

RESEARCH

Open Access



Identification of immune cell infiltration and effective biomarkers of polycystic ovary syndrome by bioinformatics analysis

Mengge Gao^{1,2,3†}, Xiaohua Liu^{1†}, Mengxuan Du^{1,2}, Heng Gu¹, Hang Xu^{1,2} and Xingming Zhong^{1,2*}

Abstract

Background Patients with polycystic ovary syndrome (PCOS) exhibit a chronic inflammatory state, which is often accompanied by immune, endocrine, and metabolic disorders. Clarification of the pathogenesis of PCOS and exploration of specific biomarkers from the perspective of immunology by evaluating the local infiltration of immune cells in the follicular microenvironment may provide critical insights into disease pathogenesis.

Methods In this study, we evaluated immune cell subsets and gene expression in patients with PCOS using data from the Gene Expression Omnibus database and single-sample gene set enrichment analysis.

Results In total, 325 differentially expressed genes were identified, among which *TMEM54* and *PLCG2* (area under the curve = 0.922) were identified as PCOS biomarkers. Immune cell infiltration analysis showed that central memory CD4⁺ T cells, central memory CD8⁺ T cells, effector memory CD4⁺ T cells, $\gamma\delta$ T cells, and type 17 T helper cells may affect the occurrence of PCOS. In addition, *PLCG2* was highly correlated with $\gamma\delta$ T cells and central memory CD4⁺ T cells.

Conclusions Overall, *TMEM54* and *PLCG2* were identified as potential PCOS biomarkers by bioinformatics analysis. These findings established a basis for further exploration of the immunological mechanisms of PCOS and the identification of therapeutic targets.

Keywords Polycystic ovary syndrome, Immune cell infiltration, Biomarker, Single-sample gene set enrichment analysis

Introduction

Polycystic ovary syndrome (PCOS) is a common reproductive endocrine metabolic disease in women of childbearing age and is often characterised by chronic anovulation and hyperandrogenaemia. Its clinical manifestations include menstrual disorders, hirsutism, acne, and polycystic ovarian changes, and these symptoms can be accompanied by metabolic diseases, such as obesity, insulin resistance (IR), and dyslipidaemia [1]. However, the pathogenic mechanisms of PCOS are still unclear. In addition to the ovarian-pituitary-hypothalamic-gonadal axis, studies of PCOS pathogenesis must also consider ovarian local cytokines, immunology, and genetics.

[†]Mengge Gao and Xiaohua Liu contributed equally to this work.

*Correspondence:

Xingming Zhong
xingmingzh@126.com

¹NHC Key Laboratory of Male Reproduction and Genetics, Guangdong Provincial Reproductive Science Institute (Guangdong Provincial Fertility Hospital), Guangzhou 510600, China

²Department of Public Health and Preventive Medicine, School of Medicine, Jinan University, Guangzhou, Guangdong Province, China

³Department of Clinical Nutrition, Huadu District People's Hospital, Southern Medical University, 48 Xinhua Road, Huadu, Guangzhou 510800, Guangdong, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Recent studies have found that the inflammatory immune mechanism is closely related to the occurrence and development of PCOS. Under physiological conditions, appropriate inflammatory stress is conducive to the growth and development of oocytes [2, 3]. However, under pathological conditions, the inflammatory response is enhanced, the development of oocytes may be limited, and the further development of chronic inflammation leads to decreased ovum quality, thereby affecting ovulation [4]. Systemic and ovarian cytokines (e.g., tumour necrosis factor [TNF]- α , interleukin [IL]-6, and IL-18) can change the local microenvironment in the ovary, regulate ovarian function, induce excessive androgen production, and promote IR through various mechanisms [5]. Inflammation in the follicular microenvironment may be involved in the dysfunction of the hypothalamic-pituitary-gonad axis and the occurrence and development of follicular dysplasia.

In studies of the distribution of white blood cells in the ovary and cytokine mRNA expression in the follicular fluid (FF) of patients with PCOS and non-PCOS women undergoing fertilisation embryo transfer in vitro, Wu et al. [6] showed that T lymphocytes play important roles in the local pathological mechanisms of PCOS. T lymphocytes secrete various inflammatory and immunomodulatory molecules that participate in the regulation of ovarian function. T-cell subsets have also been shown to be dysregulated in the peripheral blood and ovaries of patients with PCOS owing to disruption of sex hormone levels in these patients [7].

The pathogenesis of PCOS is multifactorial and complex. In addition to reproductive abnormalities, the pathogenic mechanisms also include interactions between the immune system and reproduction, resulting in a variety of changes to cytokines and immune cells. Therefore, evaluation of immune cell infiltration in patients with PCOS based on changes in the expression levels of genes may be essential for elucidating the immunological mechanisms of PCOS and identification of novel biomarkers.

In this study, we aimed to identify the roles of immune cell subsets and related gene expression changes in the pathogenesis of PCOS. We downloaded four PCOS datasets from the Gene Expression Omnibus (GEO) database and analysed differentially expressed genes. Two different machine learning algorithms were then used to further identify PCOS biomarkers. We also studied PCOS from the perspective of immunology using single-sample gene set enrichment analysis (ssGSEA) to evaluate the differences in the compositions of 28 immune cell subsets between patients with PCOS and healthy women of reproductive age. In addition, the relationships between PCOS biomarkers and immune cell infiltration were

studied to improve our understanding of the immunological mechanisms of PCOS occurrence and development.

Materials and methods

Data downloading

First, the “*GEOquery*” package in R software (version 3.6.1, <http://r-project.org/>) was used to download the PCOS expression profile datasets GSE84958, GSE106724, GSE137684, GSE114419 and GSE193812 from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). A description of all databases is presented in Supplementary Table S1.

Data preprocessing

We combined the GSE84958, GSE106724, GSE137684, and GSE114419 gene expression matrices, and interbatch differences were removed using the “*sva*” package [8]. Boxplots [9] and two-dimensional principal component analysis (PCA) cluster plots were used to visualise the effects of removing interbatch differences. The flow chart for this study is shown in Supplementary Figure S1.

Evaluation of the distribution of immune cell subtypes

ssGSEA was used to quantify the infiltration of immune cells in patients with PCOS. The abundances of the following 28 types of immune cells were obtained from the R package “*GSVA*” [10]: activated B cells, activated CD4⁺ T cells, activated CD8⁺ T cells, activated dendritic cells (DCs), CD56⁺ natural killer (NK) cells, CD56⁻ NK cells, central memory CD4⁺ T cells, central memory CD8⁺ T cells, effector memory CD4⁺ T cells, effector memory CD8⁺ T cells, eosinophils, $\gamma\delta$ T cells, immature B cells, immature DCs, mast cells, myeloid inhibitory cells, memory B cells, monocytes, NK cells, NK T cells, neutrophils, plasmacytoid DCs, macrophages, regulatory T cells (Tregs), follicular helper T cells, type 1 T helper (Th1) cells, type 17 T helper (Th17) cells, and type 2 T helper (Th2) cells.

Screening of differentially expressed genes (DEGs)

DEGs were filtered through the “*limma*” package [11]. The DEG threshold point was an adjusted *P* value less than 0.05 and a $|\log_2$ fold change| greater than 0.5.

