scientific reports

Check for updates

OPEN

Identification of potential diagnostic genes for atherosclerosis in women with polycystic ovary syndrome

Yujia Luo[1,6], Yuanyuan Zhou[2,6], Hanyue Jiang[3], Qiongjun Zhu[4], Qingbo Lv[4], Xuandong Zhang[1], Rui Gu[1], Bingqian Yan[1], Li Wei[1], Yuhang Zhu[5] & Zhou Jiang[1]✉

Polycystic ovary syndrome (PCOS), which is the most prevalent endocrine disorder among women in their reproductive years, is linked to a higher occurrence and severity of atherosclerosis (AS). Nevertheless, the precise manner in which PCOS impacts the cardiovascular well-being of women remains ambiguous. The Gene Expression Omnibus database provided four PCOS datasets and two AS datasets for this study. Through the examination of genes originating from differentially expressed (DEGs) and critical modules utilizing functional enrichment analyses, weighted gene co-expression network (WGCNA), and machine learning algorithm, the research attempted to discover potential diagnostic genes. Additionally, the study investigated immune infiltration and conducted gene set enrichment analysis (GSEA) to examine the potential mechanism of the simultaneous occurrence of PCOS and AS. Two verification datasets and cell experiments were performed to assess biomarkers' reliability. The PCOS group identified 53 genes and AS group identified 175 genes by intersecting DEGs and key modules of WGCNA. Then, 18 genes from two groups were analyzed by machine learning algorithm. Death Associated Protein Kinase 1 (DAPK1) was recognized as an essential gene. Immune infiltration and single-gene GSEA results suggest that DAPK1 is associated with T cell-mediated immune responses. The mRNA expression of DAPK1 was upregulated in ox-LDL stimulated RAW264.7 cells and in granulosa cells. Our research discovered the close association between AS and PCOS, and identified DAPK1 as a crucial diagnostic biomarker for AS in PCOS.

Keywords Polycystic ovary syndrome, Atherosclerosis, Bioinformatics analysis, Machine learning, Immune infiltration

Polycystic ovary syndrome (PCOS), a prevalent endocrine-metabolic disorder, affects a significant proportion (15% to 65%) of women in their reproductive years[1]. PCOS is a complex condition caused by genetics, hormones, metabolism, and environmental factors[2]. Diagnosing it involves clinicians primarily concentrating on anovulation, the morphology of polycystic ovaries, and hyperandrogenemia as key criteria[3]. According to recent research, it has been demonstrated that this condition is an endocrine and metabolic dysfunction that could make patients more susceptible to atherosclerosis (AS) and elevate the risk of cardiovascular problems[4]. The AS, the buildup of a fibrofatty plaque in the artery wall, along with the presence of immune cells like macrophages, T cells, and monocytes, can potentially lead to coronary and carotid artery disease[5,6]. Jabbour et al. reported that PCOS women had significantly increased intima-media thickness (IMT), a marker of AS, compared with healthy controls, which is a strong predictor of future cardiovascular events, including myocardial infarction and stroke[7]. Meanwhile, a meta-analysis by Lorenz et al. concluded that for every 0.1 mm increase in IMT, the risk of myocardial infarction increased by 10–15% and the risk of stroke increased by 13–18%[8].

Although the exact connection between PCOS and AS remains unidentified, it is evident that both disorders encompass persistent inflammation and immune infiltration[9,10]. For instance, metabolic, endocrine, and immune

[1]Department of NICU, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, China. [2]Department of Reproductive Endocrinology, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China. [3]Wenzhou Medical University, Wenzhou, China. [4]Department of Cardiology, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, China. [5]Women's Hospital, School of Medicine, Zhejiang University, Hangzhou 310006, Zhejiang, China. [6]These authors contributed equally: Yujia Luo and Yuanyuan Zhou. ✉email: 5200013@zju.edu.cn

dysfunction are caused by patients who are overweight and have PCOS. The impact of adipose tissue on PCOS pathology is enhanced by the presence of active immune cells (macrophages, T cells), which further contribute to endocrine and metabolic effects [11]. The immune cells residing in fat produce different cytokines (such as Interleukin-1, Tumor necrosis factor alpha [TNF-α], and Interleukin-6 [IL-6]), which play a role in the inflammatory cascade associated with PCOS[12]. AS is a harmful procedure that may result in the rupture of plaque, formation of blood clots, and blockage of blood vessels[13]. Furthermore, IL-6 derived from peripheral blood mononuclear cell secretion is significantly elevated in insulin-resistant PCOS patients, leading to an increased risk of AS[14]. Hence, the early detection of immune infiltration and related inflammatory substances could be beneficial in diagnosing PCOS patients with AS, thus helping to prevent serious cardiovascular complications.

This study was conducted to determine which biomarkers and potential pathways may be associated with PCOS and AS progression. In order to accomplish the objective, we screened transcriptomic information associated with PCOS and AS acquired from Gene Expression Omnibus (GEO). Subsequently, we utilized the 'limma' package and weighted gene co-expression network analysis (WGCNA) to detect genes that were expressed differently and crucial modules in each disease. By intersecting the results of three machine learning techniques, we successfully characterized Death Associated Protein Kinase 1 (DAPK1), which exhibited excellent performance that was confirmed by validation datasets. Additionally, we performed gene set enrichment analysis (GSEA) on DAPK1 to discover shared pathways linked to PCOS and AS. By analyzing immune infiltration, this study revealed the involvement of immune cells in the co-development of these two diseases. There is a possibility that elevated levels of T cell-mediated immune responses could be involved in the co-pathogenesis of PCOS and AS. Finally, we gathered granulosa cells from women who were healthy and had PCOS, along with ox-LDL stimulated macrophage, in order to confirm the dependability of DAPK1. To summarize, this research offers a valuable understanding of the common molecular processes that contribute to PCOS and AS, emphasizing the potential of DAPK1 as a diagnostic indicator for these disorders.

Methods

Gathering and organizing data

As both PCOS and AS were included in this study, GEO datasets related to both conditions were searched[15]. The study population included PCOS patients with high androgen levels, excluded patients with normal androgen levels, and did not consider the impact of obesity. The samples for PCOS were collected from granulosa cells of patients with hyperandrogenemia, chronic oligo/anovulation and polycystic ovarian morphology, which were considered typical cases. In order to reduce the impact of small sample size on the general applicability of the research results, there were three datasets numbered GSE137684, GSE114419, and GSE106724, which contained a total of 11 patients with PCOS and 11 controls. The research investigation centered on AS and employed the dataset GSE28829, comprising 29 samples of atherosclerotic plaques (13 early intimal thickening and intimal xanthoma, and 16 advanced thin or thick fibrous cap atheroma lesions) obtained from autopsied carotid arteries of human subjects from the Maastricht Pathology Tissue Collection. The patient information was handled in accordance with the "Code for Proper Secondary Use of Human Tissue". Additionally, GSE54250 was based on peripheral blood samples in 4 PCOS patients and 4 controls, whereas GSE43292 was conducted from paired pieces of carotid endarterectomy collected in 32 hypertensive patients. They were utilized for the purpose of externally validating PCOS and AS, correspondingly. The study utilized GEO datasets, and the details are provided in Table 1.

Analysis of differential gene expression

Next, by utilizing the R package called 'sva', which detected and constructed substitute variables for datasets with high dimensions, the batch effect was successfully eradicated. After preparing the data for each disease, we compared the PCOS and AS datasets separately using the R 'limma' package[16]. We used it to calculate differentially expressed genes (DEGs) between patients and controls. The DEG threshold of both diseases was defined as a P value less than 0.05 and an absolute log2 fold change greater than 1[17]. The heatmap and volcano plot were used to present the differential analysis results.

