Article

Discovery Medicine 2025; 37(201): 2371–2387
https://doi.org/10.24976/Discov.Med.202537201.201

Individual-Patient Networks Inform Patient-Level Differences in Phenotype and Delineate Key Candidate Genes and Signaling Pathways in PCOS Granulosa Cells

Fan Zhang[1], Yi-hua Chen[1], Li-li Sun[2], Wei Wei[1,*]

[1]Department of Reproductive Medicine Center, Affiliated Zhongshan Hospital of Dalian University, 116001 Dalian, Liaoning, China
[2]Gynecology Clinic, Affiliated Zhongshan Hospital of Dalian University, 116001 Dalian, Liaoning, China
*Correspondence: visa0451@163.com (Wei Wei)
Submitted: 19 June 2025 Revised: 5 August 2025 Accepted: 25 August 2025 Published: 20 October 2025

Background: Polycystic ovary syndrome (PCOS) is a highly prevalent and complex endocrine-metabolic syndrome, characterized by impaired folliculogenesis and follicular arrest. The marked clinical heterogeneity observed among PCOS patients is hypothesized to reflect variations in its etiopathogenesis, although the underlying mechanisms remain incompletely elucidated. This study was conducted employing individual-patient protein-protein interaction (PPI) network analysis to investigate key candidate genes and their associated signaling pathways in the granulosa cells (GCs) of PCOS patients at a single-patient resolution. The aim of this study was to provide novel insights and deepen the understanding of the pathogenic mechanisms and heterogeneity of PCOS.
Methods: Data were extracted from three mRNA expression datasets derived from high-throughput sequencing (GSE168404, GSE155489, and GSE138518) available in the Gene Expression Omnibus (GEO) database. These datasets included 24 samples, comprising 12 from individuals with PCOS and 12 controls. A gene correlation matrix was generated using control samples to assess global transcriptional alterations induced by data from individual PCOS patients. Individual-patient PPI networks were subsequently constructed to identify key candidate genes and molecular subtypes unique to each patient. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed via KOBAS-i to elucidate the underlying biological functions.
Results: A total of 7752 nodes (1978 unique) and 19,219 edges (13,626 unique) were mapped to construct the individual-patient PPI networks. These networks exhibited considerable variability in terms of nodes (median = 2128, range = 1141–3838) and edges (median = 2369, range = 903–6237). Despite some shared molecular features, the distinct network architectures observed in each patient underscored patient-specific variations in gene expression, thereby reflecting the intrinsic heterogeneity of the PCOS population. To identify key candidate genes, five overlapping nodes (fibronectin 1 (*FN1*), DNA ligase 3 (*LIG3*), vesicle associated membrane protein 2 (*VAMP2*), kinesin family member 23 (*KIF23*), pescadillo ribosomal biogenesis factor 1 (*PES1*)) were extracted from these networks, and *KIF23*, along with protein regulator of cytokinesis 1 (*PRC1*), were identified as common hub genes when analyses were restricted to more than 50% of patients. Furthermore, four signaling pathways, cyclic Guanosine Monophosphate-Protein Kinase G (cGMP-PKG), Wingless/Int-1 (Wnt), Relaxin, and Apelin, were differentially enriched across all individual-patient PPI networks. These pathways were distinct from those enriched with differentially expressed genes (DEGs) between the PCOS and control groups.
Conclusion: This study revealed patient-specific variations in gene expression by examining distinct network characteristics across individuals, highlighting the molecular heterogeneity inherent in PCOS. In addition to identifying candidate genes and signaling pathways enriched with DEGs, *KIF23* emerged as a potential hub gene due to its ubiquitous presence. Additionally, the cGMP-PKG, Wnt, relaxin, and apelin signaling pathways were identified as potential core signaling pathways. However, these findings require further experimental validation through comprehensive *in vivo* and *in vitro* studies to establish their biological relevance.

Keywords: polycystic ovary syndrome; Gene Expression Omnibus; individual-patient networks; key candidate genes; signaling pathways

Introduction

Polycystic ovary syndrome (PCOS) is a prevalent endocrine-metabolic disorder affecting 5–20% of reproductive-aged women worldwide, which is characterized by hyperandrogenaemia (HA), insulin resistance (IR), gonadotropin imbalance (elevated Luteinizing Hormone (LH)/Follicle-Stimulating Hormone (FSH) ratios), and increased cardiometabolic risks [1–3]. Central to its pathogenesis is aberrant folliculogenesis and follicular arrest, wherein impaired granulosa cell (GC) function disrupts follicle maturation and ovulation. Recent studies have demonstrated that GC dysfunction in PCOS involves profound metabolic reprogramming, including altered glycolysis-oxidative phosphorylation balance [4], defective branched-chain amino acid (BCAA) catabolism due to protein phosphatase, Mg^{2+}/Mn^{2+} dependent 1K (*PPM1K*) downregulation [5], and endoplasmic reticulum stress-induced ferroptosis [2]. These alterations directly compromise follicular development and oocyte competence.

GCs orchestrate folliculogenesis through multiple mechanisms: (1) regulating steroidogenesis and gonadotropin sensitivity via autocrine/paracrine signaling (e.g., the Interleukin-22/Signal Transducer and Activator of Transcription 3 signaling pathway (IL-22/STAT3 pathways)) [1,6]; (2) supplying oocytes with nutrients (glutamine, BCAA) and metabolic intermediates [3,5]; and (3) sustaining gap junction-mediated transfer of regulatory molecules (e.g., Connexin 43 (Cx43)/Gap Junction Protein, Alpha 1 (*GJA1*)) essential for oocyte development [3,5,7]. In PCOS, hyperandrogenaemia disrupts these functions by inducing GC apoptosis via C-X-C Motif Chemokine Ligand 10 (CXCL10)-Janus Kinase (JAK) signaling [7], inhibiting exocytosis of essential factors (Syntaxin 6 (Stx6), Synaptosomal Nerve-Associated Protein 25 (SNAP25)) [8], and impairing metabolic homeostasis (e.g., mechanistic Target of Rapamycin-mediated signaling pathways (mTOR-mediated pathways)) [3,9]. Consequently, follicular arrest and anovulation arise from GC dysfunction, highlighting therapeutic targets such as rapamycin [9] and liraglutide [7] that restore GC-oocyte crosstalk.

Although significant progress has been made in understanding granulosa cell pathophysiology in PCOS, the core mechanisms underlying this disorder remain incompletely understood [10]. Transcriptomic analysis in PCOS granulosa cells has emerged as a powerful approach to identify molecular signatures of ovulatory dysfunction and to discover potential biomarkers and therapeutic targets [11, 12]. Recent transcriptomic studies revealed dysregulation of key signaling pathways, including mitogen-activated protein kinase (MAPK), Wingless/Int-1 (Wnt), and transforming growth factor-beta (TGF-β) pathways in granulosa cells [13]. Furthermore, Toll-like receptor and NOD-like receptor pathways were enriched and mechanistically linked to macrophage pyroptosis in PCOS ovaries [10]. Biomarker investigations identified mevalonate diphosphate decarboxylase (*MVD*) and patatin like domain 3, 1-acylglycerol-3-phosphate O-acyltransferase (*PNPLA3*) as significantly dysregulated genes in granulosa cells associated with metabolic dysfunction [11,14], whereas multi-omics approaches implicated ferroptosis mediated by Solute Carrier Family 7 Member 11 (SLC7A11)/Glutathione Peroxidase 4 (GPX4) in granulosa cells [15]. Regarding apoptosis regulation, DLG associated protein 5 (*DLGAP5*) has emerged as an anti-apoptotic regulator overexpressed in PCOS GCs, promoting cell viability and proliferation [16], in contrast with earlier findings implicating forkhead box O3 (*FoxO3*). Transcriptomic profiling consistently identified hundreds of differentially expressed genes (e.g., 545 deregulated genes identified [11]; 705 differentially expressed genes (DEGs) with pathway enrichment [13]) and diverse noncoding RNAs, including microRNAs (miRNA) (e.g., *miR-205-5p* [14]), circular RNAs (circRNA), and long non-coding RNAs (lncRNA) that form complex regulatory networks. Alternative splicing analysis further uncovered >1200 differential events affecting endoplasmic reticulum stress responses and Yes-associated protein 1 (YAP1) signaling in granulosa cells [12]. Collectively, these findings have advanced the understanding of PCOS pathology but underscore the need for further integration of transcriptomic dysregulation, metabolic reprogramming [4,17], and immune-microenvironment interactions [10] within granulosa cells.

Conventional approaches that focus solely on rare genetic variants or transcriptomic fold changes often overlook the intricate molecular interactions underlying the heterogeneity of PCOS. Given the considerable variability in clinical manifestations among individuals, personalized pathway analysis is crucial for the development of effective therapeutic interventions. A landmark study by Maron *et al.* [18] was the first to apply patient-specific protein-protein interaction (PPI) networks to uncover unique pathological mechanisms in hypertrophic cardiomyopathy. Their findings demonstrated that individualized network analyses provide essential insights into disease mechanisms and clinically relevant phenotypic diversity at the single-patient level [18]. Building on this innovative methodology, the present study aimed to adapt the framework established by Maron *et al.* [18] to investigate the molecular pathology of granulosa cells in PCOS patients through personalized PPI network analysis, thereby facilitating patient-specific characterization of disease mechanisms.