Functional enrichment analysis

The gene names of DEGs were converted to gene IDs using the R package “*org.hs.eg.db*”. Gene ontology (GO), Kyoto Encyclopaedia of Genes and Genomes [12, 13] (KEGG), and disease ontology (DO) analyses were performed using the R package “*clusterProfiler*” [14], and *p* value correction from multiple trials was performed using the BH method. GO, DO, and KEGG pathways with significant enrichment of DEGs were screened. GO annotations of DEGs included cellular components,

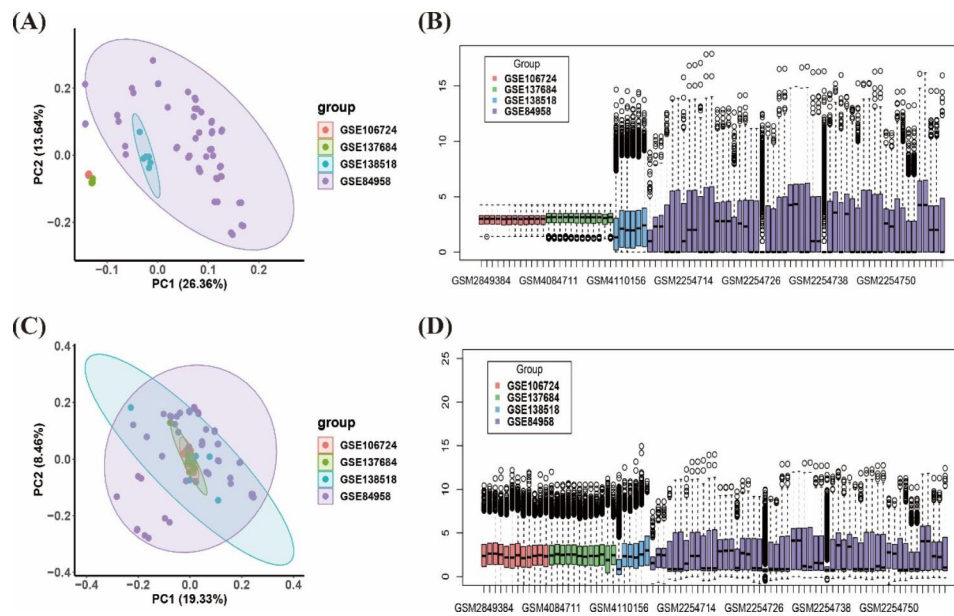


Fig. 1 Two-dimensional PCA and boxplot before and after removing interbatch effects. (A, B) Data before removing interbatch effects in datasets GSE84958, GSE106724, GSE114419, and GSE137684. (C, D) Data after removing interbatch effects

biological processes, and molecular functions, which were used to analyse the functional enrichment of DEGs.

Screening of biomarkers

Protein-protein interaction (PPI) network results for DEGs were obtained using STRING, and LASSO regression and BORUTA algorithm analyses were used for feature selection to screen biomarkers of PCOS. The LASSO algorithm was implemented through the R package “glmnet” [15]. The BORUTA algorithm was implemented through the R package “Boruta” [16]. Similar to a random forest classifier, this approach reduced the error caused by random fluctuation and correlation by adding randomness to the system and collecting results from a random sample set. The candidate biomarkers obtained using the two methods were intersected using a Venn diagram to extract the final candidate biomarkers.

Correlation analysis between biomarkers and immune cells

Spearman rank correlation analysis was used to investigate the correlations between the selected biomarkers and the level of immune cell infiltration.

Analysis of the predictive value of biomarkers

In order to determine the predictive value of biomarkers for PCOS, receiver operating characteristic (ROC) curve analysis was carried out to explore the sensitivity and specificity of candidate biomarkers for PCOS prediction. Screening of the ROC area under the curve (AUC) values identified the dataset that could distinguish PCOS from

normal control samples. An external data set and some peripheral blood samples were then used for testing. In this study, 6 PCOS patients who visited Guangdong Provincial Fertility Hospital from March 2022 to September 2022 were included and 6 women of reproductive age who visited the hospital and passed physical examination during the same period were included as the control group. Using peripheral blood RNA as template, cDNA was synthesized using Evo M-MLV reverse transcription reagent premix configuration system. Three groups of replicates were set for each sample, $2^{-\Delta\Delta CT}$ formula was used to calculate the relative change of each gene expression in each sample.

Statistical analysis and graphical visualisation

All analyses in this paper were carried out by R version 4.0.5. The analysis results were visualised using the R packages “corrplot” [17], “ggplot2” [18], and “pheatmap” [19] and spliced with “patchwork”.

Results

Data preprocessing

In this study, we selected 83 samples from four different datasets, including 34 PCOS samples and 49 normal control samples. A two-dimensional PCA diagram and boxplot showed the effects before and after treatment (Fig. 1A–D). The results demonstrated that clustering of the two groups of samples was more obvious after normalisation, indicating that differences between batches had been eliminated and that the samples were from reliable sources.

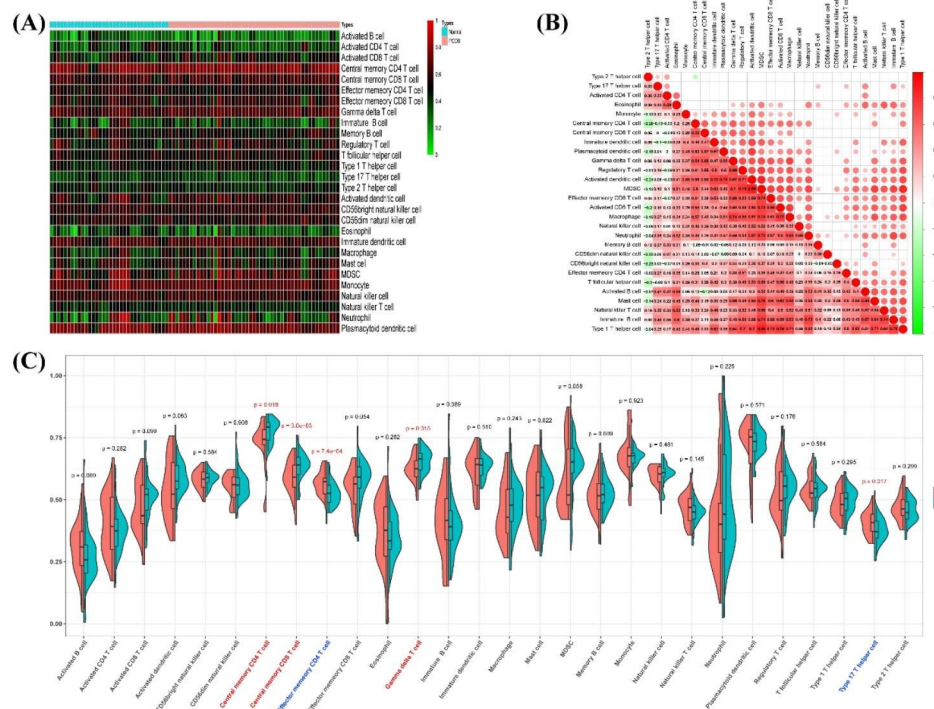


Fig. 2 Differences in immune cell infiltration between patients with PCOS and normal controls. **(A)** Differences in the enrichment of immune cells between samples from the two groups. The normalised relative abundances of immune cells in each sample are shown, with red representing high enrichment degree and green representing low enrichment degree. **(B)** Correlations of 28 types of immune cells in the dataset. Colour blocks with a correlation coefficient p value greater than 0.05 in the upper part of the graph are not displayed; red represents a positive correlation, and green represents a negative correlation, with darker colour indicating a stronger correlation. **(C)** Differences in the enrichment of 28 types of immune cells between patients with PCOS and normal controls. Results with p values less than 0.05 are indicated in red, and immune cells with different enrichment degrees on the axis are indicated as red or blue. Red indicates a significant increase in the PCOS group, whereas blue indicates a significant decrease in the PCOS group

