantly reduced (Fig. 1E). In addition, the monocyte fraction was substantially expanded in PCOS samples, suggesting a potential chronic inflammatory milieu [15]. UMAP embeddings further underscored pronounced compositional differences between control and PCOS groups (Fig. 1F).

Intercellular communication networks, inferred via CellChat, revealed extensive ligand-receptor interactions among granulosa sub-clusters-excepting GC8-thereby underscoring their coordinated functional state (Fig. 1G). Pseudo-time reconstruction with Monocle 3 positioned GC9 at the trajectory origin, indicating a possible stem-like or progenitor phenotype (Fig. 1H).

Finally, GSVA disclosed pathway enrichment patterns unique to each sub-population (Fig. 1I). GC6 was enriched in the WNT/β-catenin axis, implicating roles in tissue morphogenesis and proliferation. GC7 exhibited signatures of myogenesis, bile-acid metabolism, and epithelial-mesenchymal transition. GC9 displayed activation of E2F targets, G2/M checkpoint, mitotic-spindle, and DNA-repair pathways, hence suggesting a highly proliferative, cell-cycle-active state. Monocytes, however, were dominated by inflammatory cascades-IL6-JAK-STAT3, IL2-STAT5, interferon-γ, TNF-α via NF-κB, interferon-α, allograft rejection, and KRAS signaling-thereby confirming immune activation. Collectively, these data therefore provide single-cell resolution insights into the molecular and functional heterogeneity that underpins PCOS.

Identification and functional interrogation of key PCOS-related differential genes

To systematically uncover pivotal genes and their biological roles in PCOS, we first interrogated the bulk transcriptome dataset GSE34526, thereby identifying 1,169 up-regulated genes (Fig. 2A, logFC > 1, pvalue < 0.05). Subsequently, single-cell differential analysis via the FindMarkers function yielded 1,363 up-regulated genes in PCOS. The intersection of these two gene sets therefore comprised 55 co-up-regulated genes (Fig. 2B). Functional

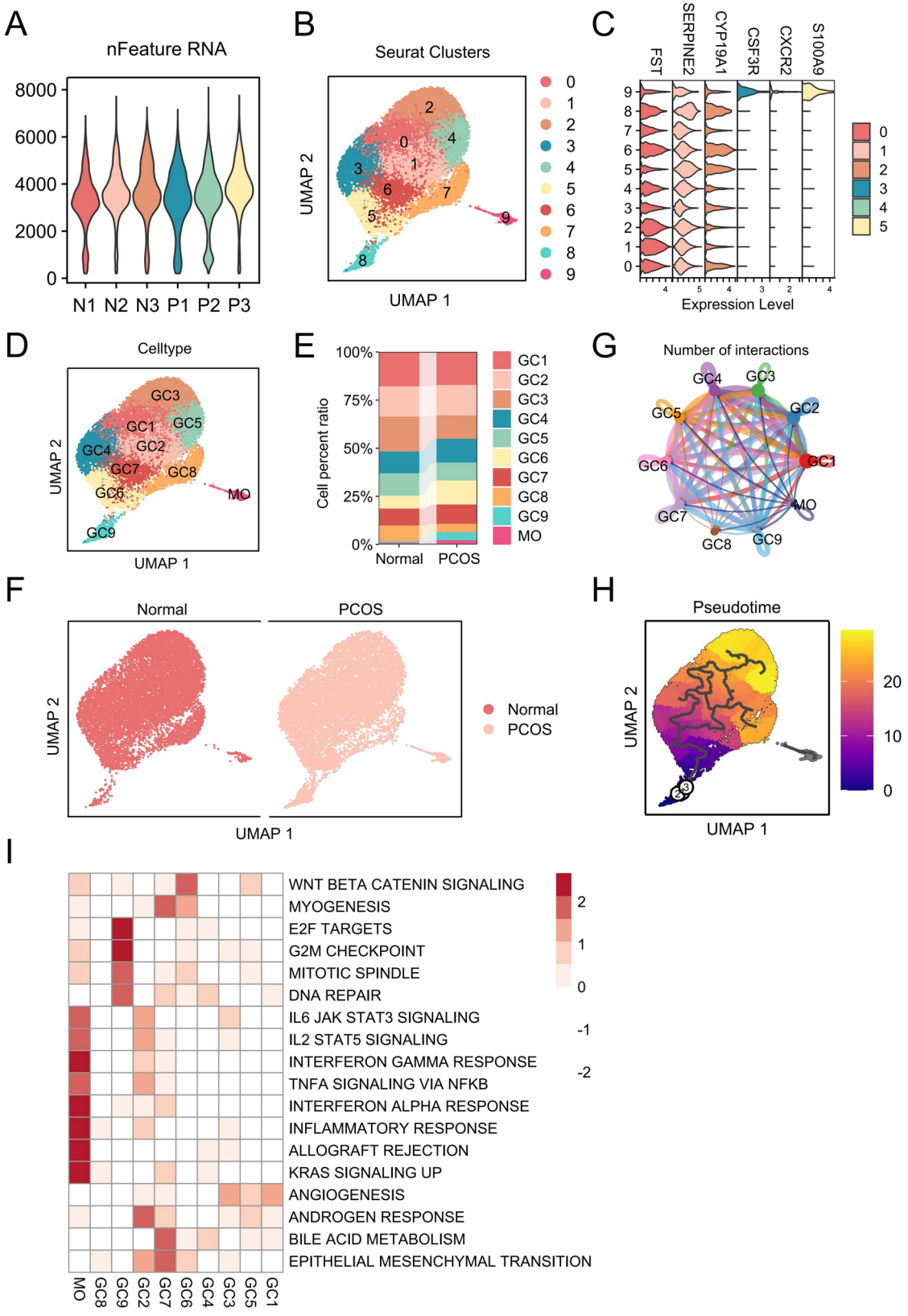

Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Single-cell sequencing analysis of PCOS and normal groups (**A**) Number of genes detected across different sample groups. Each violin plot represents one sample (N1-N3/P1-P3), illustrating the distribution characteristics of cell-detected genes within that sample (**B**) UMAP visualization of cell clusters. Different colors represent distinct clusters (0–9), each corresponding to a different cell type (**C**) Expression levels of signature genes across clusters. Each column represents one signature gene, and each row represents one cluster; violin plots display gene expression distribution within the corresponding cluster, with colors indicating cluster numbers (**D**) Distribution of different cell types in the UMAP dimensionality reduction plot. Different colors represent different cell types (GC1-GC9, MO) (**E**) Stacked bar chart showing the proportion of different cell types in the normal and PCOS groups (**F**) Distribution of the normal and PCOS groups in the UMAP dimensionality reduction plot. Different colors represent different groups: red indicates the normal group, and pink indicates the PCOS group (**G**) Interaction network among different cell types. Lines of different colors represent interactions between different cell types, with line thickness indicating interaction strength (**H**) Pseudo-time distribution of different cell types. Yellow indicates high pseudo-time, purple indicates low pseudo-time, reflecting the temporal sequence of cells in the differentiation process (**I**) Significantly enriched signaling pathways across different cell types. Rows represent signaling pathways, columns represent cell types, and color intensity indicates enrichment level—red denotes high enrichment, white denotes low enrichment

enrichment revealed that these convergent genes were significantly enriched in processes such as immune-system response, stimulus responsiveness, multicellular organismal processes, cellular processes, and developmental programs (Fig. 2C). The corresponding PPI network further illustrated their intricate cross-talk across these biological domains (Fig. 2D). Moreover, network visualization highlighted additional pathways-including antigen processing and presentation, cytokine-mediated signaling, lumen morphogenesis, negative regulation of immune processes, and steroid-hormone responsiveness-thereby underscoring a highly interconnected regulatory landscape (Fig. 2E). Collectively, these findings not only illuminate the activated molecular circuitry in PCOS but also nominate potential therapeutic targets.