Materials and Methods

In this study, we developed PPI networks to investigate distinct pathological mechanisms present in the GCs of patients with PCOS, using RNA sequencing (RNA-Seq) data obtained from the Gene Expression Omnibus (GEO) database (a public functional genomics data repository, maintained by the National Center for Biotechnology Infor-

Table 1. GEO datasets used in the study.

GSE IDs	Contributors	Submission Date	Overall Design	Sample Size		Platform ID	Platform Description
				C	P		
GSE168404	Zhao *et al.* [19]	8-Mar-2021	Granulosa cells from five women with PCOS and five age- and BMI-matched controls	5	5	GPL16791	Illumina High-Throughput Sequencing 2500 System (HiSeq 2500) (*Homo sapiens*)
GSE155489	Li *et al.* [20]	8-Sep-2020	Oocytes and GCs from three PCOS patients compared with age-matched non-PCOS women	4	4	GPL20795	Illumina High-Sequencing X Ten System (HiSeq X Ten) (*Homo sapiens*)
GSE138518	Mao *et al.* [21]	8-Oct-2019	Ovarian granulosa cells from PCOS patients and healthy controls	3	3	GPL11154	Illumina High-Throughput Sequencing 2000 System (HiSeq 2000) (*Homo sapiens*)

Notes: C, control; P, PCOS; GEO, Gene Expression Omnibus.

mation (NCBI) in Bethesda, MD, USA. https://www.ncbi.nlm.nih.gov/geo/). Applying a meta-analytical approach, we systematically standardized and integrated multiple case-control datasets. A gene correlation matrix derived from control samples served as a reference to detect systemic transcriptional alterations, incorporating data from individual PCOS patients. Individual-patient PPI networks were subsequently constructed to classify molecular subtypes and identify key regulatory elements within relevant biological processes and signaling pathways.

The overall experimental design and analytical pipeline employed in this study are schematically illustrated in Fig. 1.

Data Collection

On November 5, 2021, the GEO database was systematically searched to identify high-throughput sequencing datasets reporting GC gene expression in patients with PCOS. The search strategy used the following query: ("polycystic ovary syndrome"[MeSH Terms] OR "polycystic ovarian syndrome"[All Fields]) AND "granulosa cells"[All Fields] AND "*Homo sapiens*"[porgn] AND ("gse"[Filter] AND "Expression profiling by high-throughput sequencing"[Filter]).

This search initially retrieved six datasets. Three were excluded due to the presence of pharmacological interventions (GSE146856), composite data (GSE138575), or lack of control groups (GSE124968). The remaining three datasets (GSE168404, GSE155489, and GSE138518) satisfied all inclusion criteria, as detailed in Table 1 (Ref. [19–21]), **Supplementary Table 1**, and Fig. 1. These datasets comprised study populations divided into PCOS and control groups. Although GSE168404 contained DNA methylation and lncRNA-miRNA-mRNA data, only RNA sequencing (RNA-Seq) results were analyzed to ensure consistency with PPI analysis. Similarly, while GSE155489 included both oocytes and GC data, only the GC-specific expression profiles were retained to focus exclusively on GC-specific molecular signatures.

RNA-Seq Data Preprocessing and Analysis

The RNA-Seq data (detailed in **Supplementary Table 2**) were obtained from the European Nucleotide Archive (ENA a comprehensive public nucleotide sequence database, maintained by the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) in Hinxton, Cambridge, UK. https://www.ebi.ac.uk/ena/browser/home) for the selected GEO datasets. Initial quality control was performed using FastQC software (version 0.11.9), which developed by the Babraham Institute (Cambridge, UK) and was used for sequencing quality control. Sequencing reads were trimmed using Trimmomatic (version 0.39, an open-source bioinformatics tool developed by Björn Usadel *et al.* and based in Aachen, Germany) with stringent parameters (-PE -phred33 -SLIDINGWINDOW:4:15 -LEADING:10 -TRAILING:3 -MINLEN:36 -ILLUMINACLIP: TruSeq3 -PE -2.fa:2:30:10:1:TRUE) to ensure high-quality sequencing data. Transcript abundances were quantified using kallisto (version 0.46.2, an open-source RNA-seq quantification tool developed by Lior Pachter and based at the California Institute of Technology in Pasadena, CA, USA) with default settings, employing an index built from the human transcriptome fasta file (GRCh38.p13, RefSeq, https://www.ensembl.org).

Transcript-level abundance estimates were imported into R (version 4.1.2, an open-source programming language for statistical computing developed by the R Core Team and maintained by the R Foundation in Vienna, Austria) and aggregated into gene-level counts and expression values using the *tximport* package (version 1.22.0, an R/Bioconductor package for transcript-level data import and summarization developed by Michael Love *et al.* and maintained as part of the Bioconductor project based in the USA) with the parameters: type = "kallisto", countsFromAbundance = "lengthScaledTPM". Batch effects across datasets were corrected using *ComBat_seq* from the *sva* package (version 3.42.0, an R/Bioconductor package for removing batch effects in genomic data devel-

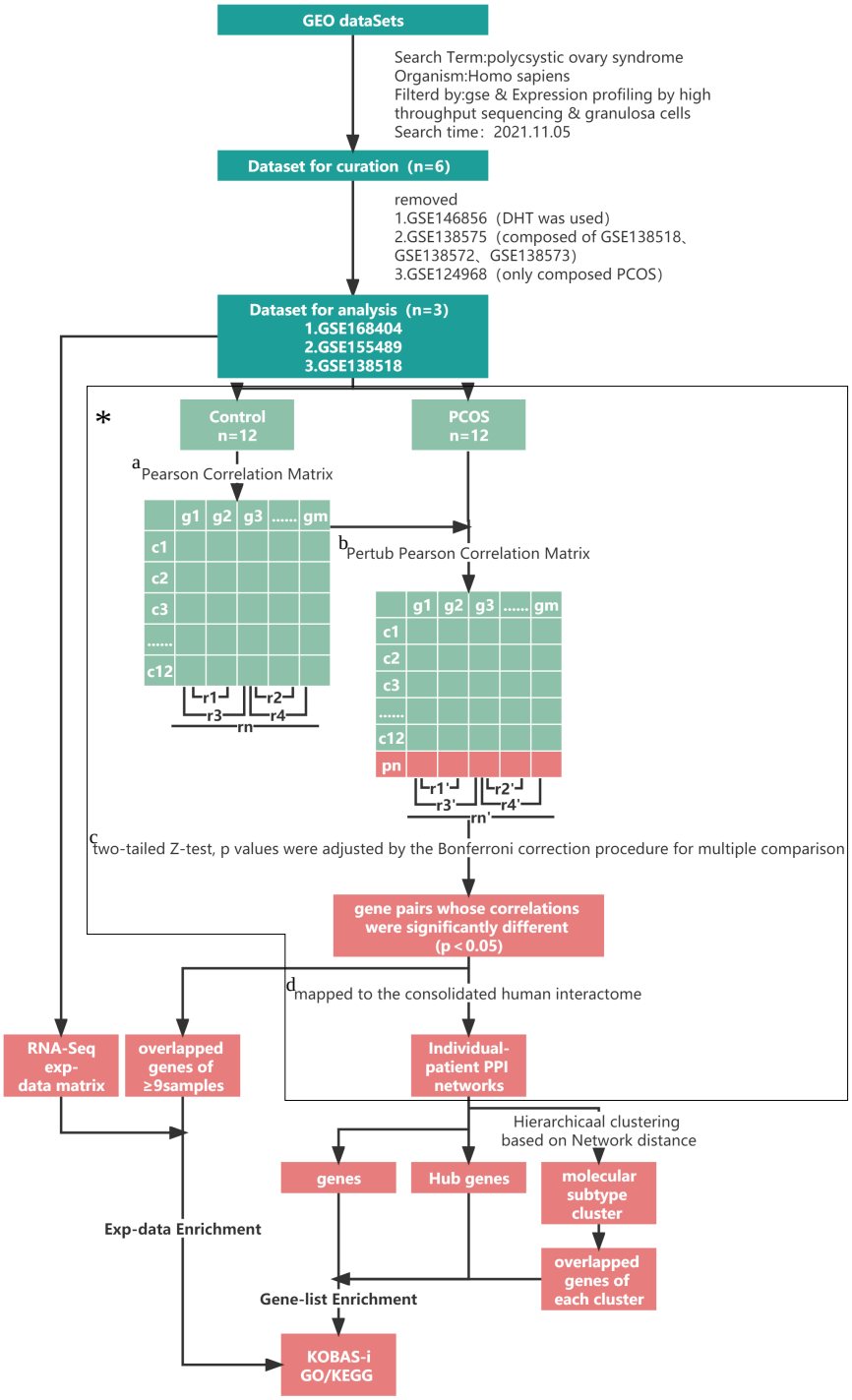

Fig. 1. Workflow for constructing patient-specific PPI networks in PCOS. Schematic overview of the methodological framework applied in this study. RNA sequencing (RNA-Seq) data were computationally processed to classify samples into control (non-PCOS, n = 12, labeled C1-C12) and PCOS groups based on clinical diagnoses. (a) For the control group, normalized read counts were used to calculate Pearson correlation coefficients (r) between all possible gene pairs following quality filtering, where N represents the total number of pairwise correlations and m the total gene count in the dataset. (b) To evaluate network disruptions, an iterative perturbation analysis was performed by sequentially integrating each PCOS patient's transcriptomic profile into the control expression matrix, yielding modified correlation coefficients (r'). This process was repeated for all 12 PCOS patients. (c) Differential network analysis was conducted to identify statistically significant alterations between baseline (r) and perturbed (r') states using two-tailed Z tests with Bonferroni-adjusted p-values to correct for multiple testing. (d) Significantly altered gene pairs (g1 to g2) were subsequently mapped onto a comprehensive human interactome database, comprising 19,304 proteins and 1,157,032 documented interactions. PPI, protein-protein interaction; PCOS, Polycystic ovary syndrome.

oped by Jeffrey T. Leek *et al*. and maintained by the Bioconductor project in the USA). The processed data were subjected to differential expression analysis using DESeq2 (version 1.34.0, an R/Bioconductor package for differential gene expression analysis developed by Michael Love *et al*. and maintained by the Bioconductor project in the USA). Significantly differentially expressed genes (DEGs) were defined as those with |log2 fold change| $\geq$1.5 and *p*-values $<$ 0.05. Finally, dimensionality reduction was performed with the uniform manifold approximation and projection (UMAP) algorithm using the *umap* package (version 0.2.7.0, a dimensionality reduction package available for R developed by Leland McInnes *et al*. and maintained as an open-source project with global contributors), employing the "naive" method with *n_neighbors* set to 6. This approach facilitated visualization of batch correction outcomes and provided a low-dimensional representation of transcriptomic profiles.