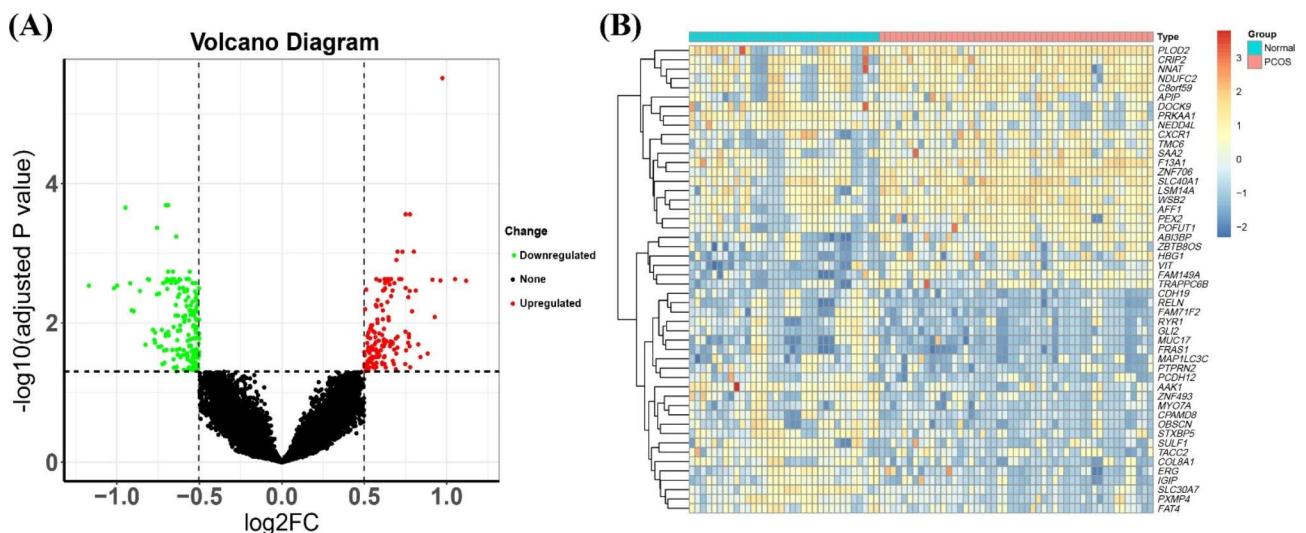


Fig. 3 Analysis of DEGs. **(A)** Volcano diagram of DEGs between patients with PCOS and normal controls. Green represents significantly downregulated genes in the PCOS group, red represents significantly upregulated genes in the PCOS group, and black represents genes with no difference between groups. **(B)** Heat map of the top 50 genes with larger absolute value differences in expression between sample groups. The colour of the block represents the normalised expression of corresponding genes in the sample

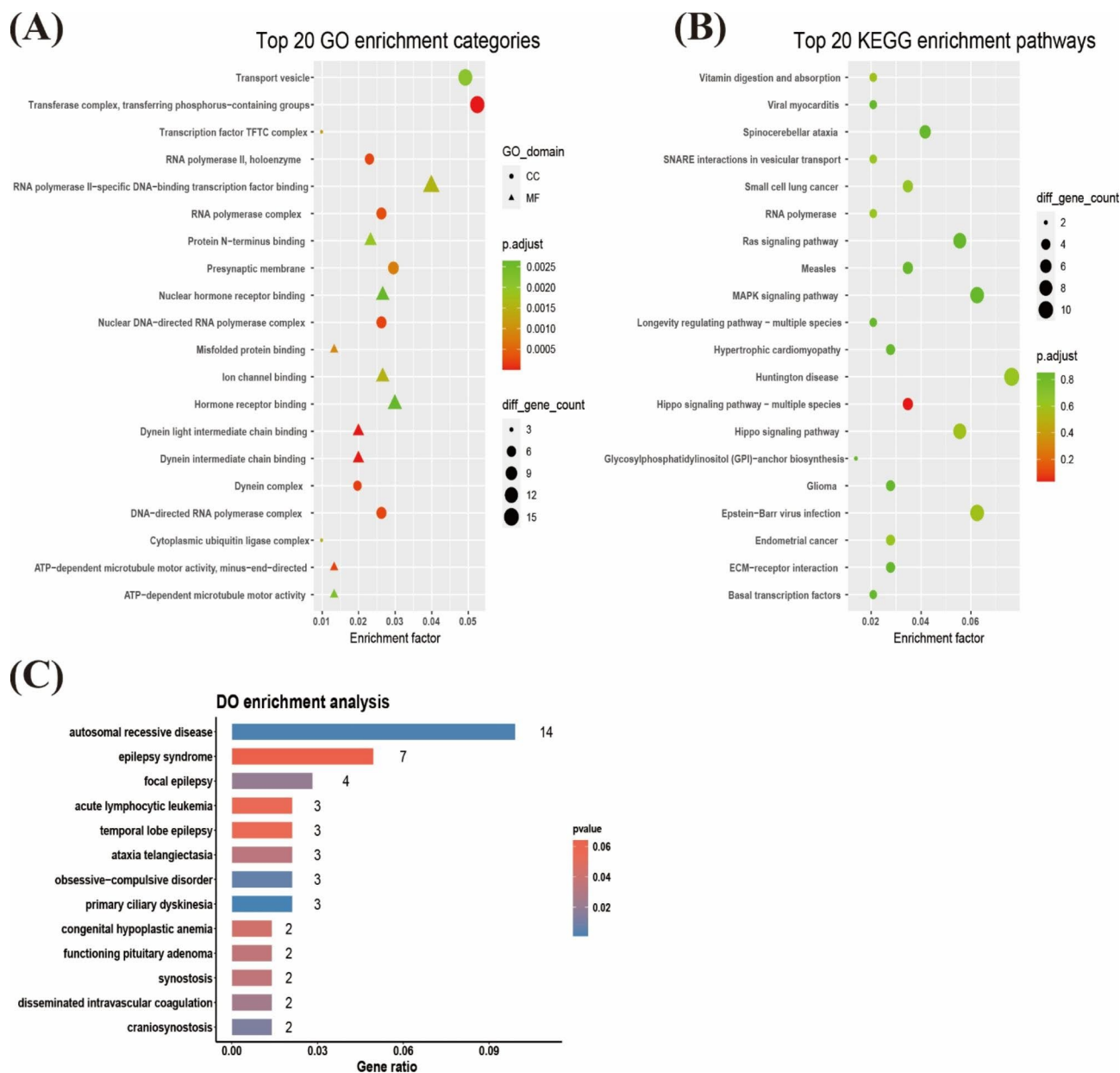


Fig. 4 Functional correlation analysis of DEGs. **(A)** Top 20 results from GO enrichment analysis of DEGs between the PCOS and normal control groups. The shape of the pattern indicates the annotation of GO, the size of the pattern indicates the number of enriched genes, and the colour of the pattern indicates the p value. **(B)** Top 20 results from KEGG enrichment analysis of DEGs between the PCOS and normal control groups. The size of the pattern indicates the number of enriched genes, and the colour of the pattern indicates the p value. **(C)** Partial results from DO analysis of DEGs between the PCOS and normal control groups. The abscissa indicates the number of enriched genes, and the colour indicates the p value

Evaluation of the distributions of immune cell subtypes

ssGSEA revealed differences in the compositions of 28 immune cell subsets in the follicular microenvironment between healthy controls and patients with PCOS (Fig. 2A and Supplementary Table S2). Spearman co-expression analysis of 28 types of immune cells (Fig. 2B and Supplementary Table S3) showed that macrophages were significantly positively correlated with activated

CD8⁺ T cells. Myeloid inhibitory cells were positively correlated with Tregs, Th1 cells, and activated DCs. Central memory CD4⁺ T cells were negatively correlated with Th2 cells. Additionally, compared with normal controls, patients with PCOS showed increased infiltration of central memory CD4⁺ T cells, central memory CD8⁺ T cells, and $\gamma\delta$ T cells, but decreased infiltration of effector memory CD4⁺ T cells and Th17 cells (Fig. 2C).