Construction of a PCOS diagnostic model

Five machine-learning algorithms were next employed to distill diagnostic gene signatures. Specifically, LASSO regression selected 9 features (Fig. 3A-B), whereas SVM-RFE achieved optimal accuracy (0.82) with 7 genes (Fig. 3C-D). Random Forest identified 9 influential predictors (Fig. 3E), XGBoost ranked the top 10 (Fig. 3F), and Boruta retained 14 (Fig. 3G-H). A Venn intersection consequently converged on three consensus genes-HLA-DRA, SRM, and CTSL-as the core diagnostic triad (Fig. 3I). A nomogram integrating these genes was constructed to estimate individual PCOS probability (Fig. 4A). Bootstrap calibration (1,000 iterations) demonstrated excellent agreement between predicted and observed probabilities (Fig. 4B), and the ROC curve confirmed robust discriminatory performance (Fig. 4C). Furthermore, individual ROC analyses revealed strong diagnostic efficacy for each of the three genes (Fig. 4D-F).

External validation of the PCOS diagnostic model

To appraise model robustness, we externally validated CTSL, HLA-DRA, and SRM expression in two independent bulk datasets. In GSE138518, all three genes were significantly up-regulated in PCOS (Fig. 5A-C; all $P<0.05$). Analogous results were observed in GSE155489 (Fig. 5D-F; all $P<0.05$). Consequently, the combined validation cohort yielded AUCs of 0.698 for CTSL, 0.846 for HLA-DRA, and 0.805 for SRM (Fig. 5G-I), thereby confirming their reliable diagnostic utility. These data therefore establish CTSL, HLA-DRA, and SRM as trans-cohort biomarkers for PCOS.

Immune-infiltration landscape in PCOS

Prior evidence implicates immune dysregulation in PCOS; consequently, we interrogated the immune-cell composition and its association with the three core diagnostic genes. Compared with controls, PCOS samples exhibited markedly elevated proportions of monocytes, naïve B cells, and M2 macrophages, thereby corroborating our single-cell findings. Conversely, memory-resting $CD4^+$ T cells and activated mast cells were significantly reduced (Fig. 6A; all $P<0.05$).

Moreover, the expression levels of the diagnostic genes displayed distinct correlations with immune subsets. CTSL expression was positively associated with M2 macrophages, eosinophils, and activated $CD4^+$ memory T cells, whereas it was inversely correlated with M0 macrophages, plasma cells, and resting $CD4^+$ memory T cells (Fig. 6B). HLA-DRA exhibited a positive correlation with plasma cells, M0 macrophages, and resting $CD4^+$ memory T cells, but a negative correlation with T follicular helper cells, activated $CD4^+$ memory T cells, and eosinophils (Fig. 6C). Similarly, SRM expression aligned positively with M2 macrophages, naïve B cells, and eosinophils, yet negatively with plasma cells, M0 macrophages, and resting $CD4^+$ memory T cells (Fig. 6D). Collectively, these data indicate that CTSL, HLA-DRA, and SRM may modulate PCOS pathogenesis by shaping immune-cell infiltration patterns.

Identification of PCOS subtypes and immune-infiltration divergence

To dissect patient heterogeneity, consensus clustering was performed, yielding two robust subgroups (C1 and C2) at $k=2$, as evidenced by a dark consensus matrix (Fig. 7A), stable cumulative distribution functions across multiple indices (Fig. 7B), and a maximal delta area at

k = 2 (Fig. 7C). Principal-component analysis further separated C1 and C2 along PC1 and PC2 axes (Fig. 7D).

Subsequent analyses revealed marked disparities in the infiltration of distinct immune cell subsets between clusters C1 and C2 (Fig. 7E; all $p < 0.05$). Cluster C1 was significantly enriched for activated dendritic cells (DCs), myeloid-derived suppressor cells (MDSCs), macrophages, Th1/Th17 cells, and multiple effector NK subsets, implying an "acute/effector-immune" dominant milieu. Conversely, these populations were markedly reduced in C2, suggesting either a lowered activation threshold or a shift toward an "exhausted/tolerogenic" state. Previous studies have demonstrated that sustained hyperandrogenism or danger-associated signals within the PCOS follicular niche can trigger DC maturation and IL-12 secretion, thereby driving Th1/Th17 polarization and recruiting NK cells to establish a local pro-inflammatory feedback loop [16]. Therefore, C1 likely represents an "immune-hyperactivated" PCOS endotype characterized by massive effector-cell infiltration, whereas C2 denotes an "immune-hypoactivated/tolerogenic" profile. Future studies must determine whether C2 corresponds to immune dysregulation or adverse long-term reproductive outcomes, hence providing a stratification framework for immune-targeted interventions and validating its clinical utility.

Validation of CTSL, HLA-DRA, and SRM up-regulation by qRT-PCR and Western blot

Finally, to corroborate our in silico findings, we quantified mRNA and protein levels in granulosa cells from three PCOS patients and three controls. qRT-PCR demonstrated significantly elevated mRNA expression of CTSL, HLA-DRA, and SRM in PCOS (Fig. 8A-C). Western blot analyses further confirmed marked up-regulation of the corresponding proteins (Fig. 8D-F). Consequently, these results substantiate the pathogenic relevance of CTSL, HLA-DRA, and SRM in PCOS and pave the way for future investigations into their mechanistic roles and clinical utility as biomarkers.

Discussion

PCOS is a prevalent endocrine-metabolic disorder that affects women of reproductive age worldwide [17]. Although its cardinal manifestations-menstrual irregularity, hyperandrogenism, and polycystic ovarian morphology-are well recognized, the clinical phenotype remains markedly heterogeneous [18]. Beyond reproductive dysfunction, PCOS is associated with long-term metabolic derangements, cardiovascular disease, and malignancy [19]. Insulin resistance, chronic low-grade inflammation, and genetic susceptibility have been proposed as contributory mechanisms [20]; however, a comprehensive understanding of the precise pathophysiology is still lacking. Consequently, a deeper dissection of disease mechanisms is imperative to refine diagnostic accuracy and therapeutic strategies.

Single-cell transcriptomics enables high-resolution interrogation of gene expression at the individual-cell level, thereby revealing cellular heterogeneity and lineage-specific transcriptional signatures [21]. Recent single-cell investigations of PCOS have yielded mechanistic insights, Ye et al. demonstrated that glutathione peroxidase 4 (GPX4) deficiency accelerates endometrial fibrosis via TGF-β1/Smad2/3-mediated collagen deposition [8], whereas Wei et al. uncovered angiotensin II-IRS1/PI3K/AKT crosstalk in murine PCOS models [22]. Furthermore, integrated Mendelian randomization and single-cell analyses have identified protective loci and key fibroblast subsets linked to PCOS risk [23]. In the present study, we delineated ovarian cellular taxonomy and identified a hyper-proliferative granulosa-cell subpopulation characterized by activation of E2F targets, G2/M checkpoint, mitotic-spindle, and DNA-repair pathways-suggesting heightened cell-cycle activity. Therefore, single-cell transcriptomics not only clarifies subpopulation dynamics but also illuminates previously unrecognized molecular circuits that may propel PCOS progression.

Although single-cell approaches offer unparalleled resolution, bulk RNA-seq datasets provide larger patient cohorts and richer clinical annotation [24]. By integrating bulk transcriptomes, we examined granulosa-cell gene-expression patterns in follicular fluid and delineated disease-relevant pathways. Immune-related cascades were prominently up-regulated in PCOS, consistent with prior evidence of immune dysregulation [16]. Hyperandrogenism may exert partial immunosuppressive effects by attenuating immune-cell activation [25]; however, estrogen-mediated modulation can counteract this protection [26]. Indeed, estrogen stimulates cytokine secretion and immune activation, whereas progesterone exerts inhibitory effects [27]. Consequently, the relative progesterone deficiency observed in PCOS patients may permit unchecked immune stimulation. Additionally, enriched pathways related to stimulus responsiveness imply heightened sensitivity to endogenous and exogenous cues [28]. Metabolic and hormonal perturbations-such as insulin resistance and hyperandrogenemia-trigger aberrant responses in adipose, skeletal muscle, and hepatic tissues, ultimately culminating in systemic metabolic dysfunction [29]. Hence, detailed characterization of gene and pathway activation states may elucidate intricate pathogenic networks and inform targeted therapeutic interventions for PCOS.