The analysis of gene co-expression networks using weights

The algorithm of WGCNA can detect gene modules exhibiting comparable expression patterns, explore a correlation between module and disease, and pinpoint gene biomarkers[18]. To obtain gene modules related to PCOS and AS, we utilized WGCNA to analyze the processing of raw data. A scale-free distribution network was

Diseases	GEO number	Number of patients	Number of controls	Group
PCOS	GSE114419	3	3	Discovery cohort
	GSE106724	4	4	Discovery cohort
	GSE137684	4	4	Discovery cohort
	GSE54250	4	4	Validation cohort
Atherosclerosis	GSE28829	13	16	Discovery cohort
	GSE43292	32	32	Validation cohort

Table 1. Details of GEO datasets used in the study.

constructed by calculating the suitable soft powers β (ranging from 1 to 20) utilizing the pickSoftThreshold function in the WGCNA package. To detect connectivity between gene modules, a topological overlap matrix (TOM) was derived using the adjacency matrix. Additional hierarchical clustering was conducted, leading to the identification of gene co-expression modules. After adjusting the minModuleSize to 60 and the mergeCutHeight to 0.25, we combined the comparable modules. In conclusion, we computed the eigengene module to depict the expression pattern of each module and examined its correlation with the clinical phenotype.

Functional enrichment of intersected genes

The combination of DEGs and modules identified by WGCNA allowed us to identify genes that contributed to PCOS and AS pathogenesis, respectively. To analyze the enrichment of Gene Ontology (GO), Disease Ontology (DO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, we employed the 'clusterProfier' package to identify the biological functions of genes[19]. The dot plots displayed the significant enrichment outcomes with a P-value less than 0.05.

Feature selection using machine learning algorithms

In order to explore the potential candidate genes for diagnosing PCOS with AS, we conducted analyses utilizing machine learning algorithms. A Least Absolute Shrinkage and Selection Operator (LASSO) can be used to improve prediction accuracy, which incorporates variable selection and regularization[20]. The Random Forest (RF) algorithm is a predictive model that can make predictions without any apparent variations by not imposing restrictions on variable conditions[21]. The identification of critical genes was achieved by applying a support vector machine recursive feature elimination (SVM-RFE) algorithm [22]. To perform machine learning analyses, we utilized the 'glmnet', 'randomforest', and 'e1071' packages of the R software. The overlap of the three findings can be considered potential hub genes for diagnosis[17].

Performance of forecasting for validation datasets

For testing the effectiveness of the diagnostic gene, the GSE54250 and GSE43292 were screened for validation of PCOS and AS. These original data were normalized utilizing 'RMA' package[23]. In validation cohorts, the boxplot demonstrated the diagnostic gene expression levels, and the calculation of area under curve (AUC) was performed.

Application of GSEA for individual biomarker

Using the 'clusterProfiler' package, a single-gene GSEA was performed to explore the functional signaling pathway in both groups after the diagnostic gene was acquired[24]. The gene sets were obtained by downloading MSigDB (c5.go.bp.v7.5.1.entrez.gmt)[25]. The enriched plots were utilized to display the top six pathways activated and inhibited in both disease categories.

Abundance of immune cells

The CIBERSORT (http://cibersort.stanford.edu/) analysis was performed on every disease sample to ascertain the relative quantities of immune cells[26]. By utilizing gene expression data, the CIBERSORT algorithm accurately determines the composition of immune cells through deconvolution. A gene signature based on LM22 has been used to quantify 22 types of immune cells with CIBERSORT and 500 iterations. For the subsequent examination, we chose samples for analysis that had a P value below 0.05. The results were visualized using the R packages 'corplot', 'vioplot', and 'ggplot2'[27]. Moreover, Spearman's correlation was applied to evaluate the association between immune infiltrating cells and biomarkers.

Collection of human samples and cell lines for verification

Approval for this research was granted by the Ethics Committee of the Obstetrics and Gynecology Hospital affiliated with Zhejiang University School of Medicine (Approval No. IRB-20230285-R). Granulosa cells were obtained from 5 patients with polycystic ovary syndrome (PCOS) and 5 control individuals. These patients were infertile women who received assisted reproductive technology treatment at the reproductive endocrinology department of Zhejiang University Women's Hospital in China. The 2023 Rotterdam Criteria were used as the basis for determining the inclusion criteria for patients with PCOS[28]. The AS cellular model is identical to the one described previously[29], and the summarized procedure is as follows: mouse macrophage RAW264.7-cell lines were obtained from ATCC (Passage 20–25, American Type Culture Collection, Rockville, MD, USA). They were converted into foam cells within 24 h using 40 μg/mL of ox-LDL (4.0 mg/mL, L34357, Thermo Fisher Scientific, Germany). Cells were collected for verification by reverse transcription polymerase chain reaction (RT-PCR). TRIzol Reagent (Invitrogen™, Japan) was used to extract the RNA from these samples, and HiScript III RT SuperMix for qPCR (Vazyme, Nanjing, China) was employed for cDNA synthesis. Next, the expression level of DAPK1 was quantified using real-time PCR by calculating 2 to the power of -ΔΔCt (2 -ΔΔCt). The primer sequences for DAPK1 were provided in Table 2.

Statistical analysis

Statistical analyses were conducted utilizing the R software (version 4.0.2). The Wilcoxon rank-sum test was utilized to assess differences among the groups, while Spearman correlation analysis was employed to calculate correlation coefficients between DEFRGs and immune cells. A significance level of two-tailed $p < 0.05$ was chosen to determine statistical significance.

Primers		Sequence (5′→3′)
DAPK1-Human	Forward	5′-ACGTGGATGATTACTACGACACC-3′
	Reverse	5′-TGCTTTTCTCACGGCATTTCT-3′
DAPK1-Mouse	Forward	5′-CAGATTCTCAGCGGCGTTTAC-3′
	Reverse	5′-GATCCGAGGTTTGGGCACATT-3′
GAPDH-Human	Forward	5′-GCCTCAAAATCCTCTCGTTGTG-3′
	Reverse	5′-GGAAGATGGTGATGGGATTTC-3′
GAPDH-Mouse	Forward	5′-AGGTCGGTGTGAACGGATTTG-3′
	Reverse	5′-TGTAGACCATGTAGTTGAGGTCA-3′

Table 2. Primers applied to Quantitative Real-Time PCR.

Ethics approval and consent to participate

The GEO database is publicly accessible. Ethics approval has been obtained for the patients involved in the database. The studies involving human participants were reviewed and approved by the ART Ethnics Committee of the Women's Hospital, School of Medicine, Zhejiang University (Approval No. IRB-20230285-R). The patients/participants provided their written informed consent to participate in this study. The study was completed in accordance with the declaration of Helsinki.

Results

Identification of DEGs in PCOS and AS

Figure 1 illustrates the complete analytical process. Next, the DEGs were identified utilizing 'limma' package in different groups. PCOS was connected to 69 DEGs, of which 55 were upregulated and 14 downregulated. Among 124 DEGs for AS, 113 were upregulated and 11 were downregulated. In Fig. 2A and C, volcano plots indicate all DEGs related to PCOS and AS. The heatmaps were used to visualize the DEGs in both groups collectively (Fig. 2B,D). The occurrence and development of PCOS and AS may be influenced by DEGs.

Screening for key modules by WGCNA

To investigate the potential relationship between diseases and crucial genes, we conducted WGCNA and also analyzed the differential expression in both groups. Our research built a co-expression network using the soft-threshold method. Maintaining a scale-free topology required the inclusion of parameter β in co-expression networks. The fit index exceeded 0.9 for PCOS group, indicating the presence of a scale-free topology, and β was set to 4 (Fig. 3A). Subsequently, we produced an adjacency matrix by utilizing the adjacency function. The TOM dissimilarity measure was utilized to construct hierarchical clustering (Fig. 3B). In total, we have discovered a total of 7 co-expression modules. Among them, the turquoise module exhibited the highest positive correlation