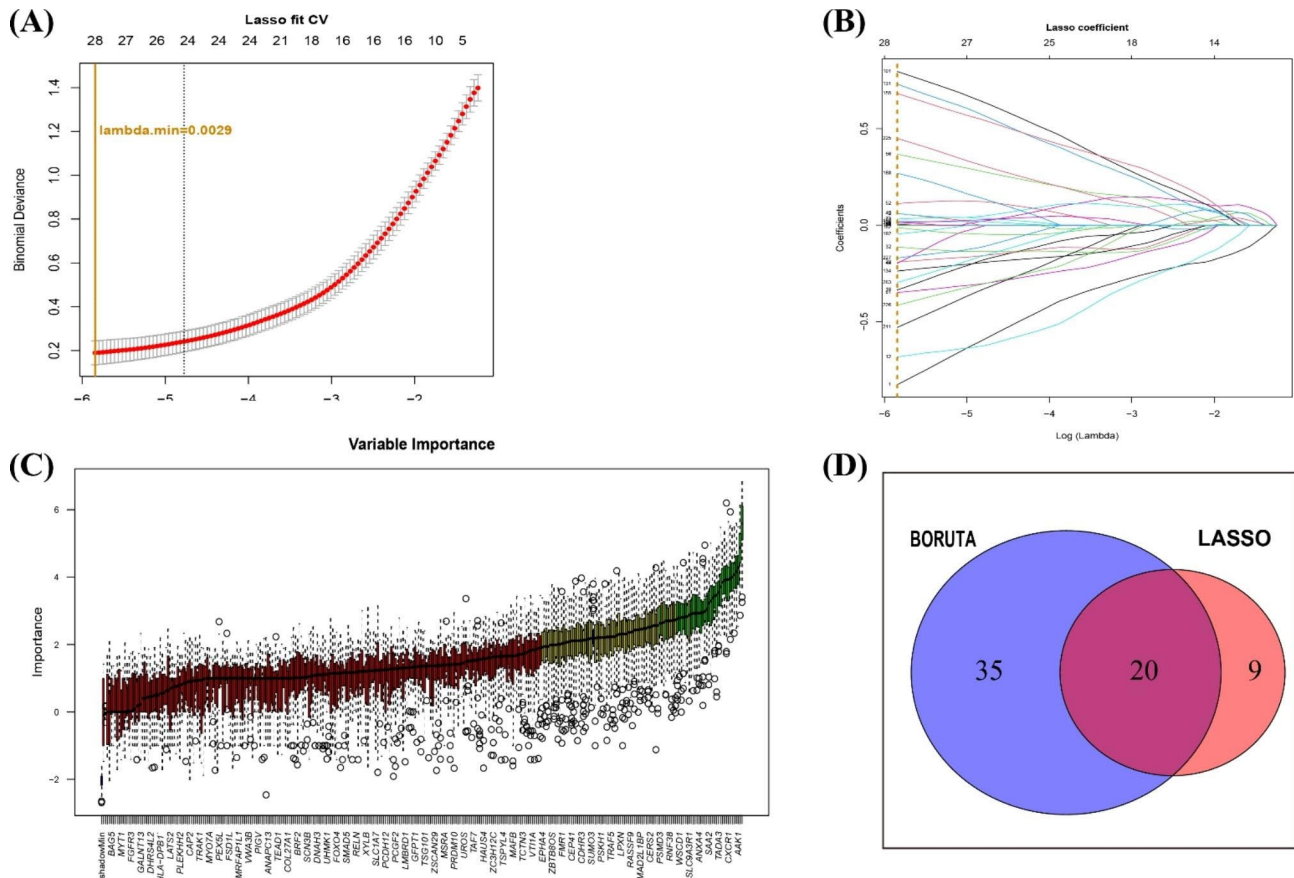


Fig. 5 Screening of PCOS biomarkers. **(A, B)** LASSO regression analysis was used to screen for biomarkers. Different coloured lines represent different genes. **(C)** BORUTA algorithm for feature variable screening. **(D)** The intersection results of the two algorithms were obtained using a Venn diagram to yield candidate biomarkers

Screening of DEGs

The above 83 samples were used to identify DEGs. In total, 325 DEGs were obtained, including 148 upregulated genes and 177 downregulated genes (Supplementary Table S4, Fig. 3A and B).

Functional enrichment analysis

DEG functional enrichment analysis was performed using GO, KEGG, and DO analysis. GO analysis showed that DEGs were mainly related to signal transduction and biological functions of cell activities, such as hormone receptor binding and ATP-dependent microtubule motility (Fig. 4A and Supplementary Table S5). KEGG analysis confirmed that DEGs were mainly related to growth metabolic signalling pathways, such as the Hippo signalling pathway, vesicle transportation, and vitamin digestion and absorption (Fig. 4B and Supplementary Table S6). DO analysis showed that DEGs were mainly related to autosomal recessive diseases and were also related to neurological and systemic haematological diseases (Fig. 4C and Supplementary Table S7).

Screening of biomarkers

PPI network results for DEGs were obtained using STRING (Supplementary Figure S2). LASSO regression analysis was used to screen out 29 candidate biomarkers (Fig. 5A and B and Supplementary Table S8), the BORUTA algorithm was used to extract genes with high importance, and 55 candidate biomarkers were obtained (Fig. 5C and Supplementary Table S6). After analysis of the overlap of biomarkers obtained using the two algorithms, 20 candidate biomarkers were finally obtained (Fig. 5D and Supplementary Table S8), as follows: *AAK1*, *AIFM1*, *CCDC114*, *PTPRN2*, *TMEM54*, *CDH15*, *CXCR1*, *ABI3BP*, *FAM149A*, *SLC9A3R1*, *ISPD*, *VIT*, *TEAD3*, *BAK1*, *PSMD3*, *FUT11*, *PLCG2*, *CFP*, *TRAPPC6B*, and *JDP2*. Most genes were upregulated in PCOS, except *AAK1*, *CCDC114*, *PTPRN2*, *CDH15*, *ISPD*, *FUT11*, and *CFP* (Fig. 6A and B). Correlation analysis between genes showed that the *CXCR1* gene was significantly positively correlated with *SLC9A3R1* and *PSMD3*, whereas *CCDC114* was significantly negatively correlated with *JDP2* (Fig. 6C and Supplementary Table S9). Moreover, correlation analysis between the candidate biomarkers