Identifying the Personalized Gene Pairs in Each Patient

To ensure robust analysis, only genes with counts $\geq$10 across all control samples were retained for further examination. The Pearson correlation coefficient (PCC, *r*) was calculated for each gene pair using the filtered normalized read counts from all control cohorts (n = 12), thereby establishing the control gene-pair correlation matrix (Fig. 1). The filtered normalized read counts from each PCOS patient were subsequently integrated into the control correlation matrix, and a new PCC (*r'*) was calculated for all gene pairs (Fig. 1). Afterwards, the PCOS patient-filtered read counts were excluded from the matrix, and this procedure was iteratively repeated for each PCOS patient (n = 12). Gene pairs exhibiting significant perturbations were identified by comparing the *r* and *r'* values (ΔPCCs) using two-tailed Z tests, with *p*-values adjusted via Bonferroni correction (significance threshold: adjusted $p < 0.05$) [18].

To visualize patient-specific dysregulated gene pairs, UMAP (implemented with the *umap* package, version 0.2.7.0, a dimensionality reduction package available for R developed by Leland McInnes *et al*. and maintained as an open-source project with global contributors) was employed with the following parameters: method = "naive", n_neighbors = 2.

Building Individual–Patient Networks and Analysis

For network construction, the significant gene pairs were mapped to a comprehensive human interactome encompassing 19,304 proteins and 1,157,032 interactions. This interactome was compiled from three curated databases: HINT (a bioinformatics resource for histone modification interactions, maintained by the Baylor College of Medicine in Houston, Texas, USA. http://hint.yulab.org), PICKLE (a protein interaction and compound knowledge database developed and maintained by the Institute of Biosciences & Applications at the National Centre for Scientific Research "Demokritos" in Athens, Greece. http://www.pickle.gr), and STRING (version 11.5, a protein-protein interaction network database developed and maintained by the Swiss Institute of Bioinformatics (SIB) and consortium partners in Lausanne, Switzerland. https://cn.string-db.org), all maintaining confidence scores $>$0.4. Data were retrieved on December 12, 2021 (Fig. 1). The mapped gene pairs facilitated the construction of an individual-patient PPI network for each PCOS patient. The resulting networks were analyzed and visualized using Cytoscape (version 3.9.0, an open-source network visualization and analysis platform developed by an international consortium led by the University of California San Diego and maintained primarily from La Jolla, CA, USA). To identify key regulators, the cytoHubba plug-in was applied, ranking hub genes based on degree centrality, with a specific focus on those exhibiting degrees $>$20.

Identification of Molecular Subtypes

Agglomerative hierarchical clustering was employed to stratify PCOS patients based on their individual-patient PPI networks, using network distances and the *ConsensusClusterPlus* R package (version 1.58.0, an R/Bioconductor package for consensus clustering analysis developed by Matt Boehmke *et al*. and maintained by the Bioconductor project in the USA). The parameters were configured as follows: maximum number of clusters (*maxK*) = 6, repetitions (*reps*) = 10,000, proportion of items sampled (*pItem*) = 0.8, proportion of features sampled (*pFeature*) = 1, clustering algorithm (*clusterAlg*) = hierarchical clustering (hc), distance metric = network distances, and random seed = 1234.

The network distance between samples was defined analogously to the Jaccard similarity coefficient:

$$d_{ij} = 1 - \frac{|E_I \cap E_J|}{|E_I \cup E_J|}$$

Where d_{ij} represents the network distance between samples *i* and *j*; E_i and E_j denote the sets of edges for the networks of samples *i* and *j*, respectively; $|E_i \cap E_j|$ is the number of overlapping edges between samples *i* and *j*; and $|E_i \cup E_j|$ is the total number of edges for samples *i* and *j*.

GO and KEGG Analysis

To elucidate the underlying biological functions, the RNA-Seq expression data matrix, the expression matrix of genes from overlapping gene pairs, genes within individual-patient-PPI networks, hub gene list for each patient, and the overlapping gene pairs of each cluster were analyzed using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses through KOBAS-I (a web-based tool for gene function annotation and pathway enrichment analysis, developed and maintained by the Center for Bioinformatics at Peking University in Beijing, China.

http://bioinfo.org/kobas). These analyses were performed using expression data enrichment and gene-list enrichment approaches.

Statistical Methods

Statistical analyses were conducted via SPSS (version 25.0, a proprietary statistical analysis software suite developed and maintained by IBM in Armonk, NY, USA). Continuous variables are summarized based on their distribution characteristics. Normally distributed data are presented as the mean ± standard deviation (SDs), and between-group comparisons were evaluated via two-tailed Student's unpaired *t*_tests. Non-normally distributed variables were expressed as median (interquartile range, IQR) with group comparisons performed via Mann-Whitney U tests (for two independent groups). A *p*-value of less than 0.05 and a false discovery rate of less than 0.05 were considered statistically significant.

Results

Transcriptomic Profile and Differentially Expressed Genes (DEGs) Between Control and Polycystic Ovary Syndrome (PCOS) Cohorts

The RNA sequencing (RNA-Seq) raw count matrix comprised 38,511 genomic features derived from 24 samples, including 12 PCOS patients and 12 control subjects. Transcript pseudoalignment, performed with Kallisto, yielded an average alignment rate of 57.73% ± 19.92% across all samples, with no statistically significant differences between the PCOS group (56.40% ± 21.64%) and the control group (59.06% ± 18.89%, p = 0.751) (**Supplementary Table 3**). After normalization, Uniform manifold approximation and projection (UMAP) visualization confirmed effective correction of batch effects and revealed pronounced separation among PCOS patients (Fig. 2A), suggesting transcriptional heterogeneity within the cohort.

DEGs were identified using DESeq2 to characterize molecular differences between the PCOS and control groups. Based on the criteria of |log2 fold change| ≥1.5 and adjusted p-values < 0.05, a total of 407 DEGs were identified between control and PCOS granulosa cells (GCs), including 68 downregulated and 339 upregulated genes (Fig. 2B; **Supplementary Table 4**). The expression patterns of the top 50 genes with the lowest p-values are shown in a heatmap (Fig. 2C), which demonstrates the potential to classify samples into two distinct groups.

Identification of Personalized Gene Pairs in Each Patient and Biofunctional Analysis

The RNA-Seq dataset initially comprised 38,511 genomic features. Of these, 1650 features were excluded due to zero counts across all 12 control samples. Subsequently, only genes with counts of ≥10 in all control samples were retained, resulting in the removal of 23,439 genes. This filtering process yielded a final set of 13,422 genes for downstream analysis. To construct PCOS patient networks, a Pearson's correlation matrix was generated using transcriptomic data from the control cohort, encompassing 90,068,331 correlations among the 13,422 genes (Fig. 1). The PCCs (r and r') for each gene pair were >0, indicating positive correlations between the expression levels.

By analyzing the single-patient transcriptomic profiles of 12 PCOS patients, a total of 13,275 genes (none of which were unique) and 3,274,238 correlations, including 2,250,324 unique and statistically significant gene pairs, were identified. UMAP plots of the dysregulated personalized gene pairs revealed marked separation among PCOS patients (Fig. 2D). Within this dataset, seven gene pairs were consistently dysregulated: alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase (*MGAT5*)-BMS1 ribosome biogenesis factor (*BMS1*), proprotein convertase subtilisin/kexin type 6 (*PCSK6*)-nicotinamide riboside kinase 1 (*NMRK1*), solute carrier family 2 member 1 (*SLC2A1*)-nicotinamide riboside kinase 1 (*NMRK1*), CTR9 component of Paf1/RNA polymerase II complex (*CTR9*)-SEC14 like lipid binding 2 (*SEC14L2*), glyceraldehyde-3-phosphate dehydrogenase, spermatogenic (*GAPDHS*)-GDP-mannose pyrophosphorylase B (*GMPPB*), mitochondrial ribosomal protein L13 (*MRPL13*)-kringle containing transmembrane protein 2 (*KREMEN2*), and Rho GTPase activating protein 28 (*ARHGAP28*)-ajuba LIM protein (*AJUBA*). Furthermore, 11,184 genes were common across all 12 patients, while 13,264 genes and 1314 gene pairs (involving 1302 genes) were found to be consistently dysregulated in more than 75% of the patients (Fig. 2E, **Supplementary Table 5**).