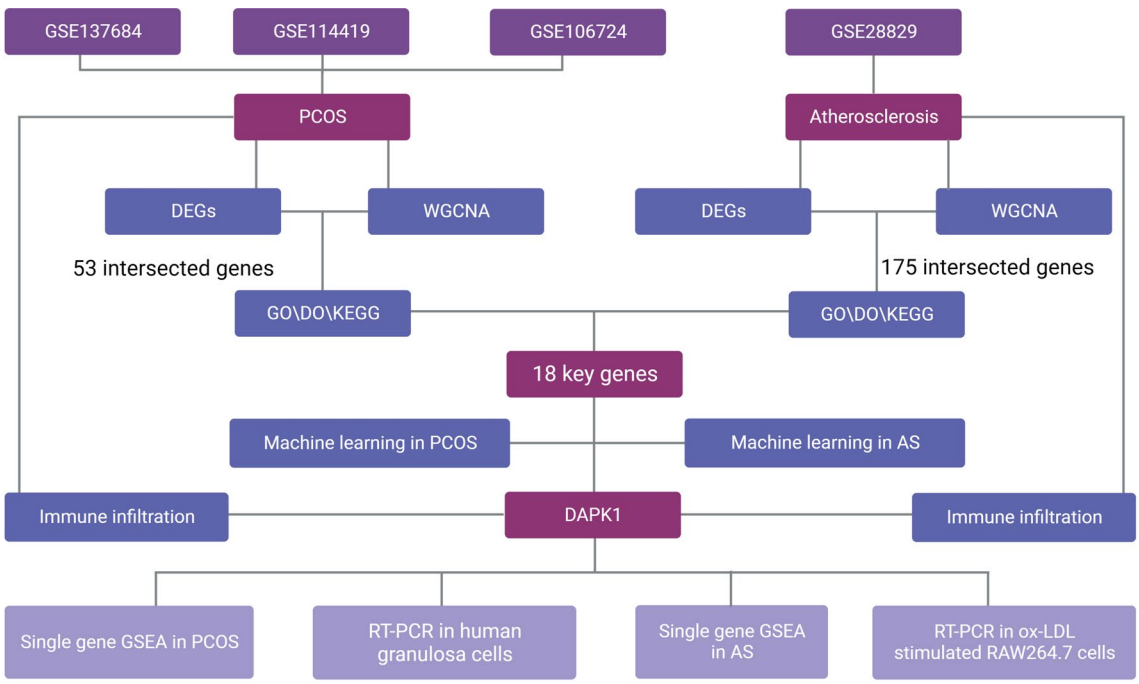

Figure 1. The flow chart for the whole design.

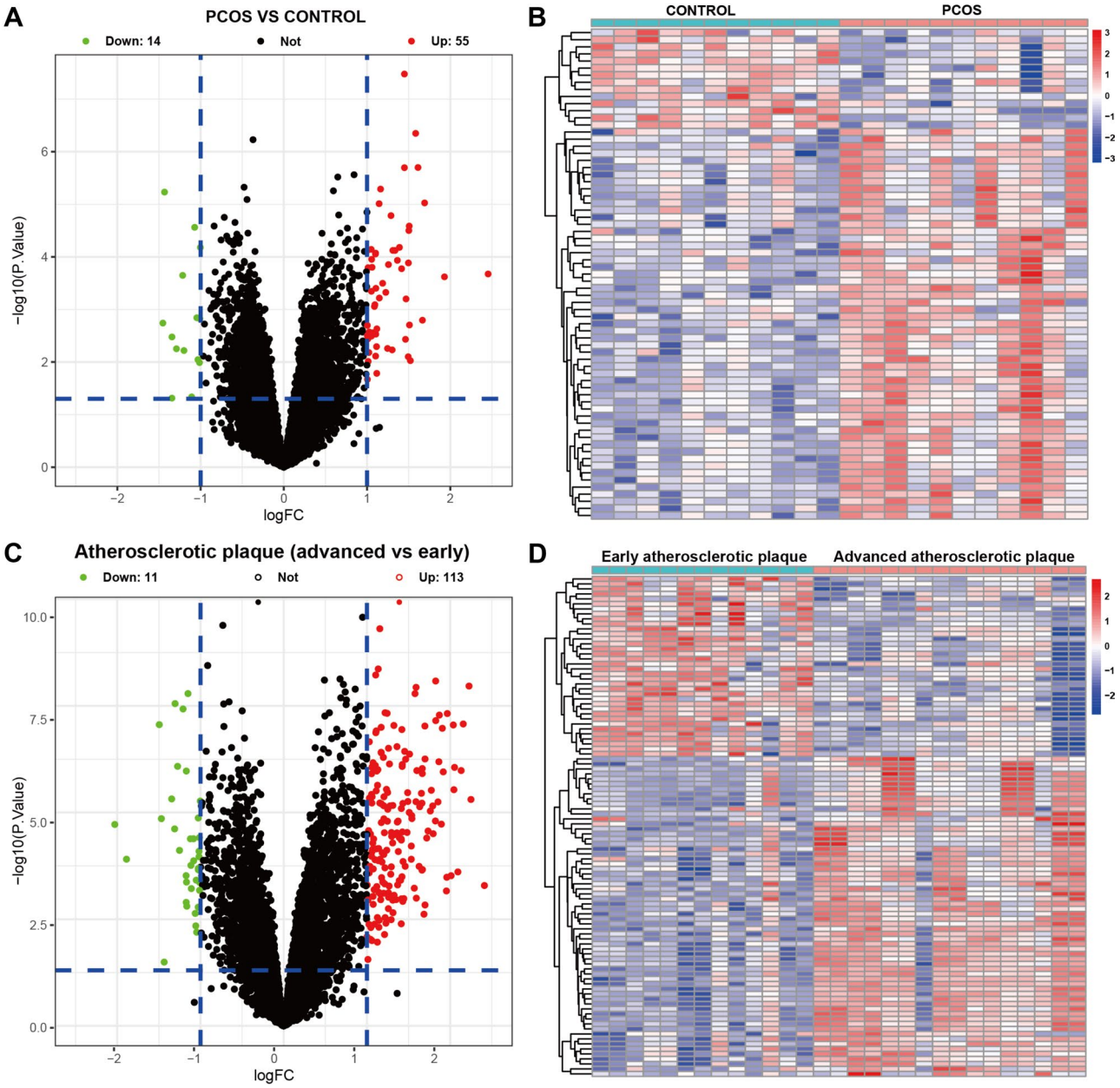

Figure 2. Identification of DEGs in PCOS and AS. (**A**,**B**) DEG heatmap and volcano plot in PCOS group. (**C**,**D**) Heatmap and the volcano plot of DEGs in AS group.

($r = 0.73$, $p = 1 \times 10^{-4}$), encompassing 242 genes (Fig. 3C). Furthermore, the AS group underwent WGCNA analysis, where the ideal soft power value ($\beta = 7$) was determined (Fig. 3D). A total of 8 modules were identified, with red ($r = 0.62$, $p = 3 \times 10^{-4}$) and blue ($r = 0.74$, $p = 5 \times 10^{-6}$) exhibiting a robust positive correlation, while modules brown and yellow displayed a significant negative correlation (Fig. 3E,F). Within AS group, we additionally chose 372 genes from the blue module that have $|MM| > 0.8$ and $|GS| > 0.5$[18].

Intersected genes and functional enrichment analysis in PCOS and AS

The intersection of DEGs and critical module genes was taken as a starting point for exploring the pathogenesis of PCOS and AS, respectively. In PCOS and AS, there was an overlap observed between the DEGs and the most crucial module (Fig. 4A,B). In order to examine the possible biological alterations in PCOS and AS, we conducted an analysis of these genes to determine their functional annotation and pathway enrichment. As expected, the GO analysis of the common genes revealed significant enrichment in cell chemotaxis, activation of T cells, and the pathway involving G protein-coupled receptors in both PCOS and AS (Fig. 4C,D). The KEGG exhibited that PCOS genes were focused on cell metabolism, while AS genes were related to T cell immune regulation. Interestingly, DO analysis discovered that PCOS was related to aortic disease and pre-eclampsia (PE), whereas AS genes were correlated to PCOS and PE (Fig. 4E,F).