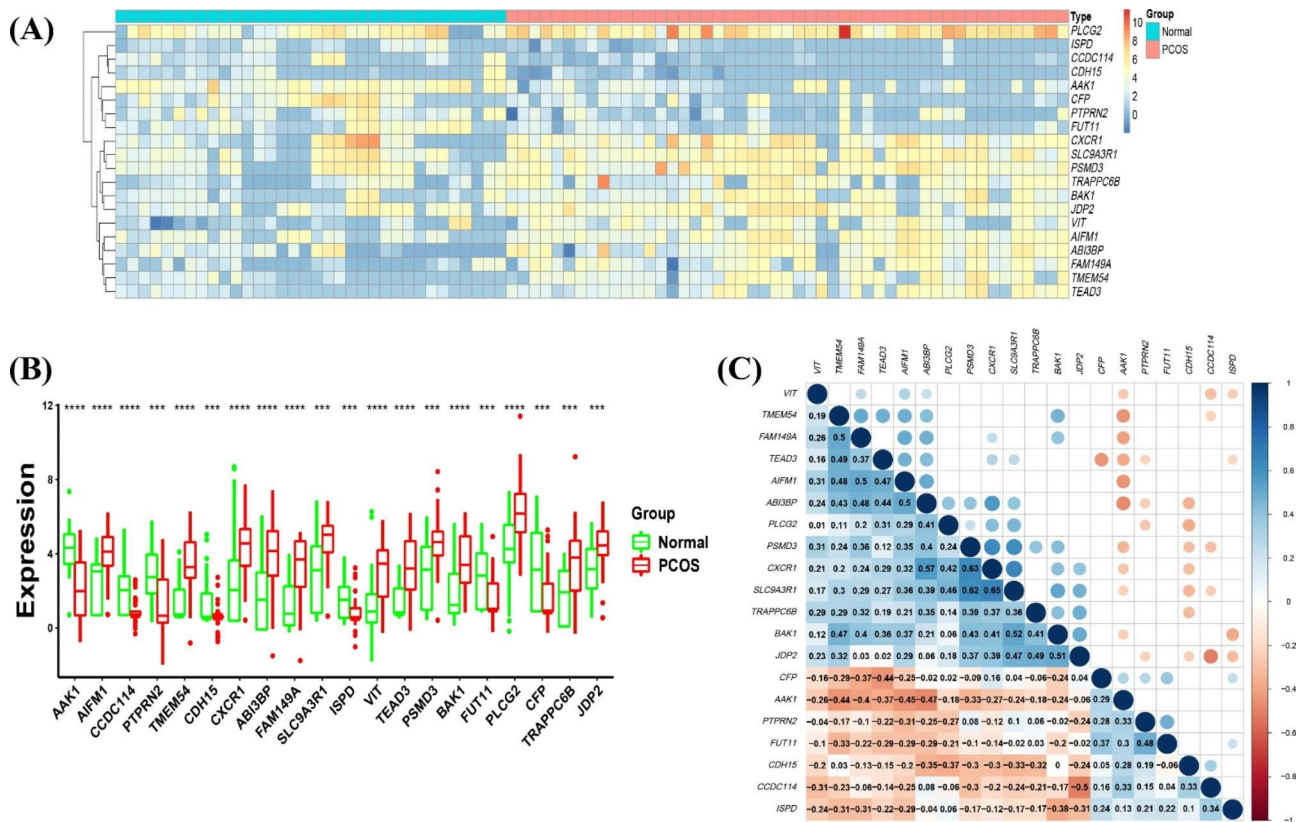


Fig. 6 Expression of candidate biomarkers. **(A)** Heatmap of differences in the expression of candidate biomarkers between PCOS and normal control samples. The colours of the blocks represent the normalised gene expression levels in the samples. **(B)** Boxplot of differences in the expression of candidate biomarkers between groups. *** $p < 0.001$, and **** $p < 0.0001$. **(C)** Heatmap of the correlations between 20 candidate biomarkers. Colour blocks with correlation coefficients greater than 0.05 in the upper part of the figure are not displayed. Blue represents positive correlations, orange represents negative correlations, and darker colours indicate stronger correlations

and 28 types of immune cells showed that *AIFM1* was negatively correlated with activated $CD4^+$ T cells ($R = -0.42$), and *CXCR1* was positively correlated with neutrophils ($R = 0.72$; Fig. 7 and Supplementary Table S9).

Analysis of the predictive value of biomarkers

ROC analysis of candidate biomarkers showed that the combination of *TMEM54* and *PLCG2* had a high value for distinguishing PCOS from healthy women of reproductive age ($AUC = 0.922$). This indicated that the expression of *TMEM54* and *PLCG2* was related to PCOS and that combination of these two genes could be used as a biomarker for distinguishing PCOS from healthy women of reproductive age and for evaluating the efficacy of PCOS treatment (Fig. 8).

The predictive value of the selected biomarkers was tested using the external dataset GSE193812 and collected peripheral blood, and it was found that the biomarkers selected in this study still have good predictive ability both in the external dataset ($AUC = 0.812$) and in the peripheral blood samples ($AUC = 0.889$), as shown in Fig. 9.

Discussion

PCOS is one of the most common ovo-derived endocrine diseases in women of reproductive age. Its basic pathophysiological features include hyperandrogenaemia and IR caused by disruption of the local internal environment, cytokine expression, and ovary function.

In this study, we identified DEGs between PCOS and normal controls and showed that the combination of *TMEM54* and *PLCG2* was a biomarker of PCOS. We also found that central memory $CD4^+$ T cells, central memory $CD8^+$ T cells, effector memory $CD4^+$ T cells, $\gamma\delta$ T cells, and Th17 cells may influence the occurrence of PCOS. *PLCG2* was highly correlated with $\gamma\delta$ T cells and central memory $CD4^+$ T cells.

The follicular microenvironment is composed of FF and granulosa cells (GCs). GCs regulate the local microenvironment of FF through various mechanisms, participate in the metabolism of oocytes, and protect oocytes from the invasion of components in the external environment [20]. The composition of the oocyte GC regulatory loop and follicular microenvironment is critical for the coordination of reproductive activities, and any changes in the composition of FF/GC molecules may affect the quality

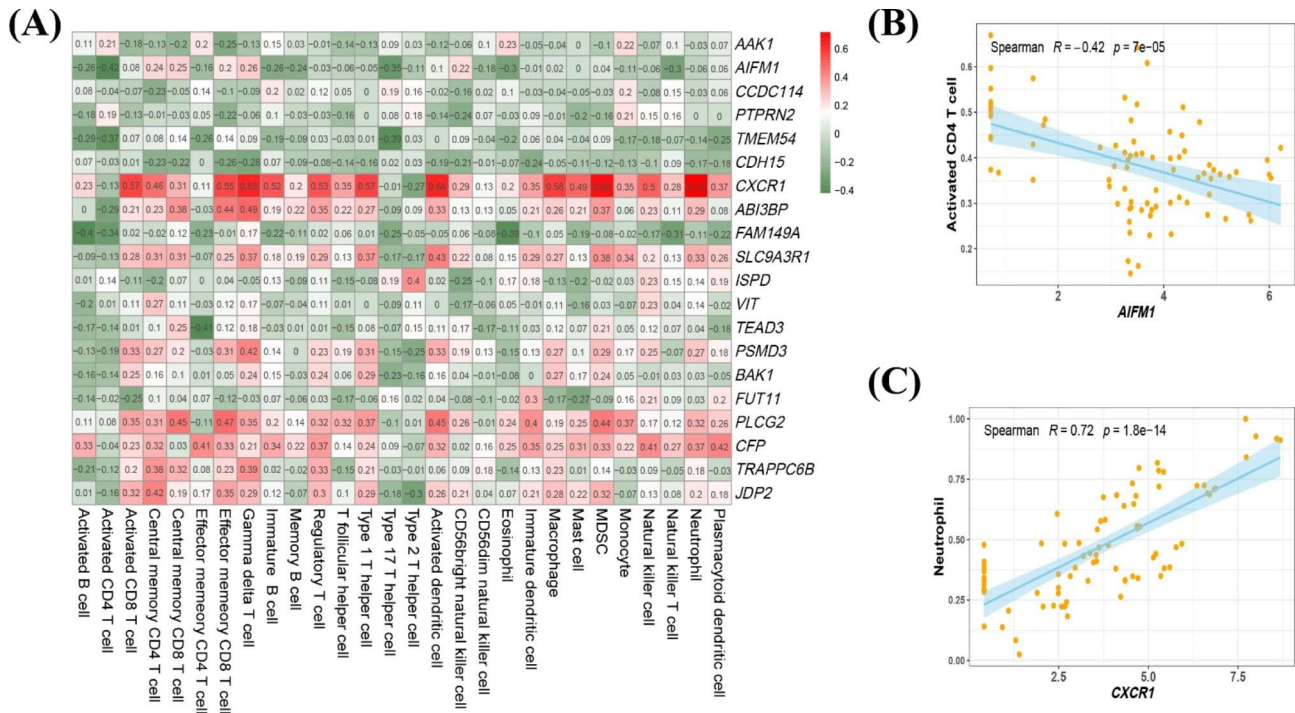


Fig. 7 Correlations between candidate biomarkers and 28 types of immune cells. **(A)** Heatmap of the correlations between 28 immune cells and 20 candidate biomarkers. The colours of the blocks indicate the magnitude of the correlation. **(B, C)** Scatter plots of the two groups of correlations with the largest absolute values of positive and negative correlations