The number of personalized gene pairs varied across patients, both in terms of genes (median = 13,063, range 12,441–13,258) and gene pairs (median = 411,989, range 164,980–765,706) (Fig. 2F). Similarly, the number of gene pairs with increased correlation (ΔPCC >0; median = 8742, range 375–34,443) and those with decreased correlation (ΔPCC <0; median = 403,248, range 164,605–756,502) also varied between patients (Fig. 2F).

In total, 97,479 perturbed gene pairs were identified, alongside the loss of 3,176,407 correlations. Upon individual analysis of the single-patient transcriptomic profiles from the 12 PCOS patients, 352 gene pairs showed increased PCCs across the remaining samples. Among these, 123 exhibited an increase in sample counts, 19 showed a decrease, and 210 maintained equal counts (**Supplementary Table 6**). For the 1314 gene pairs consistently dysregulated in more than 75% of the patients, ΔPCC values were negative in all cases (**Supplementary Table 5**).

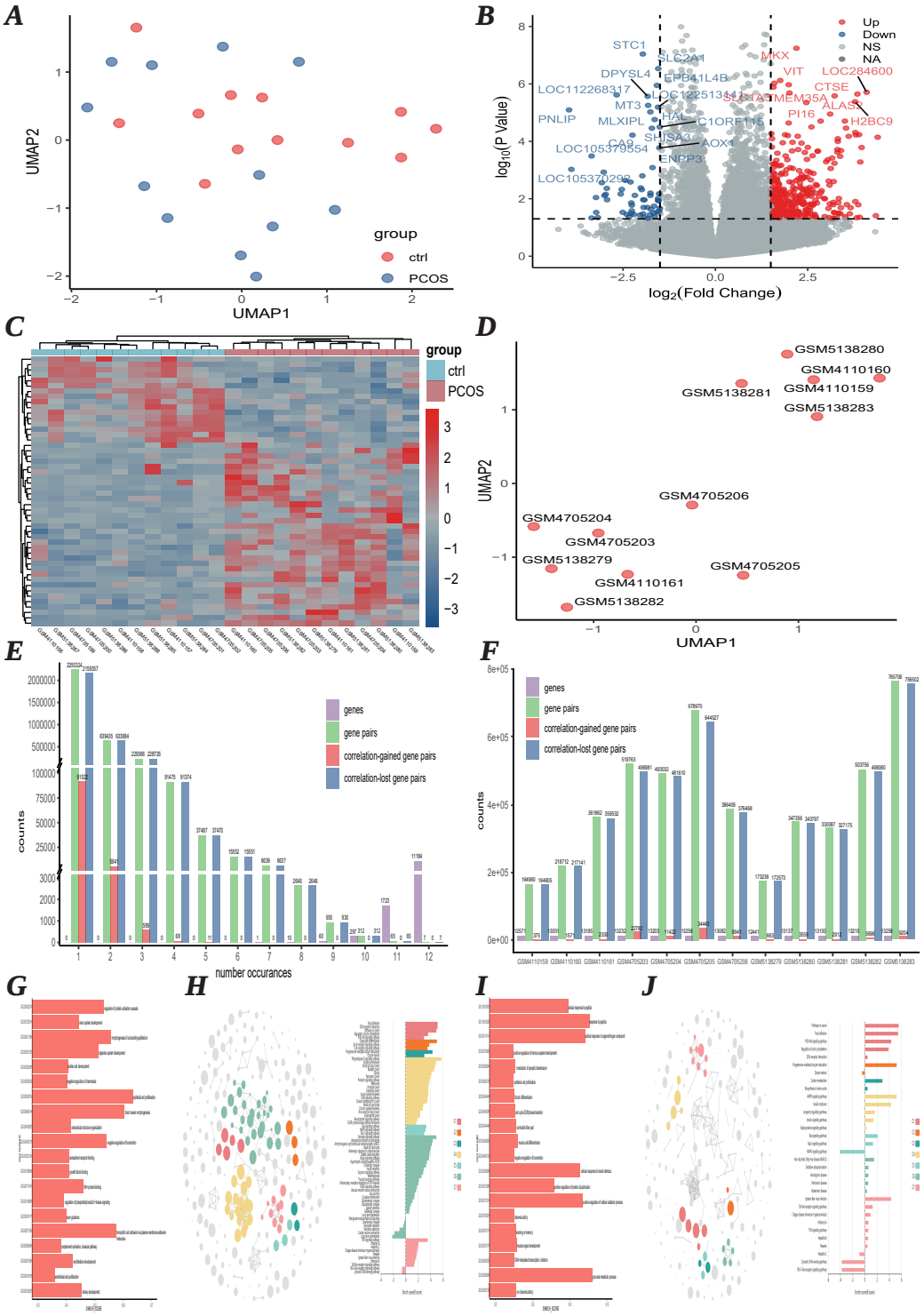

Fig. 2. Transcriptomic variance in PCOS patients compared with controls. (A) UMAP visualization of transcriptomic profiles following normalization and batch effect correction. (B) Volcano plots of DEGs. (C) Heatmaps of DEGs. (D) UMAP visualization of dysregulated personalized gene pairs. (E) Bar plot showing the frequency of genes, gene pairs, correlation-gained gene pairs, and correlation-lost gene pairs across PCOS samples (n = 12). (F) Bar plot showing the counts of genes, gene pairs, correlation-gained gene pairs, and correlation-lost gene pairs in each sample. (G,I) Bar plots of GO enrichment results for all genes in the RNA-Seq dataset (38,511 genes) and for the subset of genes in overlapping pairs (1302 genes), respectively. (H,J) cirFunMap visualizations of the KEGG pathway enrichment results for the RNA-Seq dataset (38,511 genes) and overlapping gene pairs (the 1302 genes), respectively. UMAP, uniform manifold approximation and projection; DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Table 2. Jaccard-containment of nodes and edges in individual-patient PCOS network pairs.

Node

	GSM4110159	GSM4110160	GSM4110161	GSM4705203	GSM4705204	GSM4705205	GSM4705206	GSM5138279	GSM5138280	GSM5138281	GSM5138282	GSM5138283	Key
GSM4110159		0.27	0.36	0.40	0.41	0.43	0.33	0.25	0.36	0.33	0.44	0.55	0.00
GSM4110160			0.36	0.43	0.43	0.44	0.33	0.25	0.36	0.37	0.44	0.55	0.15
GSM4110161				0.40	0.42	0.39	0.31	0.25	0.33	0.38	0.48	0.58	0.30
GSM4705203					0.46	0.39	0.33	0.25	0.31	0.32	0.43	0.54	0.45
GSM4705204						0.36	0.30	0.23	0.31	0.31	0.44	0.54	0.60
GSM4705205							0.39	0.25	0.29	0.30	0.41	0.51	
GSM4705206								0.26	0.32	0.33	0.44	0.55	
GSM5138279									0.32	0.33	0.49	0.56	
GSM5138280										0.36	0.49	0.53	
GSM5138281											0.42	0.54	
GSM5138282												0.59	
GSM5138283													

Edge

	GSM4110159	GSM4110160	GSM4110161	GSM4705203	GSM4705204	GSM4705205	GSM4705206	GSM5138279	GSM5138280	GSM5138281	GSM5138282	GSM5138283
GSM4110159		0.08	0.11	0.12	0.10	0.13	0.10	0.07	0.12	0.09	0.13	0.17
GSM4110160			0.13	0.19	0.13	0.15	0.11	0.10	0.15	0.16	0.15	0.20
GSM4110161				0.13	0.14	0.14	0.11	0.10	0.12	0.20	0.23	0.19
GSM4705203					0.14	0.08	0.08	0.07	0.08	0.10	0.11	0.16
GSM4705204						0.07	0.06	0.05	0.08	0.08	0.13	0.11
GSM4705205							0.21	0.10	0.10	0.10	0.13	0.13
GSM4705206								0.10	0.10	0.11	0.14	0.14
GSM5138279									0.10	0.11	0.20	0.13
GSM5138280										0.12	0.21	0.12
GSM5138281											0.14	0.15
GSM5138282												0.23
GSM5138283												

Notes: The top panel presents nodes (proteins in the networks) shared between individual patient-specific PCOS networks, while the bottom panel presents edges (protein-protein interactions) shared between individual patient-specific PCOS networks. The Jaccard-Containment index (JC) is defined as: $JC_{A \to B} = |A \cap B| / |A|$, where A and B are two networks being compared. A value of 1 indicates full overlap, while 0 indicates no overlap. Jaccard-Containment index (non-normal): Node = 0.385 (0.32–0.44), Edge = 0.12 (0.10–0.14); U = 1, $p < 0.0001$. The background colors in the table visually represent the Jaccard-Containment index (JC) values. This gradient color scale allows for intuitive interpretation of similarity or containment strength between datasets. The intensity of the color corresponds directly to the magnitude of JC.

Table 3. Network characteristics of individual PCOS patient-specific networks.