By employing machine-learning algorithms-such as LASSO regression and random forest-it is possible to extract key genes associated with PCOS from high-dimensional genomic data and thereby construct robust

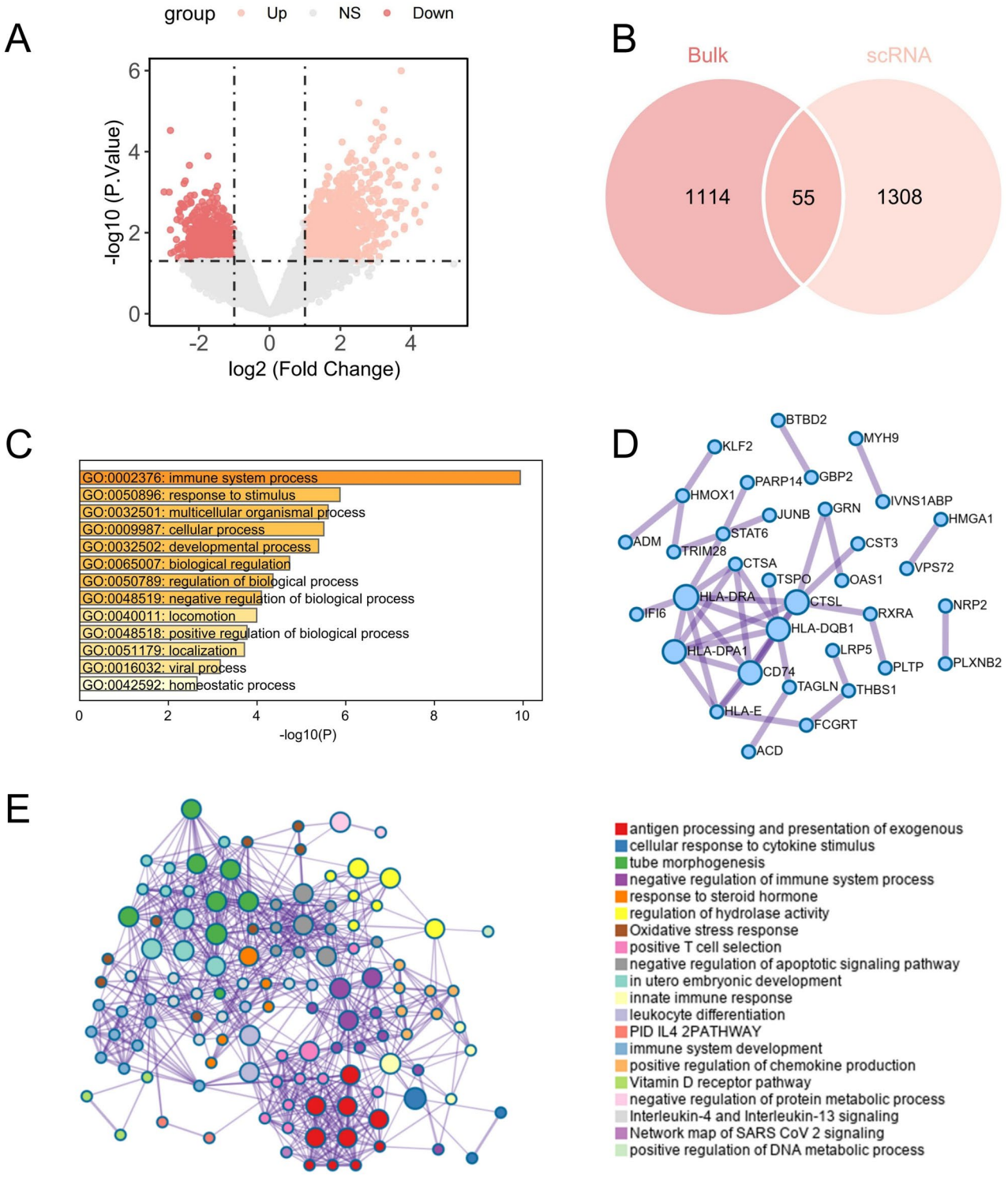

Fig. 2 Identification of Key Differentially Expressed Genes in PCOS and Analysis of Their Biological Functions (**A**) Volcano plot of differentially expressed genes. The x-axis represents the logarithm of fold change (log2 Fold Change), and the y-axis represents the logarithm of significance level (-log10 P-value). Pink dots indicate upregulated genes, red dots indicate downregulated genes, and gray dots indicate genes with no significant difference (NS) (**B**) Venn diagram of omics-specific intersections. Shows the overlap of differentially upregulated genes between bulk transcriptomics (left) and single-cell transcriptomics (right): 1114 bulk-specific upregulators, 1308 scRNA-specific upregulators, and 55 shared upregulators (**C**) GO functional enrichment bar chart for intersecting genes. The vertical axis displays significantly enriched biological processes, the horizontal axis shows enrichment significance (-log10 P-value), and color intensity corresponds to enrichment level (**D**) PPI network of intersecting genes. Nodes represent proteins corresponding to genes, and edges denote interactions between proteins (**E**) Network visualization of functionally enriched terms. Nodes of different colors represent distinct biological processes/pathways, and edges between nodes reflect the degree of association among functional modules

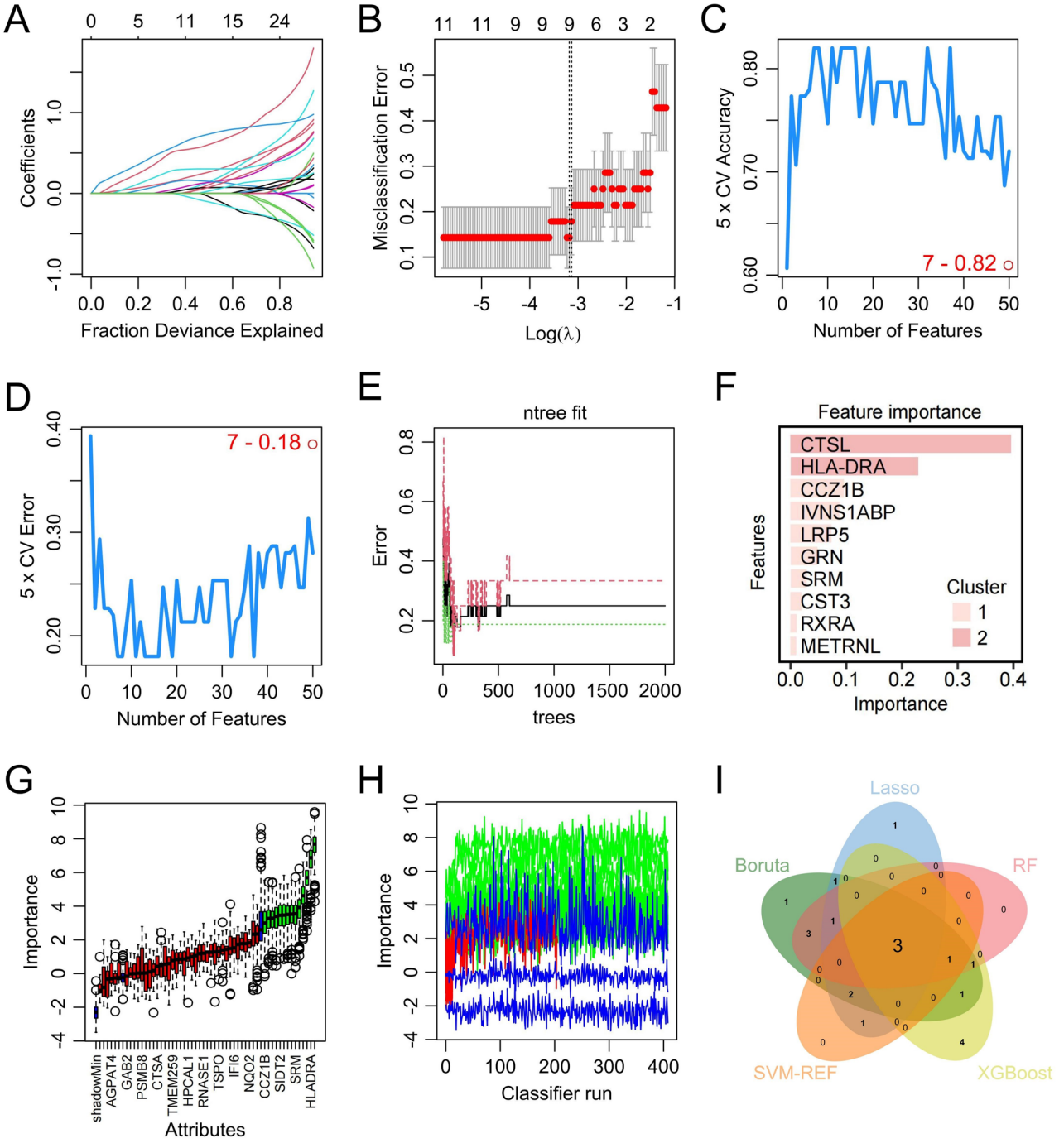

Fig. 3 (See legend on next page.)

