scientific reports

Check for updates

OPEN

Identification of MAP1LC3A as a promising mitophagy-related gene in polycystic ovary syndrome

Yizhen Yang[1,2,3], Xiaojing Chen[1,2,3], Xiuhua Liao[1], Wenwen Jiang[1], Yuan Zhou[1], Yan Sun[1,2]✉ & Beihong Zheng[1,2]✉

Increasing evidence suggests that mitophagy is crucially involved in the progression of polycystic ovary syndrome (PCOS). Exploration of PCOS-specific biomarkers related to mitophagy is expected to provide critical insights into disease pathogenesis. In this study, we employed bioinformatic analyses and machine learning algorithms to determine novel biomarkers for PCOS that may be tied with mitophagy. A grand total of 12 differential expressed mitophagy-related genes (DE-MRGs) associated with PCOS were identified. TOMM5 and MAP1LC3A among the 12 DE-MRGs were recognized as potential marker genes by LASSO, RF and SVM-RFE algorithms. The area under the ROC curve (AUROC) of MAP1LC3A were all greater than 0.8 both in the training set and validation sets. The CIBERSORT analysis indicated a potential association between alterations in the immune microenvironment of PCOS individuals and MAP1LC3A expression. In addition, we found that MAP1LC3A was positively related to the testosterone levels of PCOS patients. Overall, MAP1LC3A was identified as optimal PCOS-specific biomarkers related to mitophagy. Our findings created a diagnostic strength and offered a perspective for investigating the mitophagy process in PCOS.

Keywords Polycystic ovary syndrome, Mitophagy-related gene, MAP1LC3A, Bioinformatic analysis, Machine learning, Immune cell infiltration

Abbreviations

AUROC	Area under the ROC curve
BP	Biological process
CC	Cellular component
CRP	C-reactive protein
DE-MRGs	Differentially expressed mitophagy-related genes
DHT	Dihydrotestosterone
FSH	Follicle-stimulating hormone
GC	Granulosa cell
GEO	Gene expression omnibus
GnRH	Gonadotropin-releasing hormone
GO	Gene ontology
KEGG	Kyoto encyclopedia of genes and genomes
LASSO	Least absolute shrinkage and selection operator
LH	Luteinizing hormone
LIR	LC3-interacting region
MF	Molecular function
MRGs	Mitophagy-related genes
OMM	Outer mitochondrial membrane
PCA	Principal component analysis
PCOS	Polycystic ovarian syndrome
RF	Random forest classifier
ROC	Receiver operating characteristic
SVM-RFE	Support vector machine recursive feature elimination

[1]Reproductive Medicine Center of Fujian Maternity and Child Health Hospital, Fuzhou 350001, Fujian, China. [2]Fujian Maternal-Fetal Clinical Medicine Research Center, Fuzhou 350001, Fujian, China. [3]These authors contributed equally: Yizhen Yang and Xiaojing Chen. ✉email: sunyanteam@163.com; zhengbeihong2010@163.com

Polycystic ovarian syndrome (PCOS) is a prevalent endocrine-metabolic disorder that impacts a large population globally, the "Rotterdam criteria", defined in 2003[1], determined the diagnosis of PCOS based on oligomenorrhea or amenorrhea, hyperandrogenism, ovarian polycystic changes, and infertility. Although numerous studies have focused on PCOS, the precise origin and etiology remain incompletely comprehended.

The prime organ affected in PCOS is the ovary, in which macroautophagy/autophagy performs a pivotal role in directing the chain of events starting from oocytes origin until its fertilization. Defective autophagy in the follicular cells during different stages of follicles is observed in the PCOS ovary[2]. Lately, there has been considerable attention on the notion of selective autophagy prompted by the autophagy substrate acting as a catalyst. Mitophagy, being one of the extensive-studied forms of selective autophagy, plays a critical role in maintaining the function and genetic stability of mitochondria[3]. It is believed to be an important component in the onset of PCOS.

Moreover, the recent wide acceptance of functional mitochondrial disorders as a correlated factor of numerous diseases has led to the presupposition that abnormal mitochondrial metabolic markers are associated with PCOS[4]. Mitophagy serves as a cytoprotective mechanism to eliminate excess or malfunctioning mitochondria, ensuring a proper balance of mitochondrial numbers for intracellular stability. The PINK1-PRKN/Parkin pathway, acknowledged as the primary regulator of mitophagy[5], involves the labeling of impaired mitochondria with ubiquitin chains, initiating their selective autophagy. Accumulation of PINK1 in damaged mitochondria leads to the recruitment of parkin, resulting in ubiquitination of mitochondrial proteins. These can then be bound by the autophagic proteins p62/SQSTM1 and LC3, leading to the degradation of mitochondria through mitophagy. MAP1LC3A, our subject, is precisely one of the members of the human autophagy-related LC3/GABARAP-protein family. In addition, there has been direct evidence that indicate the relationship between mitophagy and PCOS progression. Yi et al. observed that mitophagy is significantly enhanced in dihydrotestosterone (DHT) -induced PCOS-like mice, and melatonin treatment can significantly decrease the levels of PINK1/Parkin, thus improving mitochondrial dysfunction and PCOS phenotype both in vitro and in vivo[6].