Patient ID	No. Nodes	No. Edges	Avg. No. of Neighbors	Diameter	Radius	Characteristic Path Length	Clustering Coefficient	Density	Hetero-geneity	Centra-lization	No. Connected Components
GSM4110159	1141	903	2.040	27	14	10.654	0.000	0.005	1.247	0.066	247
GSM4110160	1340	1080	2.049	31	16	12.591	0.000	0.006	1.076	0.049	274
GSM4110161	2197	2241	2.357	20	11	7.860	0.001	0.002	1.259	0.034	238
GSM4705203	2486	2897	2.677	23	13	7.585	0.008	0.001	1.165	0.016	228
GSM4705204	2562	3141	2.791	25	13	7.457	0.005	0.001	1.486	0.029	215
GSM4705205	2494	2339	2.239	27	14	9.228	0.002	0.001	1.357	0.046	343
GSM4705206	1819	1748	2.392	29	15	9.850	0.006	0.002	1.134	0.024	281
GSM5138279	1268	1154	2.225	20	11	8.064	0.003	0.003	1.328	0.041	199
GSM5138280	1836	1670	2.151	32	16	10.511	0.000	0.002	0.955	0.023	253
GSM5138281	1873	1752	2.218	22	12	8.387	0.000	0.002	1.079	0.024	251
GSM5138282	2679	3270	2.711	17	9	6.734	0.000	0.001	1.206	0.025	196
GSM5138283	3838	6237	3.481	16	8	5.916	0.003	0.001	1.316	0.019	161

Notes: Ave, average; No, number.

To further elucidate biological functions, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed on the 38,511 genes identified in the RNA-Seq dataset as well as on the 1302 genes within the overlapping gene pairs.

The GO enrichment analysis of the 38,511 genes identified 5785 GO categories, of which 1567 were significantly enriched (enrichmentRes = TRUE) (**Supplementary Table 7**). The most prominent enriched terms included epithelial cell proliferation (GO:0050673), blood vessel morphogenesis (GO:0048514), homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156), morphogenesis of a branching epithelium (GO:0061138), and negative regulation of locomotion (GO:0040013). The top 20 GO terms are shown in Fig. 2G. Similarly, KEGG enrichment analysis revealed 281 KEGG pathways, with 151 significantly enriched (enrichmentRes = TRUE) (**Supplementary Table 8**). These pathways together with their enrichment *p*-values (unfiltered) were visualized using cirFunMap, with correlation >0.25 and top n = 7 (Fig. 2H).

The GO enrichment analysis of the 1302 identified 5785 GO categories, of which 602 were significantly enriched (enrichmentRes = TRUE) (**Supplementary Table 9**). The top five terms included pyruvate metabolic process (GO:0006090), response to peptide (GO:1901652), cellular response to organonitrogen compound (GO:0071417), positive regulation of the cellular catabolic process (GO:0031331), and cellular response to insulin stimulus (GO:0032869). The top 20 enriched terms are presented in Fig. 2I. Concurrently, KEGG enrichment analysis of the same gene set revealed 281 KEGG functions, with 78 showing significant enrichment (enriched Res = TRUE) (**Supplementary Table 10**). These pathways, along with their calculated enrichment *p*-values (unfiltered), were visualized using cirFunMap (correlation >0.25; top n = 7) (Fig. 2J).

Individual-Patient Networks and Functional Enrichment Analysis

The dysregulated gene pairs were integrated into the consolidated human interactome to construct patient-specific protein-protein interaction (PPI) networks. In total, 7752 nodes (1978 unique), representing proteins within the networks, and 19,219 edges (13,626 unique), representing protein interactions, were used to construct individual-patient PPI networks (Fig. 3A, **Supplementary Table 11**). Among these nodes, five nodes (fibronectin 1 (*FN1*), DNA ligase 3 (*LIG3*), vesicle associated membrane protein 2 (*VAMP2*), kinesin family member 23 (*KIF23*), pescadillo ribosomal biogenesis factor 1 (*PES1*)) were common across all patients, whereas no edges were universally shared. Additionally, 1629 nodes and 3488 edges were shared between 2 patients (Fig. 3A).

The patient networks demonstrated marked variability in terms of genes (median = 2128, range = 1141–3838) and gene pairs (median = 2369, range = 903–6237). Similarly, the number of gene pairs with correlation gain (ΔPCC >0; median = 33, range = 1–209) and correlation loss (ΔPCC <0; median = 2337, range = 901–6028) also varied considerably (Fig. 3B). Graphical representations of the individual-patient PPI networks are shown in Fig. 3C.

Although pairwise comparisons revealed shared nodes and edges across patient networks, the proportion of shared elements varied considerably. Notably, node overlap (median Jaccard-Containment Index [JC] = 0.385, IQR = 0.32–0.44) was significantly greater than edge overlap (median JC = 0.12, IQR = 0.10–0.14; $p < 0.0001$) (Table 2).

To obtain a more comprehensive characterization of the individual-patient PPI networks, multiple topological network features, including average number of neighbors, network diameter, network radius, characteristic path length, clustering coefficient, network density, network heterogeneity, network centralization, and connected components, were calculated for each PCOS patient (Table 3).

Fig. 3. Individual protein-protein interaction (PPI) networks (reticulotypes) for all PCOS patients. (A) Bar plot showing the frequency of utilization of genes, gene pairs, correlation-gained gene pairs, and correlation-lost gene pairs used to construct individual-patient networks across samples within the PCOS cohort (n = 12). (B) Bar plot showing the counts of genes, gene pairs, correlation-gained pairs, and correlation-lost pairs employed in the construction of individual-patient networks for each sample. (C) Individual-patient PPI network for all PCOS patients, with yellow nodes indicating the top 20 hub genes ranked by node degree.

Fig. 4. cirFunMap visualization of GO and KEGG enrichment results for the top 20 hub genes. Left: bar plot of GO enrichment results for the top 20 hub genes, ranked by *p*-value of each term. Middle: circular network representation of KEGG enrichment results. Right: bar plot of the *p*-value for KEGG terms across different clusters.

Notably, the GSM5138283 network was larger than those of other PCOS patients.

To explore the biological functions of these networks, GO and KEGG enrichment analyses were performed on the nodes of each individual-patient network. The GO enrichment analysis identified 2460 enriched categories (corrected p-value < 0.01, **Supplementary Table 12**). Notably, 182 GO terms were consistently observed across all individual-patient-PPI networks. The most prevalent GO terms included protein binding (GO:0005515), cytosol (GO:0005829), nucleoplasm (GO:0005654), nucleus (GO:0005634), cytoplasm (GO:0005737), RNA binding (GO:0003723), membrane (GO:0016020), extracellular exosome (GO:0070062), ATP binding (GO:0005524), and identical protein binding (GO:0042802). The top 20 GO terms for each individual-patient PPI network are illustrated in **Supplementary Fig. 1**.

The KEGG enrichment analysis identified 329 functions (**Supplementary Table 13**), of which 278 were significantly enriched (corrected p-value < 0.01, **Supplementary Table 14**). Among these, 22 KEGG signaling pathways and 58 non-signaling KEGG functions were consistently shared across all individual-patient PPI networks. The enriched terms, along with calculated enriched p-values (unfiltered), were subsequently visualized using cirFunMap with a correlation >0.35 and top n = 7 (**Supplementary Fig. 1**).

Hub Genes of Each Individual-Patient Network and Functional Enrichment Analysis

Hub genes play crucial roles in regulating other genes within a network and are often associated with diseases, contributing significantly to the development of complex traits. In this study, hub genes were defined as those with a connectivity degree >20, indicating the number of direct interactions with other genes. From the analysis of the individual-patient PPI network, a total of 168 genes (125 unique) were identified as hub genes based on node degree (Fig. 3C, **Supplementary Table 15**). Notably, two genes, *KIF23* and protein regulator of cytokinesis 1 (*PRC1*) consistently emerged as common hub genes when restricted to patients in which these genes were detected in more than 50% of the cases (*KIF23*: GSM4110159, GSM4705203, GSM4705205, GSM5138280, GSM5138282, GSM5138283; *PRC1*: GSM4110161, GSM4705205, GSM4705296, GSM5138279, GSM5138280, GSM5138283).

To elucidate the biological functions, GO and KEGG gene-list enrichment analyses were performed for each of the top 20 nodes. GO enrichment analysis identified 267 significantly enriched GO categories (corrected p-value < 0.01, **Supplementary Table 16**). Notably, the GO term protein binding (GO:0005515) was consistently enriched across all the individual-patient PPI networks. Additionally, the terms cytosol (GO:0005829), cytoplasm (GO:0005737), nucleoplasm (GO:0005654), RNA binding (GO:0003723), membrane (GO:0016020), and nucleus (GO:0005634) were frequently enriched when restricted to symbols present in more than 50% of the networks. The top 20 GO terms for each of the top 20 nodes are shown in Fig. 4. KEGG enrichment analysis revealed 230 KEGG pathways (**Supplementary Table 17**), of which 179 were significantly enriched (corrected p-value < 0.05, **Supplementary Table 18**). Among these, pathways in cancer (hsa05200), PI3K-Akt signaling pathway (hsa04151), and endometrial cancer (hsa05213) were common in more than 50% of the patient datasets. The enriched terms were visualized using cirFunMap, with correlation >0.35 and top n = 7 (Fig. 4).

Molecular Subtypes and Functional Enrichment Analysis

The individual-patient PPI networks were clustered based on network distances using the *ConsensusClusterPlus* R package. To determine the optimal number of clusters, the average clustering consistency and intercluster variation coefficients were calculated across potential cluster numbers. The optimal clustering solution was identified via the cumulative distribution function (CDF), as illustrated in Fig. 5A, where the clustering stabilized at k = 5. Further assessment of the CDF delta area curve revealed that the area under the CDF curve plateaued beyond 5 clusters (Fig. 5B). Subsequently, a consensus matrix graph was constructed, assigning each individual-patient PPI network to one of the five clusters, enabling intuitive evaluation of cluster composition and quantity (Fig. 5C).