gene-signature models. These signatures not only facilitate early diagnosis but also enable the prediction of clinical outcomes and therapeutic responses. Although prior studies have already proposed several PCOS diagnostic models-for instance, Liu et al. developed a signature based on cuproptosis and m6A-targeted genes [30], Zhang et al. utilized peripheral-blood RNA modifications as novel diagnostic features [31], and a ferroptosis-related gene model has likewise been constructed [32]-the model established in this study, which integrates five distinct machine-learning algorithms, achieved an AUC of 0.896. Consequently, it demonstrates superior discriminatory capacity between affected and unaffected samples, thereby underscoring its clinical potential. Hence, the present work not only enhances the rate of early PCOS detection but also provides a foundation for individualized therapy. Future investigations should therefore focus

(See figure on previous page.)

Fig. 3 Construction of the PCOS Diagnostic Model (**A**) LASSO regression coefficient trajectory plot. The horizontal axis represents the fraction of deviance explained, while the vertical axis shows regression coefficients. Different colored curves indicate the trend of coefficients for distinct features as the penalty parameter varies (**B**) LASSO model cross-validation error plot. The horizontal axis shows the logarithm of the penalty parameter (Log (λ)), and the vertical axis shows the classification error. The red horizontal line indicates the threshold corresponding to the optimal error, while the dashed line marks the position of the optimal penalty parameter. The numbers above represent the number of features selected at the corresponding λ value (**C**) Accuracy curve of the SVM-RFE model. The horizontal axis represents the number of features, and the vertical axis shows the 5-fold cross-validation accuracy. The curve illustrates how model accuracy changes as the number of features decreases (**D**) Error curve of the SVM-RFE model. The horizontal axis represents the number of features, and the vertical axis shows the 5-fold cross-validation error (**E**) Fitting error curve of the Random Forest model. The horizontal axis represents the number of decision trees (trees), and the vertical axis shows the model error. The red dashed line indicates the error threshold, and the curve illustrates the trend of model fitting error as the number of trees increases (**F**) Feature importance ranking for the XGBoost model. The horizontal axis represents feature importance scores, and the vertical axis shows candidate features (**G**) Feature importance distribution for the Boruta algorithm. The horizontal axis displays candidate features, and the vertical axis shows feature importance scores; green boxes highlight important features (**H**) Iterative importance trajectory of the Boruta algorithm. The horizontal axis represents the number of algorithm iterations, and the vertical axis represents feature importance scores; different colored curves represent the importance fluctuations of different features during the iterative process (**I**) Intersection Venn diagram of feature selection results from five algorithms: LASSO, Boruta, SVM-RFE, RF, and XGBoost

on further optimization and on evaluating the model's generalizability and stability across diverse populations.

By integrating state-of-the-art machine-learning frameworks, we precisely delineated a gene-signature panel comprising HLA-DRA, SRM, and CTSL as pivotal determinants of PCOS. HLA-DRA encodes a major histocompatibility complex class II molecule that orchestrates antigen presentation; consequently, its up-regulation may fuel aberrant inflammatory circuits underlying ovarian dysfunction [33]. SRM (spermidine synthase) catalyzes a rate-limiting step in polyamine biosynthesis that governs proliferation, differentiation, and apoptosis; although originally characterized in Drosophila immunity via modulation of PGRP-LC/Toll pathways [34], recent evidence links SRM to pro-inflammatory cytokine induction in mammalian systems [35]. We therefore postulate that SRM-mediated perturbation of the ovarian immune microenvironment impairs folliculogenesis and ovulation. CTSL, a cysteine cathepsin, participates in antigen processing and apoptotic cascades [36] and has been implicated in inflammatory disorders such as rheumatoid arthritis [37]. Collectively, these observations illuminate an immune-centric dimension of PCOS pathobiology and nominate actionable therapeutic targets.

Single-cell transcriptomic profiling of follicular-fluid GCs from women with PCOS uncovered ten transcriptionally distinct subclusters, among which cluster GC9 was strikingly expanded and exhibited a hyper-proliferative phenotype. This observation implies that GC9 constitutes a functionally critical subset in PCOS pathogenesis. Research has demonstrated that unchecked GC proliferation up-regulates IL-1β and IL-18, cytokines that recruit mononuclear phagocytes and consequently perturb local immune homeostasis [38]; furthermore, our transcriptomic data reveal a concomitant up-regulation of immune-response genes across PCOS GCs, underscoring the active participation of immune signaling in disease progression. Therefore, expansion of GC9 may propagate chronic inflammation through a granulosa-cell-immune axis, hence providing a mechanistic rationale for therapeutically targeting proliferation-driven immune circuits and thereby informing patient-specific intervention strategies. Collectively, these findings not only refine our molecular understanding of PCOS but also nominate GC9-associated signatures as potential biomarkers for future precision medicine.

Although the "strict-intersection" strategy adopted herein secures high specificity, it inevitably discards weak yet potentially synergistic genes; therefore, subsequent analyses will expand the sample size and integrate gene-gene interactions within models to capture a more comprehensive regulatory landscape. Owing to technical variation across omics platforms and inherent biological heterogeneity, the overlap between bulk-RNA-seq and scRNA-seq remains limited; hence, future work will standardize specimen handling, perform synchronized sequencing with greater depth, and expand multi-center cohorts to overcome dataset-specific constraints and improve both overlap and statistical power. However, while our model demonstrated performance across independent RNA-seq and qPCR/WB datasets, its generalizability must still be evaluated in large, multi-center cohorts profiled on diverse sequencing platforms. Furthermore, the present three-gene signature remains a proof-of-concept; hence, prospective, multi-center trials (> 500 cases) and decision-curve analyses against the current Rotterdam criteria are required to quantify its incremental benefit and health-economic value before clinical translation.

In summary, this work elucidates key aspects of PCOS pathobiology by identifying critical genes and delineating their functional roles, thereby offering novel theoretical foundations for early diagnosis and targeted intervention. These findings provide a fresh perspective on the complexity of PCOS and chart a clear direction for subsequent research. Subsequent investigations should therefore focus on confirming the relevance of these key genes across different populations and on evaluating their clinical utility, thus advancing the development of personalized therapeutic strategies for PCOS.