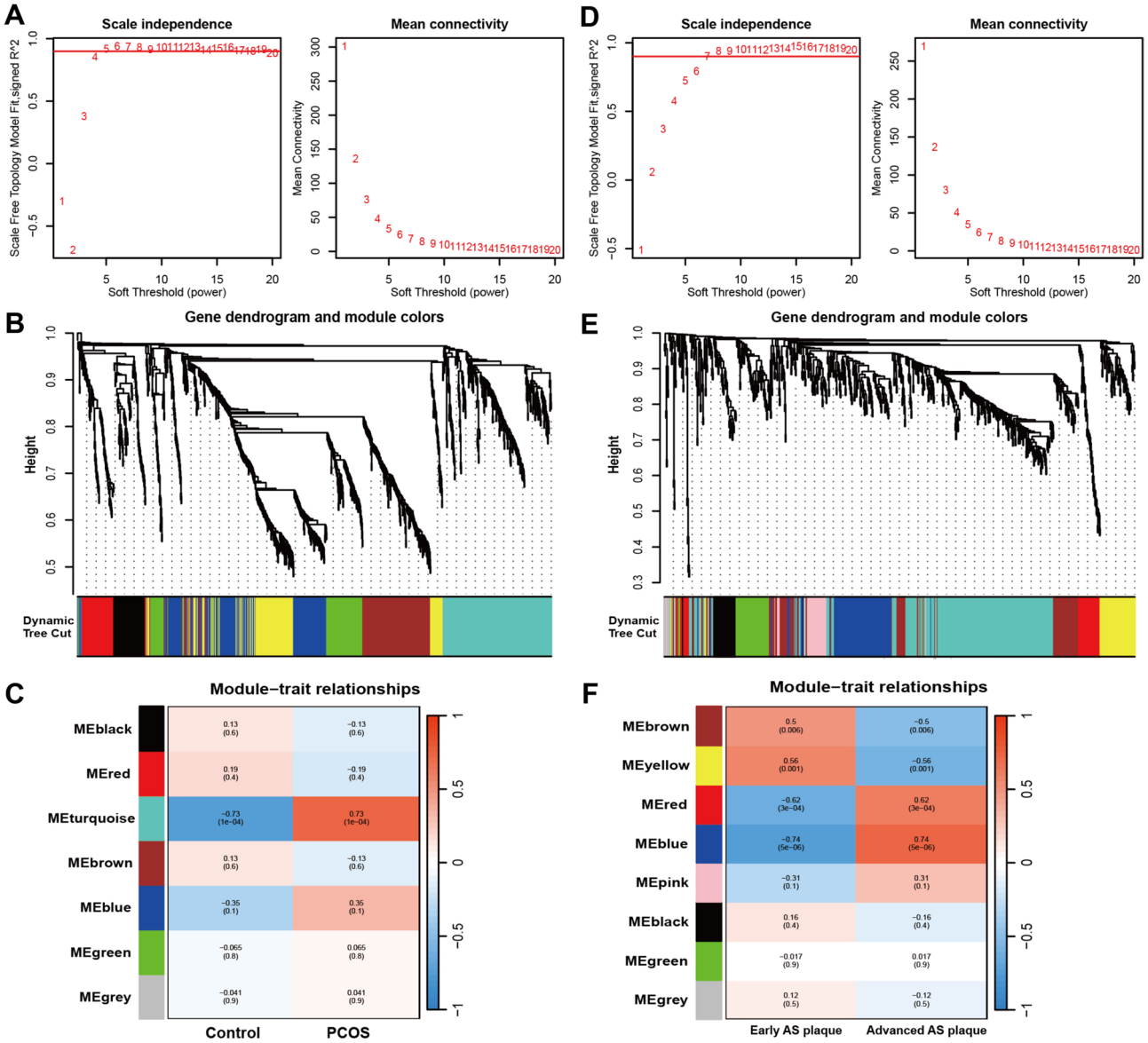

Figure 3. Analysis of PCOS and AS using weighted gene co-expression network (WGCNA). (**A**) Determination of soft-threshold power for PCOS. (**B**) Cluster dendrogram of PCOS highly connected genes in key modules. (**C**) Relationships between modules and traits in PCOS. (**D**) Calculation of soft-threshold power for AS. (**E**) Cluster dendrogram of AS modules with highly connected genes. (**F**) Module–trait relationships in AS.

Identify potential critical genes in PCOS and AS by machine learning algorithms

We intersected DEGs and key modules between PCOS and AS to obtain 17 genes, whereas CCL18 and DAPK1 were in all gene clusters (Fig. 5A). To identify additional optimal gene targets that effectively classify different groups, we utilized the algorithms of LASSO, SVM-RFE, and Random Forest to filter out optimal genes. According to LASSO coefficient profiles and an optimal tuning parameter selection map, five genes were obtained for PCOS group (Fig. 5B). After loading 17 genes to the RF classifier, we found the 6 genes with the highest significance (Fig. 5C). Afterward, the SVM-REF algorithm identified 8 genes (Fig. 5D). At the same time, similar analyses were accomplished in the AS group, and 6 genes were found in each of the 3 machine learning methods (Fig. 5E–G). Furthermore, by overlapping these three algorithms in PCOS and AS, DAPK1, and CLIC6 were the optimal genes (Fig. 5H).

Validation and single-gene GSEA of biomarkers from PCOS and AS

By utilizing the analyses of differentially expressed and machine learning between PCOS and AS, DAPK1 was recognized as a diagnostic hub biomarker. The DAPK1 mRNA expression increased dramatically in two external datasets obtained from PCOS patients ($p < 0.05$) and AS patients ($p < 0.001$) (Fig. 6A). Additionally, the diagnostic ability of DAPK1 was confirmed through the implementation of ROC curves with an AUC of 1 in PCOS. Furthermore, AS exhibited a diagnostic ability with an AUC of 0.798 (Figs. 6B). DAPK1 was then analyzed by

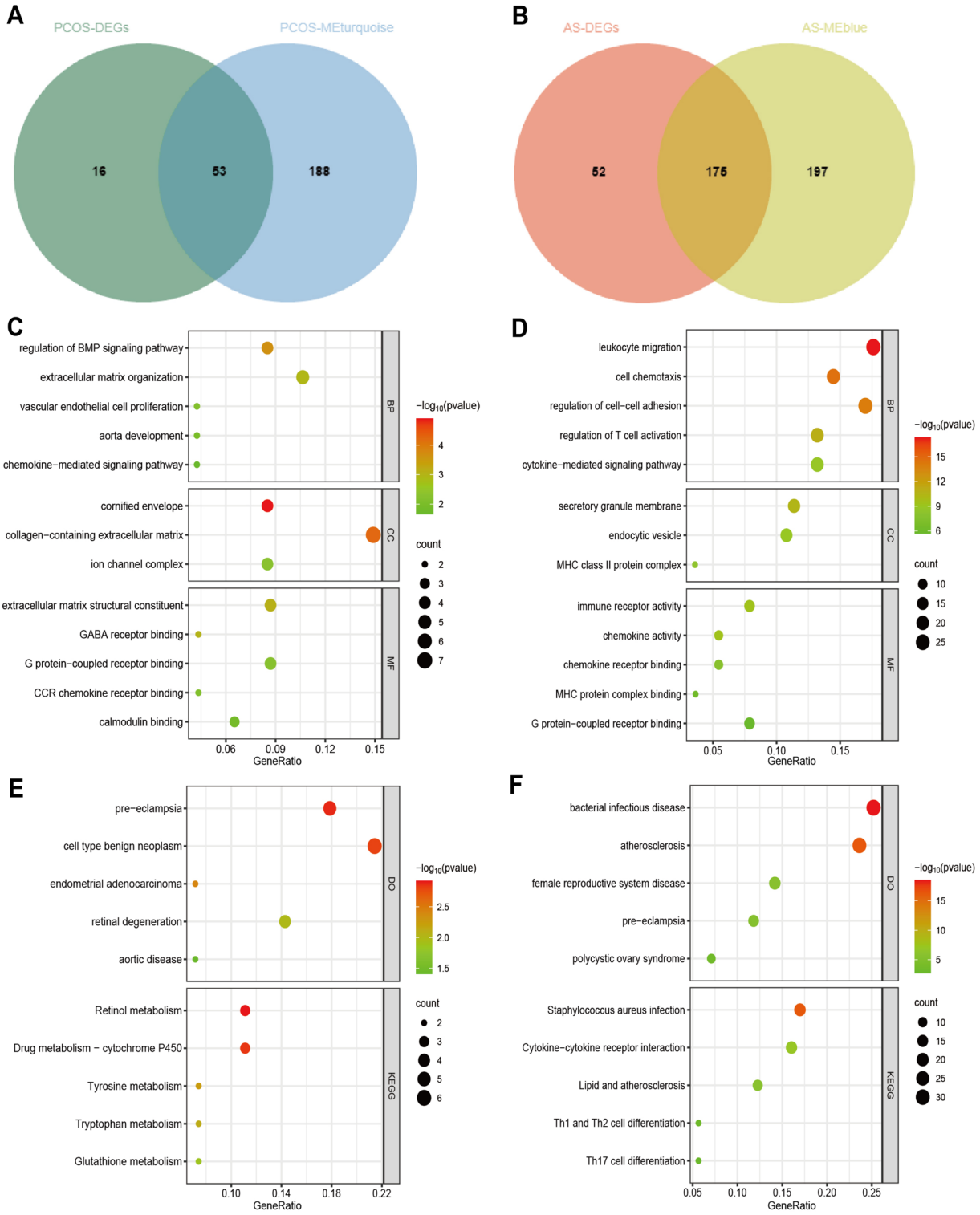

Figure 4. Shared gene signatures and functional enrichment between PCOS and AS. (**A**) The shared genes between DEGs and MEturquoise in PCOS. (**B**) The shared genes between DEGs and MEblue in AS. (**C**,**D**) Shared genes form PCOS and AS were represented by dot plots displaying GO analysis. (**E**,**F**) Shared genes form PCOS and AS were represented by dot plots displaying DO and KEGG enrichment.

single-gene GSEA in PCOS and AS groups, respectively, and 6 pathways were visualized by GSEA. In both disease groups, DAPK1 was found to be associated with inflammation and immune response pathways, including toll-like receptor signaling pathway, chemokine signaling pathway, and antigen processing and presentation (Fig. 6C,D).