To further elucidate the biological functions, GO and KEGG enrichment analyses were performed on genes that appeared in more than two samples within each cluster (clusters 3, 4, and 5). GO enrichment identified 702 significantly enriched GO categories (corrected p-value < 0.01, **Supplementary Table 19**), with 61 categories common to all clusters. The top 20 GO terms for each cluster are shown in Fig. 5D. Similarly, KEGG enrichment analysis identified 312 KEGG functional pathways (**Supplementary Table 20**), of which 179 were significantly enriched (corrected p-value < 0.01, **Supplementary Table 21**). Among these, 21 pathways were shared across all clusters, including the glucagon signaling pathway (hsa04922), HIF-1 signaling pathway (hsa04066), AMPK signaling pathway (hsa04152), and estrogen signaling pathway (hsa04915), highlighting notable common signaling pathways. The enriched terms for each symbol, along with their calculated p-values (unfiltered), were visualized using cirFunMap, with correlation >0.35 and top n = 7 (Fig. 5D).

Discussion

This study used transcriptomic data from granulosa cells (GCs) of Polycystic ovary syndrome (PCOS) patients

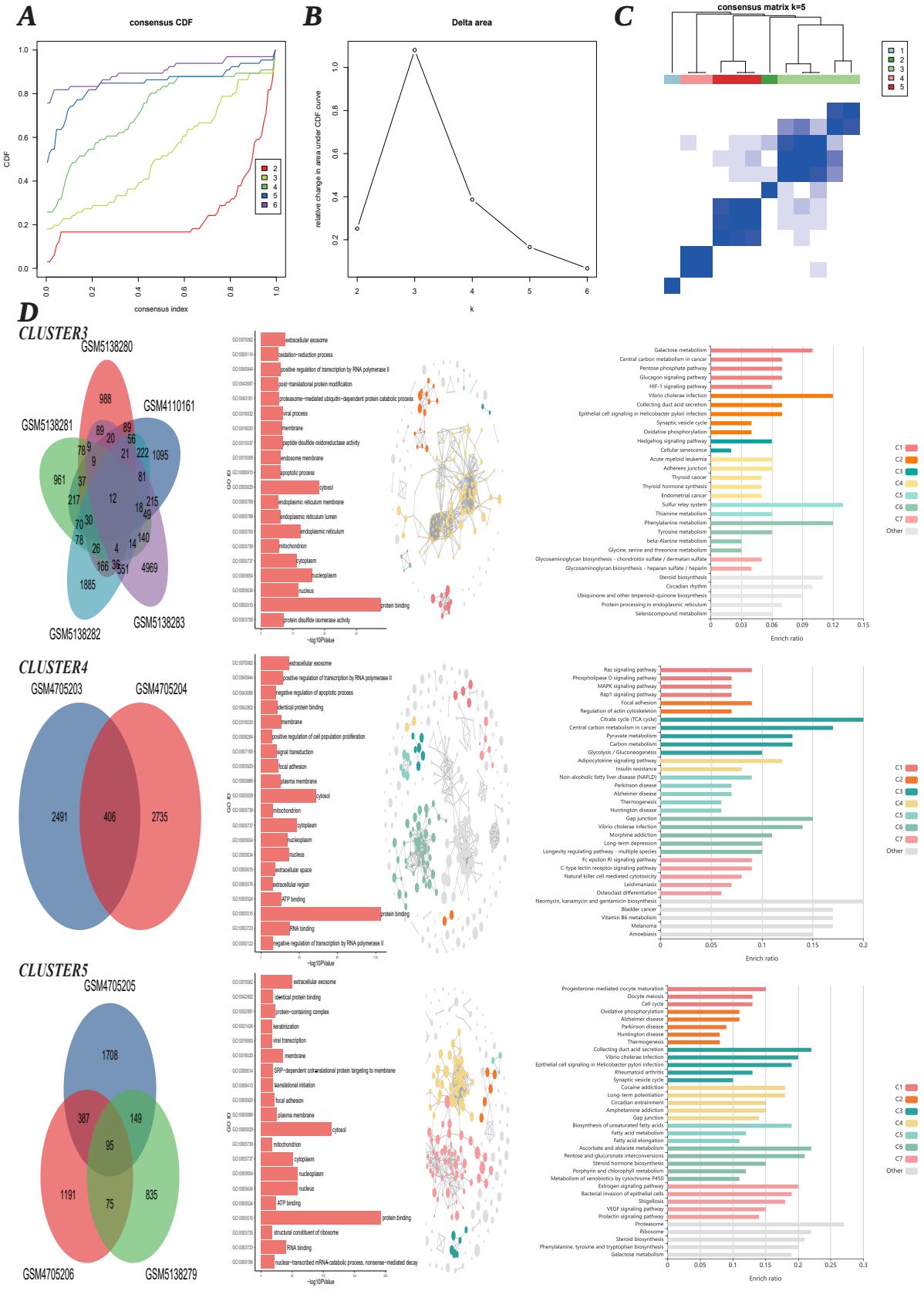

Fig. 5. Consensus clustering of PCOS molecular subgroups based on individual–patient network distances. (A) Cumulative distribution function (CDF) curve. (B) CDF delta area curve. (C) Heatmap of clustering results. (D) Venn diagram showing overlapping gene pairs across samples in clusters 3, 4, and 5. Right panel: cirFunMap visualization of GO and KEGG enrichment results for genes recurrently observed (≥2 occurrences) across samples within clusters 3, 4, and 5.

to construct personalized protein-protein interaction (PPI) networks, revealing patient-specific transcriptional heterogeneity. While common network components were observed, gene expression patterns enabled the classification of distinct transcriptional subtypes in PCOS. Clustering analysis of 12 patients identified 5 subtypes, followed by functional enrichment analysis. However, the limited sample size and lack of clinical information may limit the generalizability of these findings, underscoring the need for further investigation.

To identify key candidate genes, we initially identified five overlapping nodes (fibronectin 1 (*FN1*), DNA ligase 3 (*LIG3*), vesicle associated membrane protein 2 (*VAMP2*), kinesin family member 23 (*KIF23*), pescadillo ribosomal biogenesis factor 1 (*PES1*)) consistently present across all patient-specific networks. Hub genes within each network were then determined based on node degree, with the top 20 hubs selected for each patient. Among these, *KIF23* and protein regulator of cytokinesis 1 (*PRC1*) were the only hub genes shared by more than 50% of patients. *KIF23* was prioritized as the principal candidate hub gene in PCOS GCs due to its ubiquitous occurrence, being the sole gene detected in 100% of patient-specific networks, and ranking among the top hub genes in over half of the patients. Conversely, although *PRC1* was also common, appearing in more than 50% of patients, it was not consistently present in all networks.

Although direct experimental validation in granulosa cells from PCOS patients has not yet been conducted, accumulating evidence underscores the potential pathophysiological relevance of *KIF23*. First, as a pivotal regulator of cytokinesis and G2/M phase progression, *KIF23* is essential for mitosis, a process fundamental to follicular development [22,23]. Second, its marked downregulation in the endometrium during the implantation window suggests that *KIF23* may contribute to decidualization defects in PCOS [24]. Third, evidence from cancer models demonstrates androgen receptor (AR)-mediated transcriptional upregulation of *KIF23*, which promotes Wnt/β-catenin-dependent proliferation [25]. In contrast, in PCOS granulosa cells, proliferation is suppressed. This discrepancy is likely attributable to the failure of organism-level compensatory mechanisms to mitigate testosterone-induced epigenetic G1/S arrest mediated via the Casein Kinase 2 alpha (CK2α)-Histone Deacetylase 2 (HDAC2)-Histone H3 lysine 27 acetylation (H3K27ac) axis [26], together with Forkhead box K1 (FOXK1)/AMP-activated protein kinase (AMPK)/Mechanistic target of rapamycin (mTOR)-driven inhibition of proliferation through enhanced autophagy [27]. These findings highlight the potential role of *KIF23* in the pathophysiology of PCOS.

A comparative pathway analysis demonstrated that personalized network data significantly expanded the range of signaling pathways associated with PCOS compared to conventional RNA sequencing (RNA-Seq) differential expression methods. The analysis identified 18 consensus KEGG pathways, including Forkhead box O (FoxO), Phosphatidylinositol 3-Kinase/Ak strain transforming (PI3K-Akt), Tumor Necrosis Factor (TNF), Hypoxia-Inducible Factor-1 (HIF-1), and Mitogen-Activated Protein Kinase (MAPK), among several other signaling cascades, with four pathways, Wingless/Int-1 (Wnt), Apelin, cyclic Guanosine Monophosphate - Protein Kinase G (cGMP-PKC), and Relaxin, exhibiting patient-specific variation.

The Wnt signaling pathway is a key regulator of diverse biological processes. Wnt, a secreted glycoprotein of approximately 40 kD composed of approximately 350 amino acid residues with conserved cysteine domains, transmits signals through frizzled receptors (Fzd1-10) and low-density lipoprotein receptor-related proteins (LRP5/6). Extensive research has demonstrated its essential roles in animal growth and development, cellular metabolism, and stem cell maintenance [28,29], as well as in regulating ovarian granulosa cell proliferation, apoptosis, secretion, and interactions with growth factors [28,30–33]. Notably, the Wnt/β-catenin pathway mediates environmental toxin 2,5-hexanedione (2,5-HD)-induced cell cycle arrest in ovarian granulosa cells, involving dysregulation of Fzd1, Fzd6, LRP5/6, and β-catenin, alongside significant alterations in miRNAs such as *miR-145-5p* [34]. Furthermore, constitutive stabilization of β-catenin (CTNNB1) in pregranulosa/granulosa cells within the canonical Wnt/β-catenin pathway promotes the morphological transition of pregranulosa cells from squamous to cuboidal, a crucial step for supporting oocyte growth, though it does not influence oocyte activation [35].