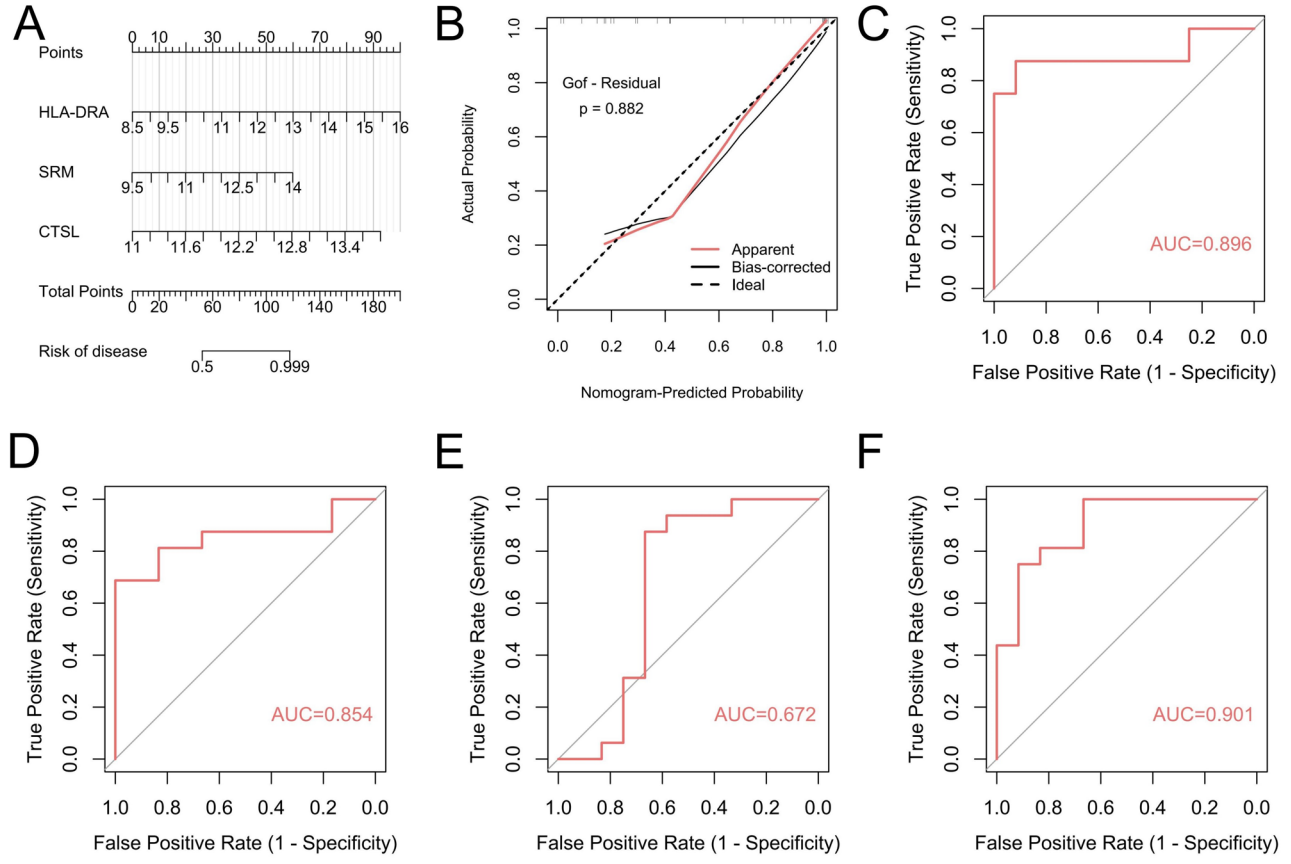

Fig. 4 Diagnostic Model Validation and Performance Evaluation (**A**) Nomogram Model. The figure displays the scoring scales for the three features: HLA-DRA, SRM, and CTSL. The right-hand scale shows the relationship between the "Total Points" (sum of scores for all features) and the PCOS risk probability (risk range 0.5–0.999) (**B**) Calibration Curve. The horizontal axis represents the predicted probability of disease onset from the nomogram, while the vertical axis shows the actual probability of disease onset. "Apparent" indicates the model's calibration performance on the training set, and "Ideal" represents the ideal calibration curve. The G4 test $P = 0.882$ indicates good consistency between predicted and actual probabilities (**C**) ROC Curve for the Nomogram Model. The horizontal axis represents the false positive rate (1 - specificity), the vertical axis represents the true positive rate (sensitivity), and the area under the curve (AUC = 0.896) reflects the model's overall discriminatory capability (**D**-**F**) ROC curves and corresponding AUC values constructed for the HLA-DRA, SRM, and CTSL genes, respectively, demonstrating the discriminatory performance of these genes as individual diagnostic markers

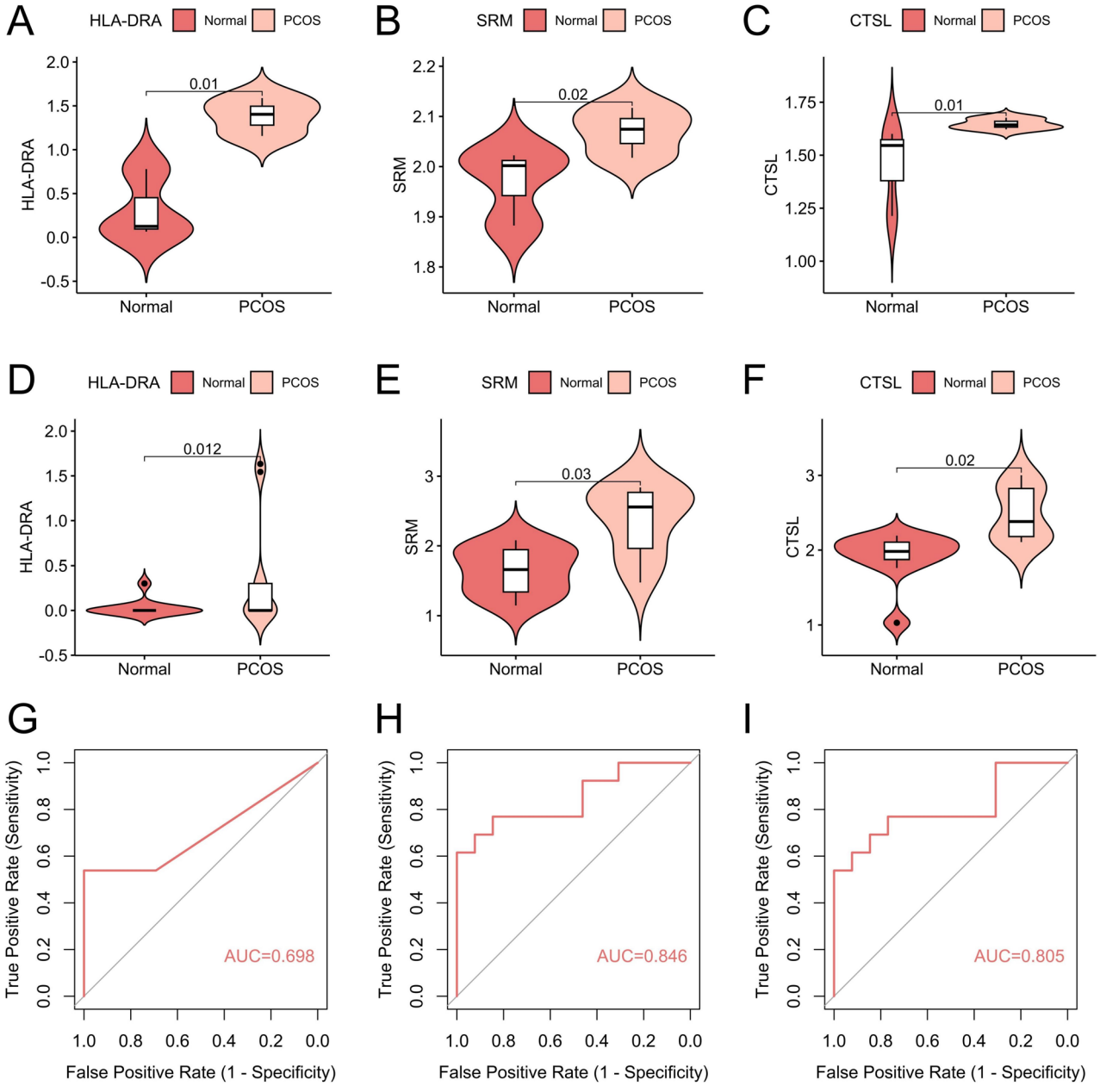

Fig. 5 External Validation of the PCOS Model (**A**-**C**) Differential gene expression in the GSE138518 dataset. Displaying expression distributions of HLA-DRA (**A**), SRM (**B**), and CTSL (**C**) in the normal group (red) versus PCOS group (pink): Violin plots show overall expression distributions, with internal boxplots indicating medians and quartiles. Labeled p-values represent statistical significance of differences between groups (**D**-**F**) Differential gene expression in the GSE155489 dataset. Shows the expression distribution of HLA-DRA (**D**), SRM (**E**), and CTSL (**F**) in the normal group (red) and PCOS group (pink), presented in the same format as (**A**-**C**). The p-values labeled in the figures indicate the significance of differences between the two groups (**G-I**) ROC curves for the external validation set. These curves demonstrate the discriminatory performance of HLA-DRA (AUC = 0.698), SRM (AUC = 0.846), and CTSL (AUC = 0.805) as diagnostic markers. The x-axis represents the false positive rate (1 - specificity), and the y-axis represents the true positive rate (sensitivity)