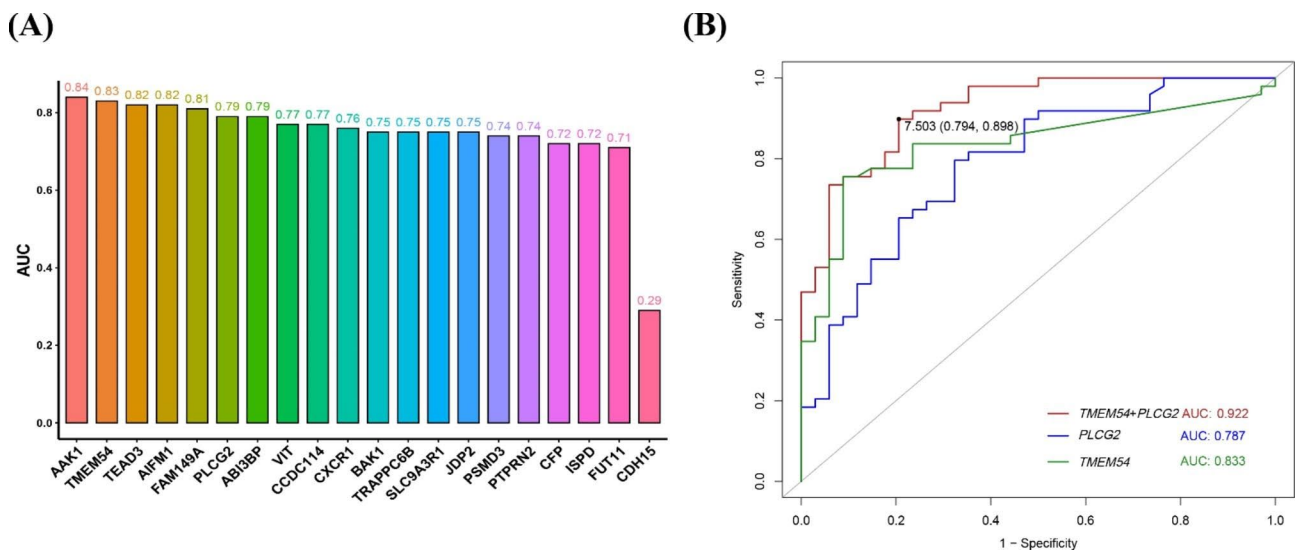


Fig. 8 Predictive value of biomarkers. **(A)** The predictive value of 20 candidate genes for PCOS was evaluated using a bar graph of the AUC. An AUC value greater than 0.85 indicated that the model differentiating effect was satisfactory. **(B)** Diagnostic efficacy comparison of the ROC curves of *TMEM54* and *PLCG2* combined and separate

of oocytes [21]. Therefore, analysis of the immune infiltrating microenvironment of GCs can provide insights into the microenvironment of FF.

As the most important adaptive immune response cells in the immune system, T lymphocytes secrete various inflammatory and immunomodulatory molecules and are involved in regulating multiple ovarian functions,

such as follicle formation, ovulation, and luteal degeneration. Activated lymphocytes secrete inflammatory cytokines, including IL-1, IL-6, IL-12, TNF- α , and insulin-like growth factor-1 [22]. In recent years, many domestic and foreign studies have reported that the levels of inflammatory factors, such as IL-6, TNF- α , and C-reactive protein, are higher in patients with PCOS than in normal women

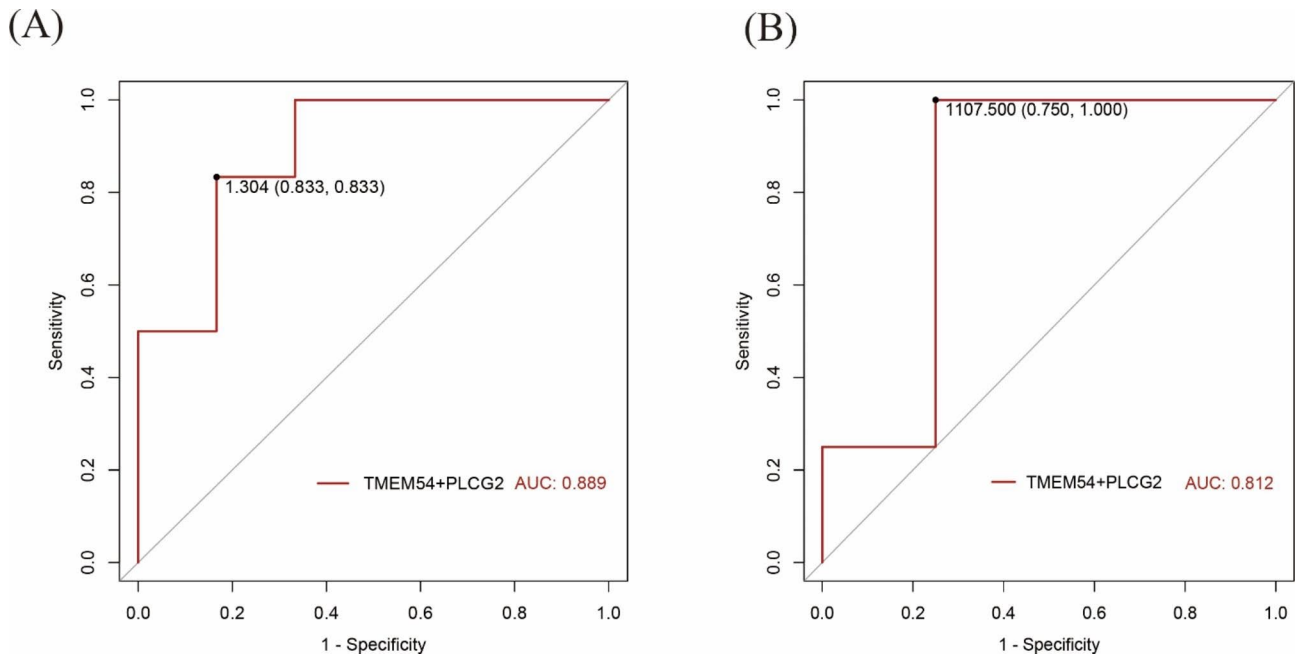


Fig. 9 (A) Diagnostic efficacy of the ROC curves of *TMEM54* and *PLCG2* combined in RT-qPCR. (B) Diagnostic efficacy of the ROC curves of *TMEM54* and *PLCG2* combined in GSE193812.

[23]. Moreover, inflammatory factors can mediate IR directly or indirectly through adipose tissue, suggesting that the occurrence and development of PCOS are closely related to inflammatory factors [24, 25]. Mature T cells can be categorised as CD4⁺ T cells or CD8⁺ T cells according to their CD molecular phenotype. Studies have shown that the ratio of CD4⁺/CD8⁺ is related to IL-2 [26, 27], IL-7 [28] and IL-16 [29]. CD4⁺ T cells and CD8⁺ T cells have different functions, which play synergistic or restrictive roles, and the ratio of CD4⁺/CD8⁺ T cells reflects changes in immune regulation.

Central memory T cells have the ability to expand, differentiate, and self-renew like stem cells and can differentiate unidirectionally into effector memory T cells and short-term effector T cells to prevent chronic infection and cancer [30]. In our study, central memory CD8⁺ T cells and central memory CD4⁺ T cells were increased in patients with PCOS, whereas effector memory CD4⁺ T cells were decreased, albeit without statistical significance. In the local ovarian microenvironment of patients with PCOS, abnormal CD4⁺/CD8⁺ T cell ratios may disrupt immune regulation. However, the differentiation of central memory T cells into effector memory T cells may also be blocked, resulting in increased central memory T cells and decreased effector memory T cells.

The functions of $\gamma\delta$ T cells were first discovered and studied in the field of autoimmune rheumatism. Although $\gamma\delta$ T cells are a highly conserved T-cell subpopulation, they have important implications in various aspects of immunobiology [31]. Animal models have

shown that $\gamma\delta$ T cells regulate classical autoantigen reactive $\alpha\beta$ T cells and B cells [32, 33] and play independent pro-inflammatory roles via direct secretion of IL-17 [34], TNF- α , and interferon (IFN)- γ [35] in a non-antigen-driven pattern. Many studies have confirmed that the levels of TNF- α , IFN- γ , and IL-18 are significantly increased in patients with PCOS and are positively correlated with IR [24, 36, 37]. In this study, the numbers of $\gamma\delta$ T cells were higher in patients with PCOS than in normal controls, this observation may be related to the observed increases in TNF- α and IFN- γ secretion by high numbers of $\gamma\delta$ T cells. Notably, CD3⁺CD4⁻CD8⁻ $\gamma\delta$ T cells are increased in women with recurrent abortion, contributing to foetal loss by regulating transforming growth factor- β and IL-17 secretion and promoting inflammation [38]. The risk of abortion after pregnancy in patients with PCOS is more than three times higher than that in normal women, which may be related to the presence of a chronic inflammatory state and autoimmune disorders, such as active autoimmunity.