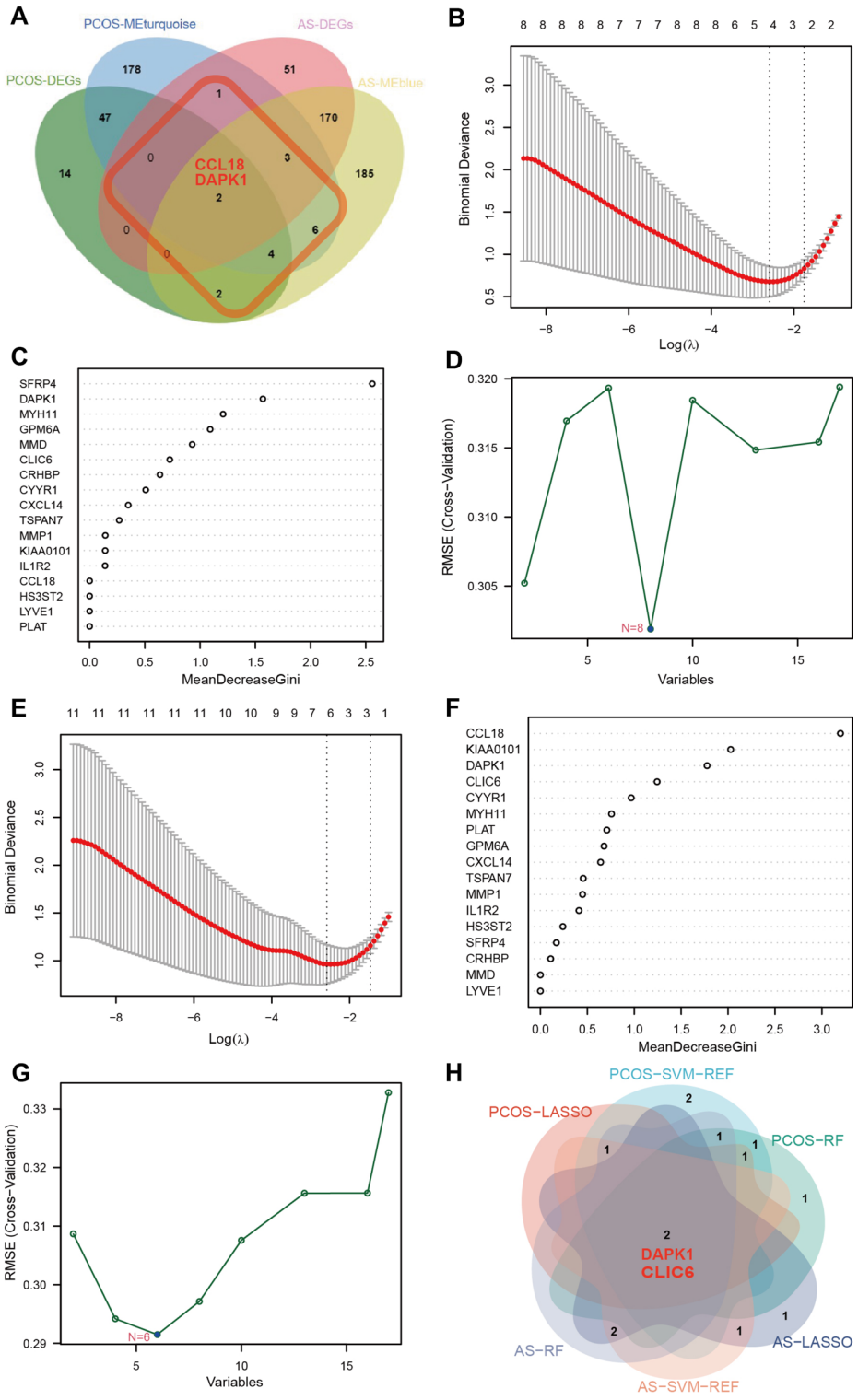

Figure 5. The screening of candidate PCOS and AS diagnostic genes using three machine learning algorithms. (**A**) 17 genes were selected between PCOS and AS for further analysis. (**B**) Coefficient profile plot of the LASSO model for PCOS showed the final parameter selection λ. (**C**) PCOS top-6 genes according to their discriminant ability in the RF algorithm. (**D**) Eight crosstalk genes were selected by using the SVM-RFE algorithm for PCOS. (**E**) Coefficient AS profile plot of the LASSO model showed the selection of the optimal parameter λ. (**F**) Top-6 RF algorithm step discriminant ability genes for AS. (**G**) Six crosstalk genes were selected by using the SVM-RFE algorithm for AS. (**H**) The Venn diagram showed two candidate diagnostic genes by intersecting the results in PCOS and AS.