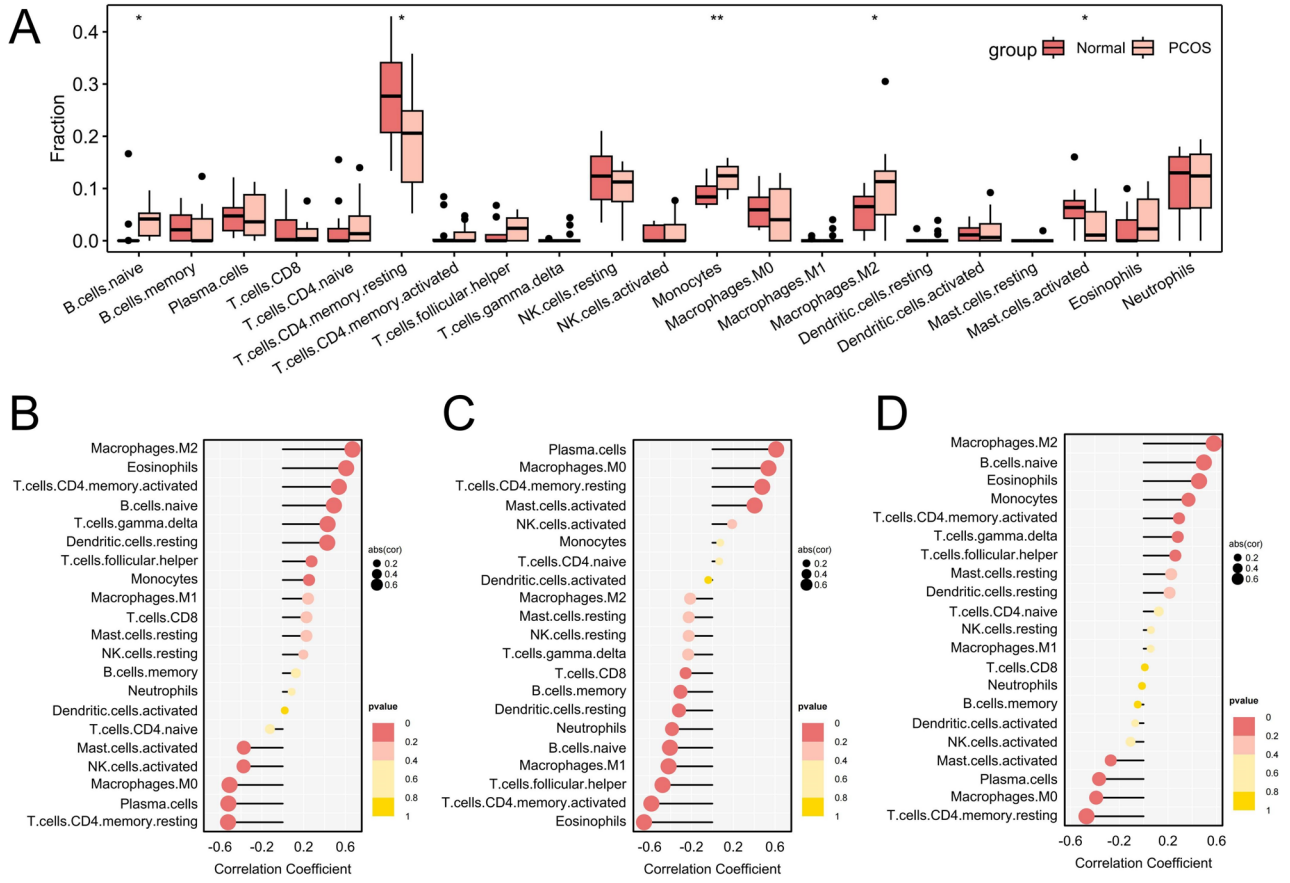

Fig. 6 Immune cell infiltration landscape in PCOS patients (**A**) Intergroup differences in immune cell infiltration proportions. Box plots show infiltration distribution of various immune cell types between the normal group (red) and PCOS group (pink). (*) indicates significance level of differences between groups (*$p < 0.05$, **$p < 0.01$) (**B**-**D**) Correlation between CTSL, HLA-DRA, and SRM with immune cell infiltration. Lollipop plots show immune cell types on the vertical axis and correlation coefficients on the horizontal axis; dot size indicates correlation strength, visually illustrating the association between gene expression and infiltration of various immune cell types