Apelin, a G protein-coupled receptor and a member of the adipokine family, participates in diverse physiological processes including fluid homeostasis, food intake, insulin secretion, and energy metabolism. It activates the AMP-dependent protein kinase (AMPK) pathway, which is central to glucose metabolism and fatty acid oxidation for energy metabolism regulation [36,37]. Both Apelin (APLN) and its receptor (APLNR) have been detected in human ovarian cells, and increasing evidence suggests that dysregulation of the APLN system contributes to PCOS pathophysiology [37,38]. Although its expression pattern in patients is complex and influenced by metabolic status, APLN has been shown to enhance Insulin-like Growth Factor 1 (IGF-1)-induced steroidogenesis in granulosa cells by upregulating 3β-Hydroxysteroid Dehydrogenase/Isomerase (HSD3B) protein and activating Mitogen-Activated Protein Kinase 3/1 (MAPK3/1) and Ak strain transforming (Akt) signaling pathways [37,39,40]. Furthermore, APLN modulates glucose and lipid metabolism and may influence insulin regulation in individuals with PCOS [37,41,42]. Hyperinsulinemia and vascular endothelial growth factor (VEGF) may also alter ovarian function in PCOS patients, potentially through direct or indirect effects on the APLN/APLNR system, contributing to infer-

tility [37,42]. Overall, the literature suggests that APLN generally functions as a protective adipokine [43]. A compensatory rise in APLN levels may occur in PCOS in response to the metabolic imbalance, serving as a modulator of disrupted energy metabolism and associated ovarian dysfunction [37,38,44].

While the Wnt and apelin pathways have been partially characterized in PCOS research, the cGMP-PKG and Relaxin pathways remain largely unexplored.

This study is subject to several limitations. The reliance on GEO datasets, which lack detailed clinical metadata such as body mass index (BMI), insulin resistance, and hormone levels, limits the ability to reliably correlate molecular subtypes with polycystic ovary syndrome (PCOS) phenotypes. Furthermore, methodological limitations restrict the capacity to gain multidimensional insights into the heterogeneity of PCOS. These limitations include the unvalidated functional roles of identified hub genes and pathways, the exclusive focus on transcriptomics without integration of complementary omics layers such as proteomics and metabolomics, limited statistical power due to small cohort sizes (n = 12 per group), and the inherent inability of RNA sequencing to fully capture posttranscriptional dynamics. Therefore, future research should: (1) utilize multicenter cohorts with standardized clinical datasets to facilitate robust stratification of phenotypic subgroups, (2) integrate multi-omics approaches, including proteomics, metabolomics, and epigenomics, alongside single-cell RNA sequencing to resolve molecular-phenotypic discordance and regulatory hierarchies, and (3) experimentally validate prioritized targets, such as *KIF23* and *PRC1*, in appropriate model systems.

Conclusion

In summary, this study revealed patient-specific variations in gene expression by characterizing distinct network properties across individuals, thereby highlighting the heterogeneity within the PCOS population. In addition to identifying candidate genes and signaling pathways enriched with DEGs, *KIF23* was highlighted as a potential hub gene. Additionally, the cGMP-PKG, Wnt, relaxin, and apelin signaling pathways were identified as potential central signaling pathways. Nonetheless, these findings warrant further experimental validation through extensive *in vivo* and *in vitro* studies to validate their biological relevance.

Availability of Data and Materials

The gene expression profile datasets GSE168404, GSE155489, and GSE138518 were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/).

Author Contributions

FZ and WW designed the research study; FZ, YHC, LLS, and WW performed the research; FZ and YHC collected and analyzed the data. FZ, LLS and WW were involved in drafting the manuscript, and all the authors were involved in critically revising it for important intellectual content. All the authors gave final approval of the version to be published. All authors have participated sufficiently in the work to take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of the work in ensuring questions related to its accuracy or integrity.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.24976/Discov.Med.202537201.201.

References

[1] Ge J, Yang N, Zhang X, Zhao X, Chen Y, Wang L. Steroid Hormone Profiling in Hyperandrogenism and Non-hyperandrogenism Women with Polycystic Ovary Syndrome. Reproductive Sciences. 2022; 29: 3449–3458. https://doi.org/10.1007/s43032-022-00985-0.

[2] Ding H, Xiang Y, Zhu Q, Wu H, Xu T, Huang Z, *et al*. Endoplasmic reticulum stress-mediated ferroptosis in granulosa cells contributes to follicular dysfunction of polycystic ovary syndrome driven by hyperandrogenism. Reproductive Biomedicine Online. 2024; 49: 104078. https://doi.org/10.1016/j.rbmo.2024.104078.

[3] Zhang KH, Zhang FF, Zhang ZL, Fang KF, Sun WX, Kong N, *et al*. Follicle stimulating hormone controls granulosa cell glutamine synthesis to regulate ovulation. Protein & Cell. 2024; 15: 512–529. https://doi.org/10.1093/procel/pwad065.

[4] Han S, Xie Y, Lv J, Sun X, Shi Y. CNOT6L regulates energy metabolism in the ovarian granulosa cells associated with polycystic ovary syndrome. Frontiers in Cell and Developmental Biology. 2025; 13: 1607161. https://doi.org/10.3389/fcell.2025.1607161.

[5] Mu L, Ye Z, Hu J, Zhang Y, Chen K, Sun H, *et al*. PPM1K-regulated impaired catabolism of branched-chain amino acids

orchestrates polycystic ovary syndrome. eBioMedicine. 2023; 89: 104492. https://doi.org/10.1016/j.ebiom.2023.104492.

[6] Liao B, Chen W, Qi X, Yun C, Pang Y. Interleukin-22 improves ovulation in polycystic ovary syndrome via STAT3 signaling. Molecular Human Reproduction. 2024; 30: gaae037. https://doi.org/10.1093/molehr/gaae037.

[7] Zhao M, Liao B, Yun C, Qi X, Pang Y. Liraglutide improves follicle development in polycystic ovary syndrome by inhibiting CXCL10 secretion. Reproductive Biology and Endocrinology. 2024; 22: 98. https://doi.org/10.1186/s12958-024-01269-9.

[8] Erol S, Zırh S, Bozdag G, Sokmensuer LK, Muftuoglu SF. In vitro evaluation of exocytosis-associated SNARE molecules in human granulosa cells in polycystic ovary syndrome. Journal of Assisted Reproduction and Genetics. 2024; 41: 49–61. https://doi.org/10.1007/s10815-023-02967-w.

[9] Yildirim E, Onel T, Agus S, Gunalan E, Yilmaz B, Aydin MS, *et al*. The effect of rapamycin treatment on mouse ovarian follicle development in dehydroepiandrosterone-induced polycystic ovary syndrome mouse model. Zygote. 2024; 32: 386–395. https://doi.org/10.1017/S0967199424000388.

[10] Huang J, Chen P, Xiang Y, Liang Q, Wu T, Liu J, *et al*. Gut microbiota dysbiosis-derived macrophage pyroptosis causes polycystic ovary syndrome via steroidogenesis disturbance and apoptosis of granulosa cells. International Immunopharmacology. 2022; 107: 108717. https://doi.org/10.1016/j.intimp.2022.108717.

[11] Zanjirband M, Baharlooie M, Safaeinejad Z, Nasr-Esfahani MH. Transcriptomic screening to identify hub genes and drug signatures for PCOS based on RNA-Seq data in granulosa cells. Computers in Biology and Medicine. 2023; 154: 106601. https://doi.org/10.1016/j.compbiomed.2023.106601.

[12] Yang L, Chen J, Miao H, Li N, Bi H, Feng R, *et al*. The landscape of alternative splicing in granulosa cells and a potential novel role of YAP1 in PCOS. PLoS ONE. 2024; 19: e0315750. https://doi.org/10.1371/journal.pone.0315750.

[13] Yuan Y, Daiterigele, Zhang Q, Du C. Whole transcriptome analysis and construction of gene regulatory networks of granulosa cells from patients with polycystic ovary syndrome (PCOS). European Journal of Medical Research. 2025; 30: 9. https://doi.org/10.1186/s40001-024-02237-0.

[14] Li Q, Sang Y, Chen Q, Ye B, Zhou X, Zhu Y. Integrated bioinformatics analysis elucidates granulosa cell whole-transcriptome landscape of PCOS in China. Journal of Ovarian Research. 2023; 16: 154. https://doi.org/10.1186/s13048-023-01223-0.

[15] Huang X, Geng H, Liang C, Xiong X, Du X, Zhuan Q, *et al*. Leonurine restrains granulosa cell ferroptosis through SLC7A11/GPX4 axis to promote the treatment of polycystic ovary syndrome. Free Radical Biology & Medicine. 2025; 226: 330–347. https://doi.org/10.1016/j.freeradbiomed.2024.11.021.