The high correlation between the CXC motif chemokine receptor 1 (*CXCR1*) gene and various immune cells was notable, and the correlation between *CXCR1* and neutrophils reached $R=0.72$. The protein encoded by the *CXCR1* gene is a member of the G-protein-coupled receptor family and acts as a receptor for IL-8. IL-8 is a powerful neutrophil chemokine that binds to receptors and promotes neutrophil activation. *CXCR1* binds to IL-8 with high affinity and transduces signals through a second messenger system activated by G proteins [39, 40].

This explains the results of our analysis from a mechanistic level and supports the accuracy of our results.

One of the selected biomarkers, transmembrane protein 54 (*TMEM54*), is a member of the transmembrane protein family, which contains many proteins with unknown functions. Studies have shown that *TMEM* expression is different in tumour tissues than in adjacent healthy tissues, and some *TMEM* family members have been identified as potential prognostic biomarkers in different types of tumours [41]. In addition, *TMEM* proteins are tumour suppressors or oncogenes and have been shown to be associated with tumour progression and invasion [42, 43] or chemotherapy resistance [44, 45]. Although there are few studies on *TMEM54*, our current findings suggest that the *TMEM54* gene may have important roles in the development of PCOS; further studies are warranted.

Phospholipase C gamma 2 (*PLCG2*) is a transmembrane signalling enzyme that is an important driver of many immunological aetiological diseases, such as inflammation, autoimmune diseases, immunodeficiencies, and allergies, as well as haematological malignancies. Some studies have demonstrated that point mutations in the *PLCG2* gene may be an important cause of severe spontaneous inflammation and autoimmunity [46]. In a meta-analysis of gene expression in patients with rheumatoid arthritis (RA), *PLCG2* was found to be upregulated in several datasets, including many pathways associated with RA inflammatory responses, e.g., inflammasome activation, platelet aggregation, and activation, indicating that *PLCG2* is a potential target for the control of RA inflammation [47]. Moreover, *PLCG2* is important in bone marrow cells, including monocytes, macrophages, NK cells, DCs, and mast cells, possibly because it promotes downstream signalling involving Fc receptors [48]. In a bioinformatics analysis of the tumour microenvironment in soft tissue sarcoma, *PLCG2* was found to be an indicator of the tumour microenvironment and patient prognosis [49]. Additionally, CD8⁺ T cells, $\gamma\delta$ T cells, monocytes, and M1 macrophages were shown to be positively correlated with *PLCG2* expression, consistent with our current results.

There were some limitations to this study. First, this study was only carried out from the perspective of gene transcriptome, and multi-omics and mechanistic studies were not performed. Further validation of our bioinformatics results through *in vitro* and *in vivo* experiments and clinical practice is needed.

In summary, in this study, ssGSEA was used for the first time to analyse immune infiltration into the follicle microenvironment in patients with PCOS. Central memory CD4⁺ T cells, central memory CD8⁺ T cells, effector memory CD4⁺ T cells, $\gamma\delta$ T cells, and Th17 cells may be involved in the occurrence of PCOS. In addition,

differences in gene expression in ovarian tissues between patients with PCOS and healthy women of reproductive age were determined, and our findings showed that the combination of *TMEM54* and *PLCG2* was a biomarker of PCOS. *PLCG2* was shown to be highly correlated with $\gamma\delta$ T cells and central memory CD4⁺ T cells. These findings provide a basis for further research on the immunological pathogenesis of PCOS.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12884-023-05693-4>.

Additional file 1: Supplementary Tables

Additional file 2: Supplementary Figures

Acknowledgements

We would like to thank all the patients for participation in this study. We would also like to thank Editage (www.editage.cn) for English language editing and the GEO Database for providing us with a shareable dataset.

Authors' contribution

All authors contributed to the article and approved the final version for submission.

Funding

This work was supported by Guangzhou Science and Technology Plan Project (grant number: 202102080062) and Guangdong Province Medical Research Funding (grant number: B2023135).

Availability of data and materials

The datasets used and/or analysed during the current study are available from NCBI Gene Expression Omnibus (GEO). GSE84958 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84958>. GSE106724 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106724>. GSE114419 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114419>. GSE137684 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE137684>. GSE193812 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE193812>.

Declarations

Ethics approval and consent to participate

The studies involving human participants were reviewed and approved by the Ethics Committee of Guangdong Provincial Fertility Hospital. All methods were performed according to the Declaration of Helsinki. Anonymized and deidentified information for participants was used for analysis, so the requirement for informed consent permission was waived by the Ethics Committee of Guangdong Provincial Fertility Hospital.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 18 December 2022 / Accepted: 9 May 2023

Published online: 24 May 2023

References

1. Qiao J, Li R, Li L. Polycystic ovary syndrome – an epidemiological study of polycystic ovary syndrome. *Chin J Practical Gynecol Obstet*. 2013;29(11):849–52.