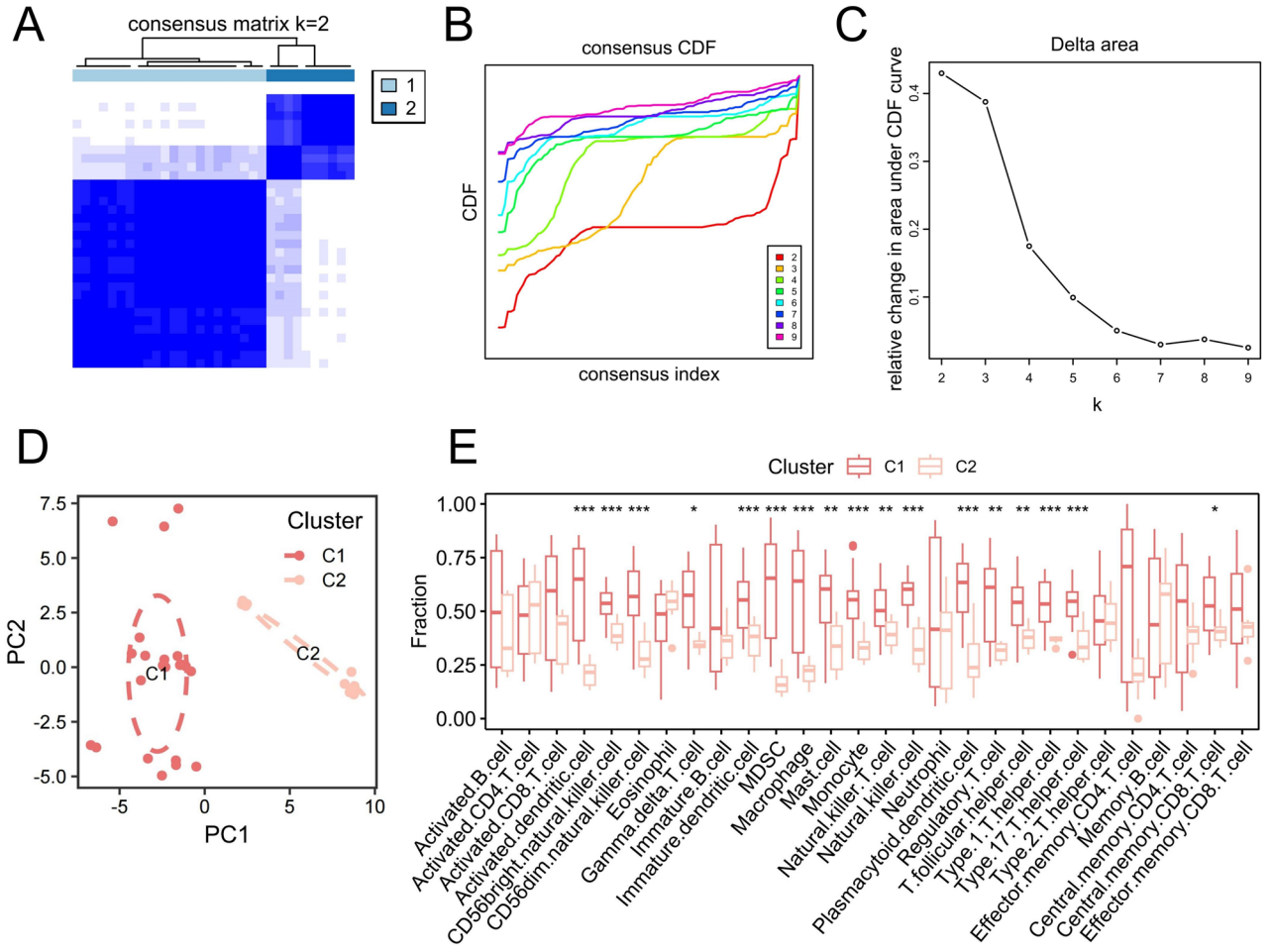

Fig. 7 Identification of PCOS Patient Subgroups and Analysis of Immune Infiltration Differences (**A**) Consensus matrix at k = 2. Colors in the matrix represent the degree of similarity between samples, with darker areas indicating higher consistency among corresponding samples within clusters (**B**) Cluster consensus cumulative distribution function (CDF) curves. Different colored curves correspond to different k values (number of clusters). Steeper curves and higher final CDF values indicate greater clustering stability at that k value (**C**) Delta area curve. The horizontal axis represents the number of clusters k, while the vertical axis shows the relative change in the area under the CDF curve. The point where the downward trend of the curve flattens indicates the optimal number of clusters (k = 2 is a favorable choice) (**D**) PCA Distribution of Subpopulations. The horizontal axis represents the first principal component (PC1), and the vertical axis represents the second principal component (PC2). Red dots indicate the C1 subpopulation, and pink dots indicate the C2 subpopulation. Ellipses outline the concentrated distribution areas of each subpopulation, illustrating the degree of separation between the two subpopulations (**E**) Differences in Immune Cell Infiltration Proportions Between Subpopulations. Box plots display the infiltration ratios of various immune cell types in C1 (red) and C2 (pink) subpopulations. (*) indicates the significance level of differences between groups (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)

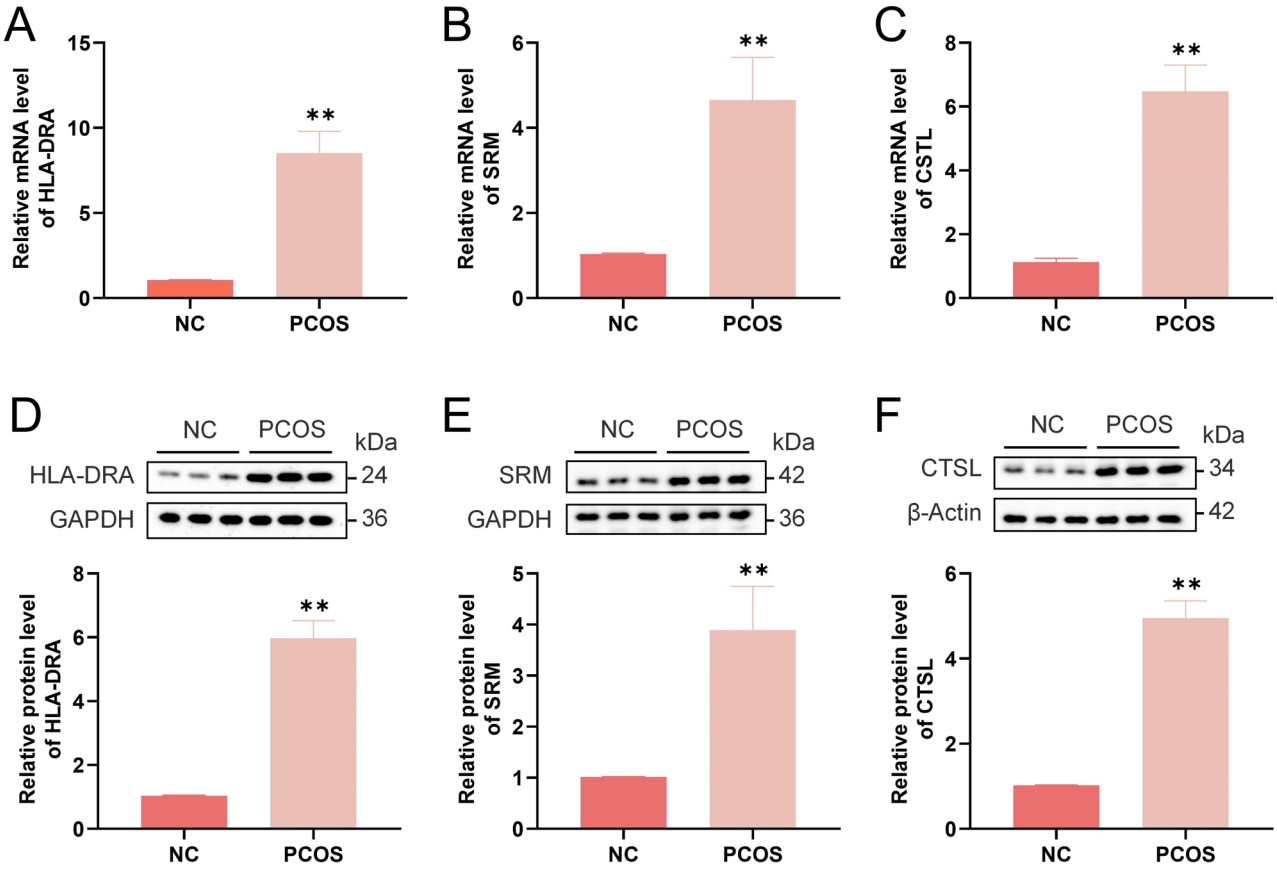

Fig. 8 qRT-PCR and Western Blot validation of CTSL, HLA-DRA, and SRM upregulation (**A**-**C**) mRNA expression levels detected by qRT-PCR. Bar charts show relative mRNA levels of HLA-DRA (**A**), SRM (**B**), and CTSL (**C**) in GCs from control (NC) and PCOS patients, respectively. $^{**}p < 0.01$ indicates statistically significant differences between groups (**D**-**F**) Protein expression levels detected by Western Blot. The upper panel displays protein bands for HLA-DRA (**D**, internal control GAPDH), SRM (**E**, internal control GAPDH), and CTSL (**F**, internal control β-Actin) in NC and PCOS groups. The lower bar charts show relative expression levels of target proteins compared to internal controls. $^{**}p < 0.01$ indicates statistically significant differences between groups

Abbreviations

PCOS	Polycystic ovary syndrome
SVM	Support-vector machine
RF	Random forest
LASSO	Least absolute shrinkage and selection operator
AUC	Area under the curve
ROC	Receiver operating characteristic
ssGSEA	Single-sample gene set enrichment analysis
qPCR	Quantitative real-time PCR
WB	Western blot
UMAP	Uniform Manifold Approximation and Projection

Acknowledgements
None.

Authors' contributions
KX drafted and revised the manuscript. SZ, LG, and TL per-dformed data analysis. YL designed the study. YZ revised the manuscript. All authors read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation (No. 82503256) and the Scientific Research Project of Yancheng Health Commission (No. YK2024065).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
The study protocol was approved by the Ethics Committee of Binhai County People's Hospital (No. 2025-BYKYLL-027).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 7 August 2025 / Accepted: 24 December 2025
Published online: 07 January 2026

References

1. Jalil AT, Zair MA, Hanthal ZR, Naser SJ, Aslandook T, Abosaooda M, et al. Role of the AMP-Activated protein kinase in the pathogenesis of polycystic ovary syndrome. Indian J Clin Biochemistry: IJCB. 2024;39(4):450–8.
2. Rosenfield RL, Ehrmann DA. The pathogenesis of polycystic ovary syndrome (PCOS): the hypothesis of PCOS as functional ovarian hyperandrogenism revisited. Endocr Rev. 2016;37(5):467–520.
3. Eriksson G, Li C, Sparovec TG, Dekanski A, Torstensson S, Risal S, et al. Single-cell profiling of the human endometrium in polycystic ovary syndrome. Nat Med. 2025. https://doi.org/10.1038/s41591-025-03592-z.