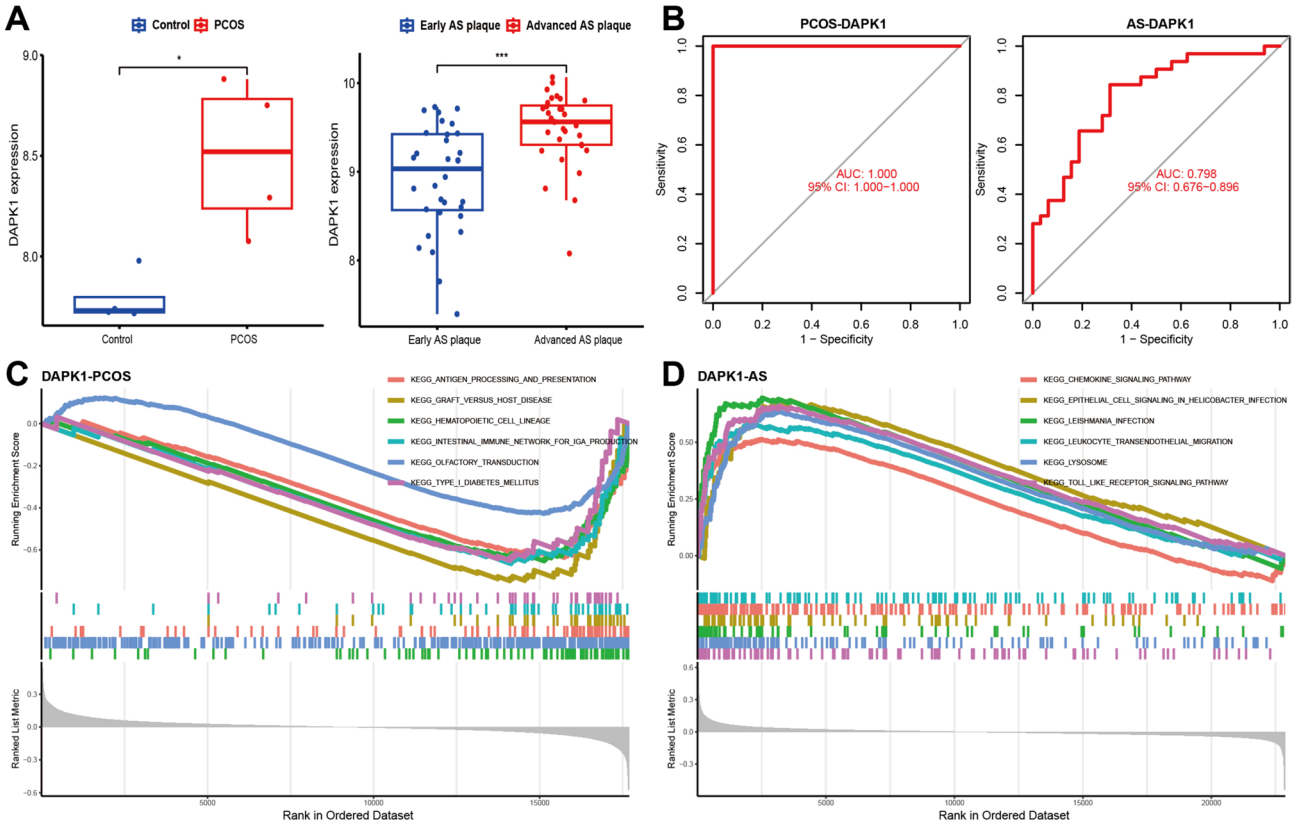

Figure 6. Validation and GSEA of the diagnostic gene. (**A**) Differential expression of DAPK1 in the validation group for PCOS and AS. (**B**) ROC curve of DAPK1 in the validation group for PCOS and AS. (**C**,**D**) GSEA analysis for DAPK1 in PCOS and AS group. *P < 0.05, ***P < 0.001.

Immune infiltration analysis of PCOS and AS

In light of the fact that PCOS and AS were marked by an excessive immune response. The distribution of immune cells among the various groups was analyzed utilizing CIBERSORT. Bar plots were drawn for each group showing the proportion of immune cells in PCOS (Fig. 7A) and AS (Fig. 7B). In comparison to the control samples, there was an elevation in neutrophils ($p = 0.023$) observed in PCOS group (Fig. 7C). Conversely, the AS group showed an elevation in gamma delta T cells ($p = 0.044$), M0 Macrophages ($p = 0.005$), and M2 Macrophages ($p = 0.028$), while experiencing a decrease in monocytes and activated dendritic cells (Fig. 7D). Furthermore, an examination was conducted on the correlation of biomarkers with the composition of immune cells. The resting dendritic cells ($p = 0.035$) were negatively related to DAPK1 in PCOS group (Fig. 7E). In the AS samples, DAPK1 exhibited a notable favorable association with gamma delta T cells ($p = 0.003$), M0 Macrophages ($p < 0.001$), and M2 Macrophages ($p < 0.001$). However, activated mast cells ($p = 0.031$), regulatory T cells ($p = 0.013$), monocytes ($p = 0.013$), and activated dendritic cells ($p < 0.001$) showed an adverse correlation with DAPK1 (Fig. 7F). In order for PCOS and AS to develop, immune function appears to be crucial.

Validation of DAPK1 by RT-PCR

RT-PCR was conducted on granulosa cells obtained from healthy females and individuals with PCOS as well as on ox-LDL stimulated RAW264.7 cells. As a result, the expression levels of DAPK1 were confirmed. In line with the data analysis, our findings indicated that the expression of DAPK1 was increased in the granulosa cells ($p < 0.01$) of individuals with PCOS and RAW264.7 cells ($p < 0.05$) stimulated by ox-LDL (Fig. 8A,B).

Discussion

Assessment for hirsutism, menstrual irregularity, and infertility is often the primary focus of initial clinical management in women with PCOS[30]. Nevertheless, they also commonly experience insulin resistance, obesity, dyslipidemia, metabolic syndrome, and hypertension, which increases their susceptibility to Cardiovascular disease[31]. Considering that PCOS is marked by excessive androgen levels, it is evident from various researches that pre- and postmenopausal women with PCOS exhibit a higher occurrence of coronary artery calcium (CAC) and notably elevated CAC score[4]. These findings persist even after accounting for age and Body Mass Index[32]. In comparison to controls, women suffering from PCOS exhibit higher carotid intimal medial thickness and a remarkerable increased occurrence of carotid plaque[33]. A different research analyzed the blood flow mediated dilation in brachial artery and discovered that women with PCOS had notably reduced levels compared to controls of the same age, suggesting a higher degree of endothelial dysfunction[34]. The precise connection between PCOS and AS remains incompletely comprehended thus far. Hence, investigating the molecular mechanisms

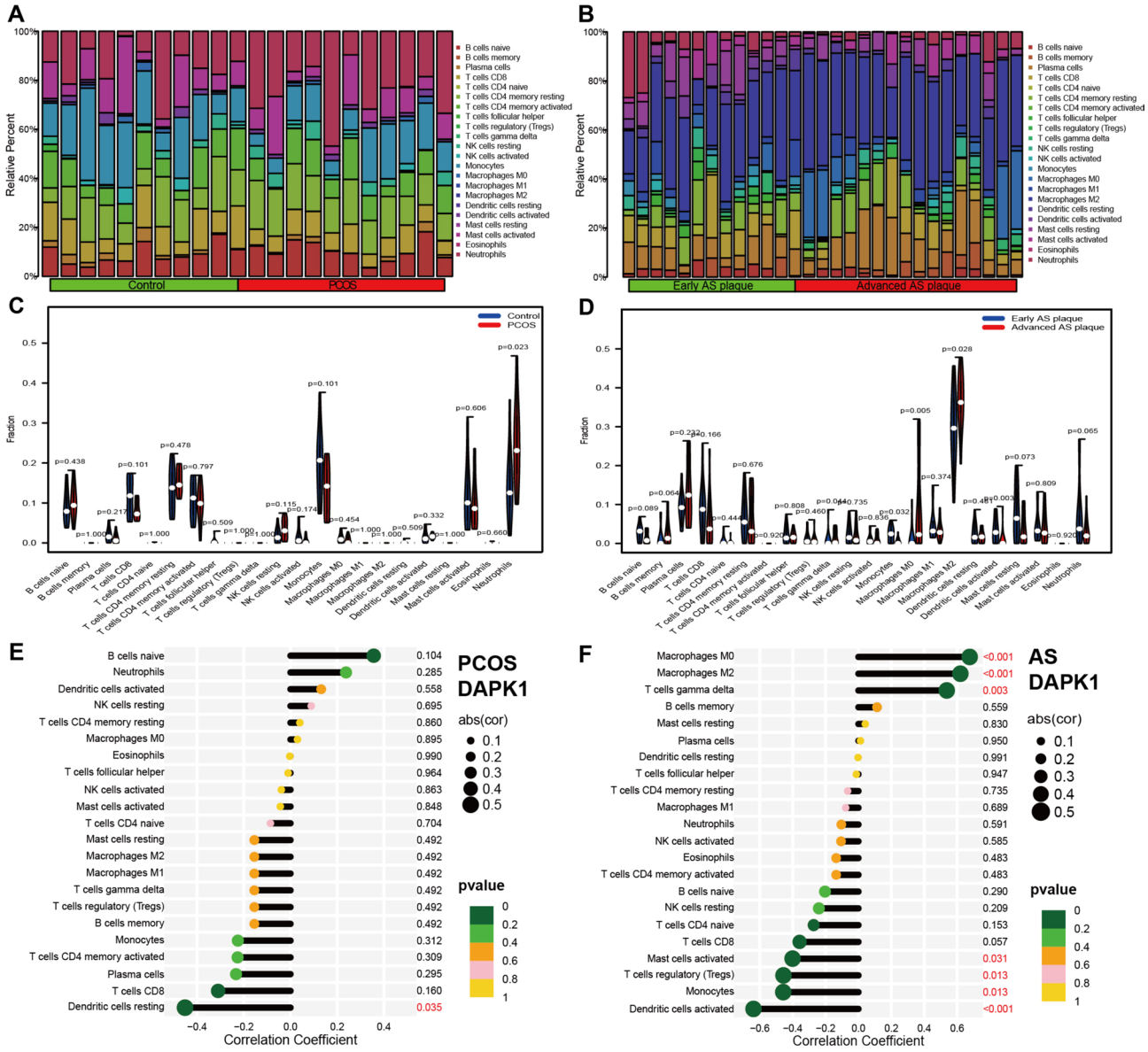

Figure 7. PCOS and AS immune cell composition. (**A**,**B**) Infiltrating immune cells were plotted in a stacked bar chart for the PCOS and AS group. (**C**,**D**) Violin diagram indicated that the PCOS and AS group exhibited a significantly different type of immune cell. (**E**) Correlation between DAPK1 expression and immune cells in the PCOS group. (**F**) Relationship between the expression of DAPK1 and immunity in AS patients.

connecting PCOS and AS, along with timely detection and intervention, unquestionably hold crucial clinical importance.