2. Brännström M, Enskog A. Leukocyte networks and ovulation. *J Reprod Immunol*. 2002;57(1–2):47–60.
3. Duffy DM, Ko C, Jo M, Brannstrom M, Curry TE. Ovulation: parallels with inflammatory processes. *Endocr Rev*. 2019;40(2):369–416.
4. Repaci A, Gambineri A, Pasquali R. The role of low-grade inflammation in the polycystic ovary syndrome. *Mol Cell Endocrinol*. 2011;335(1):30–41.
5. Barrea L, Marzullo P, Muscogiuri G, Di Somma C, Scacchi M, Orio F, et al. Source and amount of carbohydrate in the diet and inflammation in women with polycystic ovary syndrome. *Nutr Res Rev*. 2018;31(2):291–301.
6. Wu R, Fujii S, Ryan NK, Van der Hoek KH, Jasper MJ, Sini I, et al. Ovarian leukocyte distribution and cytokine/chemokine mRNA expression in follicular fluid cells in women with polycystic ovary syndrome. *Hum Reprod (Oxford England)*. 2007;22(2):527–35.
7. Solano ME, Sander VA, Ho H, Motta AB, Arck PC. Systemic inflammation, cellular influx and up-regulation of ovarian VCAM-1 expression in a mouse model of polycystic ovary syndrome (PCOS). *J Reprod Immunol*. 2011;92(1–2):33–44.
8. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinf (Oxford England)*. 2012;28(6):882–3.
9. Spitzer M, Wildenhain J, Rappsilber J, Tyers M. BoxPlotR: a web tool for generation of box plots. *Nat Methods*. 2014;11(2):121–2.
10. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
11. Smyth GK. *limma: Linear Models for microarray data*. New York: Springer; 2005.
12. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51(D1):D587–d92.
13. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
14. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OmicS-a J Integr Biology*. 2012;16(5):284–7.
15. Robert T. *Methodological B*. 1996.
16. Kursa MB, Rudnicki WR. Feature selection with Boruta Package. *J Stat Softw*. 2010;36(11):1–13.
17. Friendly M, Corrgams. Exploratory displays for correlation matrices. *Am Stat*. 2002;56(November):316–24.
18. Gómez-Rubio V. *ggplot2 - Elegant Graphics for Data Analysis (2nd Edition)*. *Journal of Statal Software*. 2017;077(Book review 2).
19. Kolde R. *heatmap: Pretty Heatmaps*. 2015.
20. Dumesic DA, Meldrum DR, Katz-Jaffe MG, Krisher RL, Schoolcraft WB. Oocyte environment: follicular fluid and cumulus cells are critical for oocyte health. *Fertil Steril*. 2015.
21. Palomba S, Daolio J, Sala GL. Oocyte competence in women with polycystic ovary syndrome. *Trends in Endocrinology & Metabolism*. 2017;28(3):186–98.
22. Walusimbi SS, Pate JL. Physiology and Endocrinology Symposium: role of immune cells in the corpus luteum. *Journal of animal science*. 2013;91(4):1650–9.
23. Deligeorgoglou E, Vrachnis N, Athanasopoulos N, Iliodromiti Z, Sifakis S, Iliodromiti S, et al. Mediators of chronic inflammation in polycystic ovarian syndrome. *Gynecol endocrinology: official J Int Soc Gynecol Endocrinol*. 2012;28(12):974–8.
24. Escobar-Morreale HF, Botella-Carretero JJ, Villuendas G, Sancho J, San Millán JL. Serum interleukin-18 concentrations are increased in the polycystic ovary syndrome: relationship to insulin resistance and to obesity. *J Clin Endocrinol Metab*. 2004;89(2):806–11.
25. Atabekoglu CS, Sönmez M, Özmen B, Yarci A, Akbiyik F, Taşçı T, et al. Increased monocyte chemoattractant protein-1 levels indicating early vascular damage in lean young PCOS patients. *Fertil Steril*. 2011;95(1):295–7.
26. Pozzesi N, Gizzi S, Gori F, Vacca C, Cannarile L, Riccardi C, et al. IL-2 induces and altered CD4/CD8 ratio of splenic T lymphocytes from transgenic mice overexpressing the glucocorticoid-induced protein GILZ. *J Chemother*. 2007;19(5):562–9.
27. Räsänen L, Jansén CT, Hyöty H, Reunala T, Morrison H. Cis-urocanic acid stereospecifically modulates human monocyte IL-1 production and surface HLA-DR antigen expression, T-cell IL-2 production and CD4/CD8 ratio. *Photodermatology*. 1989;6(6):287–92.
28. Geislerhart LA, Humphries CA, Gregorio TA, Mou S, Subleski J, Komschlies KL. IL-7 administration alters the CD4:CD8 ratio, increases T cell numbers, and increases T cell function in the absence of activation. *Journal of immunology (Baltimore, Md: 1950)*. 2001;166(5):3019–27.
29. Koike M, Sekigawa I, Okada M, Matsumoto M, Iida N, Hashimoto H, et al. Relationship between CD4(+)/CD8(+) T cell ratio and T cell activation in multiple myeloma: reference to IL-16. *Leuk Res*. 2002;26(8):705–11.
30. Pais Ferreira D, Silva JG, Wyss T, Fuertes Marraco SA, Scarpellino L, Charmoy M, et al. Central memory CD8(+) T cells derive from stem-like Tcf7(hi) effector cells in the absence of cytotoxic differentiation. *Immunity*. 2020;53(5):985–1000e11.
31. Bank I. The role of Gamma Delta T cells in Autoimmune Rheumatic Diseases. *Cells*. 2020;9(2).
32. Samuelson EM, Laird RM, Papillion AM, Tatum AH, Princiotta MF, Hayes SM. Reduced B lymphoid kinase (blk) expression enhances proinflammatory cytokine production and induces nephrosis in C57BL/6-lpr/lpr mice. *PLoS ONE*. 2014;9(3):e92054.
33. Rezende RM, Lanser AJ, Rubino S, Kuhn C, Skillin N, Moreira TG, et al. $\gamma\delta$ T cells control humoral immune response by inducing T follicular helper cell differentiation. *Nat Commun*. 2018;9(1):3151.
34. Reinhardt A, Yevsa T, Worbs T, Lienenklaus S, Sandrock I, Oberdörfer L, et al. Interleukin-23-Dependent $\gamma\delta$ T cells produce Interleukin-17 and Accumulate in the Entesis, aortic valve, and ciliary body in mice. *Volume 68. Arthritis & rheumatology*; 2016. pp. 2476–86. (Hoboken, NJ). 10.
35. Avau A, Mitera T, Put S, Put K, Brisse E, Filtjens J, et al. Systemic juvenile idiopathic arthritis-like syndrome in mice following stimulation of the immune system with Freund's complete adjuvant: regulation by interferon- γ . *Arthritis & rheumatology (Hoboken NJ)*. 2014;66(5):1340–51.
36. Choi YS, Yang HI, Cho S, Jung JA, Jeon YE, Kim HY, et al. Serum asymmetric dimethylarginine, apelin, and tumor necrosis factor- α levels in non-obese women with polycystic ovary syndrome. *Steroids*. 2012;77(13):1352–8.
37. Kim CH, Ahn JW, You RM, Kim SH, Chae HD, Kang BM. Pioglitazone treatment decreases follicular fluid levels of tumor necrosis factor- α and interleukin-6 in patients with polycystic ovary syndrome. *Clin experimental reproductive Med*. 2011;38(2):98–102.
38. Talukdar A, Rai R, Aparna Sharma K, Rao DN, Sharma A. Peripheral Gamma Delta T cells secrete inflammatory cytokines in women with idiopathic recurrent pregnancy loss. *Cytokine*. 2018;102:117–22.
39. Bishayi B, Nandi A, Dey R, Adhikary R. Expression of CXCR1 (IL-8 receptor A) in splenic, peritoneal macrophages and resident bone marrow cells after acute live or heat killed *Staphylococcus aureus* stimulation in mice. *Microb Pathog*. 2017;109:131–50.
40. Jaufmann J, Carevic M, Tümen L, Eliacik D, Schmitt F, Hartl D, et al. Enhanced IgG(1)-mediated antibody response towards thymus-dependent immunization in CXCR1-deficient mice. *Immun Inflamm Dis*. 2021;9(1):210–22.
41. Ma R, Feng N, Yu X, Lin H, Zhang X, Shi O, et al. Promoter methylation of Wnt/ β -Catenin signal inhibitor TMEM88 is associated with unfavorable prognosis of non-small cell lung cancer. *Cancer biology & medicine*. 2017;14(4):377–86.
42. Hu R, Hu F, Xie X, Wang L, Li G, Qiao T, et al. TMEM45B, up-regulated in human lung cancer, enhances tumorigenicity of lung cancer cells. *Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2016;37(9):12181–91.
43. Shen K, Yu W, Yu Y, Liu X, Cui X. Knockdown of TMEM45B inhibits cell proliferation and invasion in gastric cancer. *Biomed pharmacotherapy = Biomedecine pharmacotherapie*. 2018;104:576–81.
44. Schmit K, Michiels C. TMEM Proteins in Cancer: a review. *Front Pharmacol*. 2018;9:1345.
45. de Leon M, Cardenas H, Vieth E, Emerson R, Segar M, Liu Y, et al. Transmembrane protein 88 (TMEM88) promoter hypomethylation is associated with platinum resistance in ovarian cancer. *Gynecol Oncol*. 2016;142(3):539–47.
46. Yu P, Constien R, Dear N, Katan M, Hanke P, Bunney TD, et al. Autoimmunity and inflammation due to a gain-of-function mutation in phospholipase C gamma 2 that specifically increases external Ca²⁺ entry. *Immunity*. 2005;22(4):451–65.
47. Afroz S, Giddaluru J, Vishwakarma S, Naz S, Khan AA, Khan N. A Comprehensive Gene expression Meta-analysis identifies Novel Immune Signatures in Rheumatoid Arthritis Patients. *Front Immunol*. 2017;8:74.
48. Jakus Z, Simon E, Frommhold D, Sperandio M, Mócsai A. Critical role of phospholipase Cgamma2 in integrin and fc receptor-mediated neutrophil functions and the effector phase of autoimmune arthritis. *J Exp Med*. 2009;206(3):577–93.
49. Li Z, Zhao R, Yang W, Li C, Huang J, Wen Z, et al. PLAG2 as a potential indicator of tumor microenvironment remodeling in soft tissue sarcoma. *Medicine*. 2021;100(11):e25008.

Publisher's Note

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