[16] Deng Y, Li H, Song Y, Cen J, Zhang Y, Sui Y, *et al*. Whole Genome Transcriptomic Analysis of Ovary Granulosa Cells Revealed an Anti-Apoptosis Regulatory Gene DLGAP5 in Polycystic Ovary Syndrome. Frontiers in Endocrinology. 2022; 13: 781149. https://doi.org/10.3389/fendo.2022.781149.

[17] Liu K, Wei H, Nong W, Peng H, Li Y, Lei X, *et al*. Nampt/SIRT2/LDHA pathway-mediated lactate production regulates follicular dysplasia in polycystic ovary syndrome. Free Radical Biology & Medicine. 2024; 225: 776–793. https://doi.org/10.1016/j.freeradbiomed.2024.10.312.

[18] Maron BA, Wang RS, Shevtsov S, Drakos SG, Arons E, Wever-Pinzon O, *et al*. Individualized interactomes for network-based precision medicine in hypertrophic cardiomyopathy with implications for other clinical pathophenotypes. Nature Communications. 2021; 12: 873. https://doi.org/10.1038/s41467-021-21146-y.

[19] Zhao R, Jiang Y, Zhao S, Zhao H. Multiomics Analysis Reveals Molecular Abnormalities in Granulosa Cells of Women With Polycystic Ovary Syndrome. Frontiers in Genetics. 2021; 12: 648701. https://doi.org/10.3389/fgene.2021.648701.

[20] Li J, Chen H, Gou M, Tian C, Wang H, Song X, *et al*. Molecular Features of Polycystic Ovary Syndrome Revealed by Transcriptome Analysis of Oocytes and Cumulus Cells. Frontiers in Cell and Developmental Biology. 2021; 9: 735684. https://doi.org/10.3389/fcell.2021.735684.

[21] Mao Z, Li T, Zhao H, Qin Y, Wang X, Kang Y. Identification of epigenetic interactions between microRNA and DNA methylation associated with polycystic ovarian syndrome. Journal of Human Genetics. 2021; 66: 123–137. https://doi.org/10.1038/s10038-020-0819-6.

[22] Hutterer A, Glotzer M, Mishima M. Clustering of centralspindlin is essential for its accumulation to the central spindle and the midbody. Current Biology. 2009; 19: 2043–2049. https://doi.org/10.1016/j.cub.2009.10.050.

[23] Seguin L, Liot C, Mzali R, Harada R, Siret A, Nepveu A, *et al*. CUX1 and E2F1 regulate coordinated expression of the mitotic complex genes Ect2, MgcRacGAP, and MKLP1 in S phase. Molecular and Cellular Biology. 2009; 29: 570–581. https://doi.org/10.1128/MCB.01275-08.

[24] Sutaji Z, Elias MH, Ahmad MF, Karim AKA, Abu MA. A Systematic Review and Integrated Bioinformatic Analysis of Candidate Genes and Pathways in the Endometrium of Patients With Polycystic Ovary Syndrome During the Implantation Window. Frontiers in Endocrinology. 2022; 13: 900767. https://doi.org/10.3389/fendo.2022.900767.

[25] Xu H, Liu J, Zhang Y, Zhou Y, Zhang L, Kang J, *et al*. KIF23, under regulation by androgen receptor, contributes to nasopharyngeal carcinoma deterioration by activating the Wnt/β-catenin signaling pathway. Functional & Integrative Genomics. 2023; 23: 116. https://doi.org/10.1007/s10142-023-01044-w.

[26] Tong X, Hu Z, Zhou H, Zhang Y, Zhang YL, Zhang S, *et al*. Testosterone-Induced H3K27 Deacetylation Participates in Granulosa Cell Proliferation Suppression and Pathogenesis of Polycystic Ovary Syndrome. The American Journal of Pathology. 2024; 194: 2326–2340. https://doi.org/10.1016/j.ajpath.2024.08.012.

[27] Wang C, Yu J, Ding C, Chen C. CangFu Daotan decoction improves polycystic ovarian syndrome by downregulating FOXK1. Gynecological Endocrinology. 2023; 39: 2244600. https://doi.org/10.1080/09513590.2023.2244600.

[28] Qiao GY, Dong BW, Zhu CJ, Yan CY, Chen BL. Deregulation of WNT2/FZD3/β-catenin pathway compromises the estrogen synthesis in cumulus cells from patients with polycystic ovary syndrome. Biochemical and Biophysical Research Communications. 2017; 493: 847–854. https://doi.org/10.1016/j.bbrc.2017.07.057.

[29] Clevers H, Nusse R. Wnt/β-catenin signaling and disease. Cell. 2012; 149: 1192–1205. https://doi.org/10.1016/j.cell.2012.05.012.

[30] Gustin SE, Hogg K, Stringer JM, Rastetter RH, Pelosi E, Miles DC, *et al*. WNT/β-catenin and p27/FOXL2 differentially regulate supporting cell proliferation in the developing ovary. Developmental Biology. 2016; 412: 250–260. https://doi.org/10.1016/j.ydbio.2016.02.024.

[31] Du X, Li Q, Yang L, Liu L, Cao Q, Li Q. SMAD4 activates Wnt signaling pathway to inhibit granulosa cell apoptosis. Cell Death & Disease. 2020; 11: 373. https://doi.org/10.1038/s41419-020-2578-x.

[32] Wu XQ, Wang YQ, Xu SM, Liu JF, Bi XY, Wang ZQ, *et al*. The WNT/β-catenin signaling pathway may be involved in granulosa cell apoptosis from patients with PCOS in North China. Journal of Gynecology Obstetrics and Human Reproduction. 2017; 46: 93–99. https://doi.org/10.1016/j.jgyn.2015.08.013.

[33] Tepekoy F, Akkoyunlu G. The interaction of Wnt signaling members with growth factors in cultured granulosa cells. Animal Reproduction. 2020; 17: e20190106. https://doi.org/10.1590/1984-3143-AR2019-0106.

[34] Xu X, Pan Y, Zhan L, Sun Y, Chen S, Zhu J, *et al*. The Wnt/β-catenin pathway is involved in 2,5-hexanedione-induced ovarian granulosa cell cycle arrest. Ecotoxicology and Environmental Safety. 2023; 268: 115720. https://doi.org/10.1016/j.ecoenv.2023.115720.

[35] Habara O, Logan CY, Kanai-Azuma M, Nusse R, Takase HM. WNT signaling in pre-granulosa cells is required for ovarian folliculogenesis and female fertility. Development. 2021; 148: dev198846. https://doi.org/10.1242/dev.198846.

[36] Chen C, Zhao S, Chen Z, Wang J, Li Y, Zhang Q, *et al.* Exogenous [Pyr1]apelin-13 prevents bupivacaine-induced cardiotoxicity via the apelin (APJ) receptor. Clinical Toxicology. 2025; 63: 507–517. https://doi.org/10.1080/15563650.2025.2510528.

[37] Dravecká I, Figurová J, Lazúrová I. Is Apelin a new biomarker in patients with polycystic ovary syndrome? Physiological Research. 2021; 70: S635–S641. https://doi.org/10.33549/physiolres.934708.

[38] Pich K, Pietroń K, Szlaga A, Billert M, Skrzypski M, Pawlicki P, *et al*. Adipokines level in plasma, hypothalamus, ovaries and adipose tissue of rats with polycystic ovary syndrome. Reproductive Biomedicine Online. 2025; 50: 104693. https://doi.org/10.1016/j.rbmo.2024.104693.

[39] Roche J, Ramé C, Reverchon M, Mellouk N, Cornuau M, Froment P, Dupont J. Apelin (APLN) and Apelin Receptor (APLNR) in Human Ovary: Expression, Signaling, and Regulation of Steroidogenesis in Primary Human Luteinized Granulosa Cells. Biology of Reproduction. 2016; 95: 104. https://doi.org/10.1095/biolreprod.116.141754.

[40] Wang W, Tan L, Ge L, Gou R, Gou L, Liu L, *et al*. Circ_0043314 Modulates Proliferation and Apoptosis of Ovarian Granulosa Cells in Polycystic Ovarian Syndrome via the MicroRNA-146b-3p/Apelin 13 Axis. Gynecologic and Obstetric Investigation. 2025; 90: 18–29. https://doi.org/10.1159/000540097.

[41] Niepsuj J, Piwowar A, Franik G, Bizoń A. Impact of Smoking and Obesity on the Selected Peptide Hormones and Metabolic Parameters in the Blood of Women with Polycystic Ovary Syndrome-Preliminary Study. International Journal of Molecular Sciences. 2024; 25: 8713. https://doi.org/10.3390/ijms25168713.

[42] Bongrani A, Plotton I, Mellouk N, Ramé C, Guerif F, Froment P, *et al*. High androgen concentrations in follicular fluid of polycystic ovary syndrome women. Reproductive Biology and Endocrinology. 2022; 20: 88. https://doi.org/10.1186/s12958-022-00959-6.

[43] Ruan X, Li M, Min M, Ju R, Wang H, Xu Z, *et al*. Plasma visfatin and apelin levels in adolescents with polycystic ovary syndrome. Gynecological Endocrinology. 2023; 39: 2216807. https://doi.org/10.1080/09513590.2023.2216807.

[44] Kuai D, Tang Q, Wang X, Yan Q, Tian W, Zhang H. Relationship between serum apelin, visfatin levels, and body composition in Polycystic Ovary Syndrome patients. European Journal of Obstetrics, Gynecology, and Reproductive Biology. 2024; 297: 24–29. https://doi.org/10.1016/j.ejogrb.2024.03.034.