This study conducted various bioinformatics analyses on separate gene chip databases for PCOS and AS. The analysis of GO enrichment revealed that differentially expressed genes (DEGs) were primarily enriched in cellular chemotaxis, pathways involving cytokine-mediated signaling, and pathways involving G protein-coupled receptors (GPCRs). In one hand, Feinstein et al. discovered that GPCRs such as chemokine receptors (CCR2 and CXCR4) are involved in the recruitment of immune cells to the site of inflammation in the vascular endothelium, exacerbating the atherosclerotic process[35]. In the other hand, the luteinizing hormone receptor, a GPCR, had been shown to have altered signaling in PCOS patients, contributing to excessive androgen production and anovulation, which are hallmark features of the PCOS[36]. Meanwhile, GPCRs such as GPR40 and GPR120, which were involved in fatty acid signaling, affected insulin secretion and sensitivity, exacerbated insulin resistance[37]. Evidently, the analysis of differential expression revealed that the DEGs associated with AS were predominantly enriched in PCOS, preeclampsia, and diseases related to the female reproductive system. This strongly indicates a significant connection between PCOS and AS in terms of their development. Meanwhile, KEGG and GSEA analyses showed that leukocyte migration and cytokine-mediated signaling pathway mainly occurred in AS whereas antigen presentation and intestinal immunity were the main manifestations in PCOS. Chattopadhayay and colleagues found that individuals with PCOS exhibit heightened levels of oxidative stress,

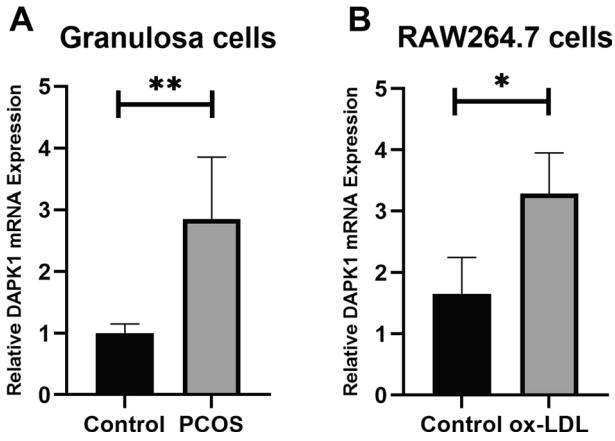

Figure 8. Validation of DAPK1 in granulosa cells and RAW264.7 cells. (**A**) Expression levels of DAPK1 in granulosa cells of normal and PCOS patients (n = 5). (**B**) Expression levels of DAPK1 in ox-LDL stimulated RAW264.7 cells. (n = 4) *P < 0.05, **P < 0.01.

which is intricately linked to inflammation. Patients with PCOS display increased levels of oxidative stress indicators, such as malondialdehyde, and diminished antioxidant capacity. This oxidative stress exacerbates inflammatory processes, resulting in elevated levels of inflammatory markers, including C-reactive protein, IL-6, and TNF-α, as well as heightened insulin resistance, thereby establishing a detrimental feedback loop[38]. Moreover, Libby reported that inflammatory cytokines play crucial roles in atherosclerosis. IL-6, produced by various cells within the plaque, promotes further recruitment of immune cells and enhances the inflammatory response. TNF-α contributes to endothelial dysfunction, increases vascular permeability, and stimulates the expression of other pro-inflammatory mediators. Additionally, chemokines such as monocyte chemoattractant protein-1 direct the migration of monocytes to sites of inflammation[39]. Thus, both diseases were primarily related to the immune response and inflammation, which contributed to our discovery of new diagnostic genes and study of the underlying molecular mechanisms.

In this research, we utilized WGCNA and machine learning techniques to discover and confirm DAPK1 as a diagnostic gene that is strongly linked to polycystic ovary syndrome and AS, indicating an unfavorable prognosis. DAPK1 is a member of a group of five Ser/Thr kinases, and it stands as the most extensive protein in the DAPK family, comprising 1,430 amino acids. The presence of the Ca^{2+}/CaM autoregulatory domain in this component leads to the regulation of apoptosis, cell motility, and autophagy via phosphorylation and ERK signaling pathways[40]. Zheng et al. indicated a notable decrease in miR-141-3p levels within the rat model of PCOS, leading to the promotion of apoptosis in ovarian granulosa cells through the targeting of DAPK1[41]. Gauster et al. found that placental tissues of early-onset pre-eclampsia showed low expression of DAPK1, and increased expression and lipolabelling of LC3B, suggesting that inflammation and oxidative stress as potential stimuli for altered placental autophagy[42]. In addition, Duan et al. discovered that DAPK1 hinders autophagy to alleviate inflammatory reactions in mice with LPS-induced acute lung injury (ALI). Thus, DAPK1 regulates PCOS and inflammatory disease processes by mediating apoptosis or autophagy.

Tian et al. reported that Gpr137b-ps regulates amino acid-mTORC1-autophagy signalling by impairing the interaction of HSC70 with G3BP, leading to defective macrophage autophagy, plaque necrotic core enlargement and lipid accumulation in patients with advanced AS[43]. Deficient autophagy in T cells without autophagy-related protein 7 hampers the development of fatty liver and the advancement of AS by enhancing the secretion of interferon-γ and interleukin–17 from CD4+ and CD8+ T cells[44]. Our immune infiltration analysis revealed a very strong association of DAPK1 with immune cells. DAPK1 showed a strong positive correlation with gamma delta T cells, M0 macrophages, and M2 macrophages in AS samples. Conversely, DAPK1 exhibited a negative association with activated mast cells, regulatory T cells, monocytes, and activated dendritic cells. Therefore, although there are no relevant studies of DAPK1 in AS, we speculate that DAPK1 may regulate the progression of AS by mediating macrophage or T cell autophagy.

The present study requires mentioning certain limitations. Initially, the relatively small sample size of selected datasets could affect the generalizability and statistical power of the findings. Future studies should aim to include larger sample sizes to enhance statistical power and generalizability. Furthermore, processing raw data is limited in our understanding, and these datasets were analyzed utilizing a microarray technology, which is significantly less advanced compared to RNA sequencing techniques, and limits the depth and accuracy of the gene expression analysis. Finally, the study also lacks experimental validation of the identified biomarkers, such as functional assays to confirm the role of DAPK1 in PCOS and AS.

Conclusion

The diagnostic biomarker DAPK1 has been identified as a crucial factor in the regulation of immune responses mediated by T cells in PCOS and AS. Our results provide molecular targets for subsequent in vitro experiments for mechanistic studies in PCOS model with AS. Our research may guide clinicians to incorporate longitudinal data to gain insight into the temporal dynamics of DAPK1 expression and its role in disease progression.

Data availability

The public datasets: GSE137684 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE137684), GSE114419 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114419), GSE106724 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106724), GSE54250 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54250), GSE28829 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28829) and GSE43292 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4329) supporting the conclusions of this article are available in the GEO database (http://www.ncbi.nlm.nih.gov/geo).