4. Harris RA, McAllister JM, Strauss JF 3rd. Single-Cell RNA-Seq Identifies Pathways and Genes Contributing to the Hyperandrogenemia Associated with Polycystic Ovary Syndrome. Int J Mol Sci. 2023;24(13):10611. https://doi.org/10.3390/ijms241310611.
5. Qi L, Li Y, Zhang L, Li S, Zhang X, Li W, et al. Immune and oxidative stress disorder in ovulation-dysfunction women revealed by single-cell transcriptome. Front Immunol. 2023;14:1297484.
6. Luo M, Yang X, Zhou M, Zhang J, Yu B, Lian H, et al. Integrated single-cell and Spatial transcriptomics reveal microenvironment disruptions by androgen in mouse ovary. iScience. 2024;27(10):111028.
7. Li S, Li Y, Sun Y, Feng G, Yang Z, Yan X, et al. Deconvolution at the single-cell level reveals ovarian cell-type-specific transcriptomic changes in PCOS. Volume 22. Reproductive biology and endocrinology: RB&E; 2024. p. 24. 1.
8. Ye Z, Cheng M, Lian W, Leng Y, Qin X, Wang Y, et al. GPX4 deficiency-induced ferroptosis drives endometrial epithelial fibrosis in polycystic ovary syndrome. Redox Biol. 2025;83:103615.
9. Wang M, Zhang S, He J, Zhang T, Zhu H, Sun R, et al. Biochemical classification diagnosis of polycystic ovary syndrome based on serum steroid hormones. J Steroid Biochem Mol Biol. 2025;245:106626.
10. Wang X, Lai X, Pan C, Yin T, Zhu W, Zhu X, et al. The association between metal exposure in follicular fluid and PCOS risk in reproductive-aged women: mediating roles of LH/FSH ratio and androgens. Ecotoxicol Environ Saf. 2025;300:118473.
11. Wang M, An K, Huang J, Mprah R, Ding H. A novel model based on necroptosis to assess progression for polycystic ovary syndrome and identification of potential therapeutic drugs. Front Endocrinol. 2023;14:1193992.
12. Yan X, Yang Z, Zhao H, Feng G, Li S, Li Y, et al. Unveiling lipoprotein subfractions signature in high-FNPO PCOS: implications for PCOM diagnosis and risk assessment using advanced machine learning models. BMC Med. 2025;23(1):289.
13. Kumariya S, Ubba V, Jha RK, Gayen JR. Autophagy in ovary and polycystic ovary syndrome: role, dispute and future perspective. Autophagy. 2021;17(10):2706–33.
14. Wang Z, Wu Z, Wang H, Feng R, Wang G, Li M, et al. An immune cell atlas reveals the dynamics of human macrophage specification during prenatal development. Cell. 2023;186(20):4454–e7119.
15. Luyckx L, Myllykangas M, Saarela U, Virtanen N, Hurskainen E, Savolainen A, et al. Prenatally androgenized PCOS mice have ovary-independent uterine dysfunction and placental inflammation aggravated by high-fat diet. Sci Adv. 2025;11(19):eadu3699.
16. Shabbir S, Khurram E, Moorthi VS, Eissa YTH, Kamal MA, Butler AE. The interplay between androgens and the immune response in polycystic ovary syndrome. J Transl Med. 2023;21(1):259.
17. Rantakallio JSS, Nevalainen JE, West SI, Ollila MM, Puukka K, Bloigu AH, et al. Association of Self-Reported polycystic ovary Syndrome, Obesity, and weight gain from adolescence to adulthood with hypertensive disorders of pregnancy: A Community-Based approach. Hypertens (Dallas Tex: 1979). 2021;77(3):1010–9.
18. Koca SB, Tengiç E, Büyükyılmaz G. Presence of hyperandrogenemia in cases evaluated due to menstrual irregularity, the effect of clinical and/or biochemical hyperandrogenemia on polycystic ovary syndrome. J Pediatr Endocrinol Metabolism: JPEM. 2025;38(6):631–6.
19. Mehta P, Saha D, Das A, Das BK. Gut microbiota in diabetic-linked polycystic ovarian syndrome: mechanisms and therapeutic insights. Tissue Cell. 2025;95:102870.
20. Marantides D. Management of polycystic ovary syndrome. The Nurse Practitioner. 1997;22(12):34–8. 40 – 1.
21. Liu X, Song J, Yuan M, Zuo F, Li H, Tang L, et al. Single-cell transcriptional dissection illuminates an evolution of immunosuppressive microenvironment during pancreatic ductal adenocarcinoma metastasis. Signal Transduct Target Ther. 2025;10(1):182.
22. Wei L, Bo L, Jiang W, Qi R, Luo C, Qian F, et al. Single cell RNA sequencing reveals the role of local renin-angiotensin system in regulating ovarian physiological cycle and promoting PCOS. Cell Death Discov. 2025;11(1):255.
23. Liu D, Liu D, Zhou K. Polycystic ovary syndrome and epithelial-mesenchymal transition: Mendelian randomization and single-cell analysis insights. J Ovarian Res. 2025;18(1):33.
24. Chen L, Hui L, Wang Y, Yao X, Li J. Elevated IGFBP7 expression in follicular granulosa cells promotes PCOS pathogenesis. Biochimica et Biophysica Acta (BBA). 2025;1871(4):167743.
25. Wang J, Yin T, Liu S. Dysregulation of immune response in PCOS organ system. Front Immunol. 2023;14:1169232.
26. Zhang Y, Hu Y, Yu J, Xie X, Jiang F, Wu C. Landscape of PCOS co-expression gene and its role in predicting prognosis and assisting immunotherapy in endometrial cancer. J Ovarian Res. 2023;16(1):129.
27. Emanuel RHK, Roberts J, Docherty PD, Lunt H, Campbell RE, Möller K. A review of the hormones involved in the endocrine dysfunctions of polycystic ovary syndrome and their interactions. Front Endocrinol. 2022;13:1017468.
28. 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.
29. Purwar A, Nagpure S. Insulin resistance in polycystic ovarian syndrome. Cureus. 2022;14(10):e30351.
30. Liu Y, Gengxiao Y, Wu Y, Li Y, Hu X, Bie J. Comprehensive analysis of cuproptosis and m6A-target gene signatures for diagnostic significance and immune microenvironment characterization in polycystic ovary syndrome. Sci Rep. 2025;15(1):19996.
31. Zhang L, Liu X, Zhang Y, Qin L, Pan S, Yan X, et al. Peripheral blood RNA modifications as a novel diagnostic signature for polycystic ovary syndrome. Sci China Life Sci. 2025. https://doi.org/10.1007/s11427-024-2913-7.
32. Zhang J, Ding N, Xin W, Yang X, Wang F. Quantitative proteomics reveals that a prognostic signature of the endometrium of the polycystic ovary syndrome women based on ferroptosis proteins. Front Endocrinol. 2022;13:871945.
33. Zuo W, Liu X, Chen J, Zuo W, Yin Y, Nie X, et al. Single-cell sequencing provides insights into the landscape of ovary in PCOS and alterations induced by CUMS. Am J Physiol Endocrinol Metab. 2023;325(4):E346-e62.
34. Yoo TJ, Sup Shim M, Bang J, Kim JH, Jae Lee B. SPS1 deficiency-triggered PGRP-LC and Toll expression controls innate immunity in Drosophila S2 cells. Biol Open. 2022. https://doi.org/10.1242/bio.059295.
35. Lin TJ, Yang SS, Hua KF, Tsai YL, Lin SH, Ka SM. SPAK plays a pathogenic role in IgA nephropathy through the activation of NF-κB/MAPKs signaling pathway. Free Radic Biol Med. 2016;99:214–24.
36. Shen X, Zhao YF, Xu SQ, Wang L, Cao HM, Cao Y, et al. Cathepsin L induced PC-12 cell apoptosis via activation of B-Myb and regulation of cell cycle proteins. Acta Pharmacol Sin. 2019;40(11):1394–403.
37. Clanchy FIL, Borghese F, Bystrom J, Balog A, Penn H, Taylor PC, et al. Disease status in human and experimental arthritis, and response to TNF blockade, is associated with MHC class II invariant chain (CD74) isoform expression. J Autoimmun. 2022;128:102810.
38. Huang R, Gao M. Circ_0070987 promotes pyroptosis of ovarian granulosa cells in polycystic ovarian syndrome through the miR-139-5/CDH1 axis. Reproductive sciences (Thousand Oaks, Calif). 2025;32(10):3361–70.

Publisher's Note