Received: 16 February 2024; Accepted: 31 July 2024
Published online: 06 August 2024

References

1. Delitala, A. P., Capobianco, G., Delitala, G., Cherchi, P. L. & Dessole, S. Polycystic ovary syndrome, adipose tissue and metabolic syndrome. *Arch. Gynecol. Obstet.* **296**(3), 405–419 (2017).
2. Stener-Victorin, E. *et al.* Polycystic ovary syndrome. *Nat. Rev. Dis. Primers.* **10**(1), 27 (2024).
3. Lizneva, D. *et al.* Criteria, prevalence, and phenotypes of polycystic ovary syndrome. *Fertil. Steril.* **106**(1), 6–15 (2016).
4. Osibogun, O., Ogunmoroti, O. & Michos, E. D. Polycystic ovary syndrome and cardiometabolic risk: Opportunities for cardiovascular disease prevention. *Trends Cardiovasc. Med.* **30**(7), 399–404 (2020).
5. Hansson, G. K. & Hermansson, A. The immune system in atherosclerosis. *Nat. Immunol.* **12**(3), 204–212 (2011).
6. Libby, P. The changing landscape of atherosclerosis. *Nature.* **592**(7855), 524–533 (2021).
7. Jabbour, R., Ott, J., Eppel, W. & Frigo, P. Carotid intima-media thickness in polycystic ovary syndrome and its association with hormone and lipid profiles. *PLoS ONE.* **15**(4), e0232299 (2020).
8. Lorenz, M. W., Schaefer, C., Steinmetz, H. & Sitzer, M. Is carotid intima media thickness useful for individual prediction of cardiovascular risk? Ten-year results from the Carotid Atherosclerosis Progression Study (CAPS). *Eur. Heart J.* **31**(16), 2041–2048 (2010).
9. Chen, X. *et al.* Sodium-glucose cotransporter 2 inhibitor ameliorates high fat diet-induced hypothalamic-pituitary-ovarian axis disorders. *J. Physiol.* **600**(21), 4549–4568 (2022).
10. Gao, L. *et al.* Polycystic ovary syndrome fuels cardiovascular inflammation and aggravates ischemic cardiac injury. *Circulation.* **148**, 1958–1973 (2023).
11. Huang, Z. H. *et al.* PCOS is associated with increased CD11c expression and crown-like structures in adipose tissue and increased central abdominal fat depots independent of obesity. *J. Clin. Endocrinol. Metab.* **98**(1), E17-24 (2013).
12. Borthakur, A., Prabhu, Y. D. & Valsala Gopalakrishnan, A. Role of IL-6 signalling in Polycystic Ovarian Syndrome associated inflammation. *J. Reprod. Immunol.* **141**, 103155 (2020).
13. Everett, B. M. *et al.* Rationale and design of the Cardiovascular Inflammation Reduction Trial: A test of the inflammatory hypothesis of atherothrombosis. *Am. Heart J.* **166**(2), 199-207.e15 (2013).
14. Fulghesu, A. M. *et al.* IL-6 serum levels and production is related to an altered immune response in polycystic ovary syndrome girls with insulin resistance. *Mediators Inflamm.* **2011**, 389317 (2011).
15. Chen, W. *et al.* Shared diagnostic genes and potential mechanism between PCOS and recurrent implantation failure revealed by integrated transcriptomic analysis and machine learning. *Front. Immunol.* **14**, 1175384 (2023).
16. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**(7), e47 (2015).
17. Luo, Y. & Zhou, Y. Identification of novel biomarkers and immune infiltration features of recurrent pregnancy loss by machine learning. *Sci. Rep.* **13**(1), 10751 (2023).
18. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics.* **9**, 559 (2008).
19. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2**(3), 100141 (2021).
20. Yang, C., Delcher, C., Shenkman, E. & Ranka, S. Machine learning approaches for predicting high cost high need patient expenditures in health care. *Biomed. Eng. Online.* **17**(Suppl 1), 131 (2018).
21. Ellis, K. *et al.* A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiol. Meas.* **35**(11), 2191–2203 (2014).
22. Mi, X., Zou, B., Zou, F. & Hu, J. Permutation-based identification of important biomarkers for complex diseases via machine learning models. *Nat. Commun.* **12**(1), 3008 (2021).
23. Harbron, C., Chang, K. M. & South, M. C. RefPlus: An R package extending the RMA algorithm. *Bioinformatics.* **23**(18), 2493–2494 (2007).
24. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.* **14**(2), 482–517 (2019).
25. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* **27**(12), 1739–1740 (2011).
26. Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M. & Alizadeh, A. A. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* **1711**, 243–259 (2018).
27. Ito, K. & Murphy, D. Application of ggplot2 to pharmacometric graphics. *CPT Pharmacometrics Syst. Pharmacol.* **2**(10), e79 (2013).
28. Teede, H. J. *et al.* Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *J. Clin. Endocrinol. Metab.* **108**(10), 2447–2469 (2023).
29. Mirza, Z. *et al.* Atheroprotective effect of fucoidan in THP-1 macrophages by potential upregulation of ABCA1. *Biomedicines.* **11**(11), 2929 (2023).
30. Sirmans, S. M. & Pate, K. A. Epidemiology, diagnosis, and management of polycystic ovary syndrome. *Clin. Epidemiol.* **6**, 1–13 (2013).
31. de Groot, P. C., Dekkers, O. M., Romijn, J. A., Dieben, S. W. & Helmerhorst, F. M. PCOS, coronary heart disease, stroke and the influence of obesity: A systematic review and meta-analysis. *Hum. Reprod. Update.* **17**(4), 495–500 (2011).
32. Calderon-Margalit, R. *et al.* Prospective association of polycystic ovary syndrome with coronary artery calcification and carotid-intima-media thickness: The Coronary Artery Risk Development in Young Adults Women's study. *Arterioscler. Thromb. Vasc. Biol.* **34**(12), 2688–2694 (2014).
33. Meyer, M. L., Malek, A. M., Wild, R. A., Korytkowski, M. T. & Talbott, E. O. Carotid artery intima-media thickness in polycystic ovary syndrome: A systematic review and meta-analysis. *Hum. Reprod. Update.* **18**(2), 112–126 (2012).
34. Sathyapalan, T., Javed, Z., Kilpatrick, E. S., Coady, A. M. & Atkin, S. L. Endocannabinoid receptor blockade increases vascular endothelial growth factor and inflammatory markers in obese women with polycystic ovary syndrome. *Clin. Endocrinol.* **86**(3), 384–387 (2017).
35. Slysz, J. *et al.* Single-cell profiling reveals inflammatory polarization of human carotid versus femoral plaque leukocytes. *JCI Insight.* https://doi.org/10.1172/jci.insight.171359 (2023).

36. Schniewind, H. A. *et al.* Autoimmunity to the follicle-stimulating hormone receptor (FSHR) and luteinizing hormone receptor (LHR) in polycystic ovarian syndrome. *Int. J. Mol. Sci.* **22**(24), 13667 (2021).
37. Lu, Z. *et al.* GPR40 deficiency is associated with hepatic FAT/CD36 upregulation, steatosis, inflammation, and cell injury in C57BL/6 mice. *Am. J. Physiol. Endocrinol. Metab.* **320**(1), E30–E42 (2021).
38. Chattopadhayay, R. *et al.* Effect of follicular fluid oxidative stress on meiotic spindle formation in infertile women with polycystic ovarian syndrome. *Gynecol. Obstet. Investig.* **69**(3), 197–202 (2010).
39. Libby, P. Inflammation in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **32**(9), 2045–2051 (2012).
40. Nair, S., Hagberg, H., Krishnamurthy, R., Thornton, C. & Mallard, C. Death associated protein kinases: Molecular structure and brain injury. *Int. J. Mol. Sci.* **14**(7), 13858–13872 (2013).
41. Li, D. *et al.* MicroRNA-141-3p targets DAPK1 and inhibits apoptosis in rat ovarian granulosa cells. *Cell Biochem. Funct.* **35**(4), 197–201 (2017).
42. Prokesch, A. *et al.* Placental DAPK1 and autophagy marker LC3B-II are dysregulated by TNF-alpha in a gestational age-dependent manner. *Histochem. Cell Biol.* **147**(6), 695–705 (2017).
43. Qu, W. *et al.* Long noncoding RNA Gpr137b-ps promotes advanced atherosclerosis via the regulation of autophagy in macrophages. *Arterioscler. Thromb. Vasc. Biol.* **43**(11), e468–e489 (2023).
44. Amersfoort, J. *et al.* Defective autophagy in T cells impairs the development of diet-induced hepatic steatosis and atherosclerosis. *Front. Immunol.* **9**, 2937 (2018).

Author contributions

Conceptualization, Z.J. and Y.L.; Formal analysis, Y.Z. and X.Z.; Methodology, Q.Z. and Q.L.; Resources, Y.Z. and R.G.; Data curation, B.Y., and L.W.; Writing—original draft, Y.Z. and Y.L.; Writing—review and editing, Z.J. and H.J. All authors have reviewed and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Z.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.