There are multiple perspectives on the screening of PCOS biomarkers, including multi-omics, DNA methylome, senescence-related genes and also autophagy-associated mRNA-miRNA-LncRNA network[7–10], but the specific genes associated with mitophagy (mitophagy-related genes, MRGs) that are linked with PCOS have yet to be explored and clearly identified. In our study, to discern the candidate mitophagy-related biomarkers for PCOS, functional enrichment analysis, application of machine-learning algorithms (least absolute shrinkage and selection operator [LASSO], random forest [RF], and support vector machine-recursive feature elimination [SVM-RFE]), evaluation of receiver operating characteristic (ROC) curve, and analysis of immune cell infiltration were successively performed[11]. We came to the conclusion that MAP1LC3A may serve as a promising mitophagy-associated biomarker in PCOS.

Materials and methods

Data sources and processing

The study flow chart is presented in Fig. 1. We obtained three trustworthy transcriptomic datasets related to PCOS from the Gene Expression Omnibus (GEO) database[12], namely GSE95728[13], GSE168404[14], and GSE155489[15]. Detailed information about the three datasets were listed in Table 1. Totally, 32 granulosa cell (GC) samples (16 PCOS, 16 Controls) and 12 oocyte samples (6 PCOS, 6 Controls) were taken into consideration in our research to assess the MRGs' expression levels. Gene symbols were matched to the array probes based on the corresponding annotation data. The normalized gene expression matrices for GSE95728 and GSE168404 were directly downloaded, whereas the total read count per sample in GSE155489 were required to be normalized to a common library size using the DESeqDataSetFromMatrix function from the R package "DESeq2". Then we merged the training datasets together (GSE95728-GC and GSE155489-GC). The ComBat function from the R package "sva"[16] was utilized to eliminate sequencing batch variation, and the impact of inter-sample correction was visualized by a two- dimensional principal component analysis (PCA) cluster plot.

Screening for differential expressed mitophagy-related genes

29 MRGs were taken out from the reactome pathway database (Table 2). The "limma" package[17] was utilized to identify the differentially expressed mitophagy-related genes (DE-MRGs) between PCOS patients and controls. We also utilized the Benjamini-Hochberg-based False Discovery Rate method to modify *P* values and identified the noteworthy DE-MRGs with an adjusted *P*-value < 0.1. Subsequently, we employed the "ggplot2" package to depict a volcano plot and the "pheatmap" package to generate a heatmap.

Annotating the functional aspects of DE-MRGs

The gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) enrichment analyses were conducted for functional annotation of DE-MRGs using the "clusterProfiler" R package[18]. The GO analysis identified three categories, namely biological process (BP), cellular component (CC), and molecular function (MF)[19]. The investigation of potential biological pathways was carried out utilizing KEGG[20]. Significant enrichment was defined as *P*-value < 0.05.

Identification of optimal hub genes for PCOS

To filter feature genes, three machine learning algorithms were performed: LASSO, RF, and SVM-RFE. LASSO regression, serving as a technique for reducing dimensionality, outperforms regression analysis when dealing with high-dimensional data and employs regularization to enhance prediction accuracy. Using the "glmnet" R package, a tenfold cross-verification method with a turning or penalty parameter was conducted[21]. The RF algorithm employed the "Random Forest" package to calculate the error rate and accuracy rate of the combination

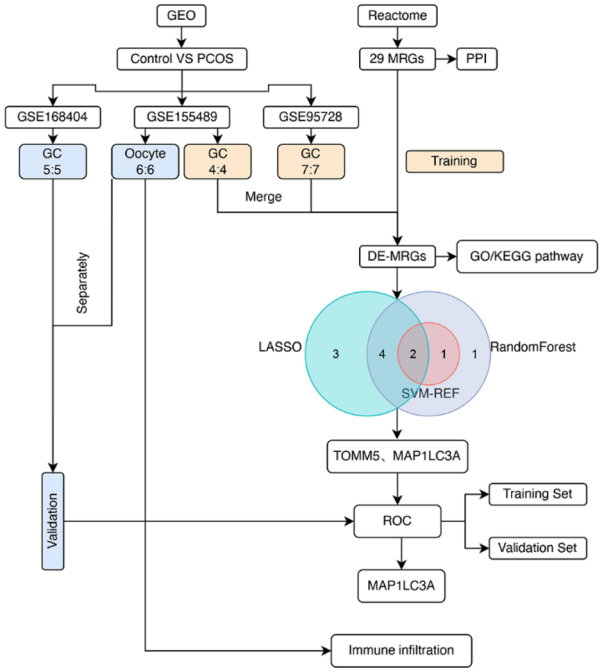


Figure 1. Study flow chart.

Data type	GEO series	Platform	Control	PCOS	Tissue
Training set	GSE95728	GPL16956	7	7	Granulosa cell
Training set	GSE155489	GPL20795	4	4	Granulosa cell
Validation set	GSE155489	GPL20795	6	6	Oocyte
Validation set	GSE168404	GPL16791	5	5	Granulosa cell

Table 1. Detailed information about the collected datasets.

ATG12	ATG5	CSNK2A1	CSNK2A2	CSNK2B
FUNDC1	MAP1LC3A	MAP1LC3B	MFN1	MFN2
MTERF3	PARK2	PGAM5	PINK1	RPS27A
SQSTM1	SRC	TOMM20	TOMM22	TOMM40
TOMM5	TOMM6	TOMM7	TOMM70A	UBA52
UBB	UBC	ULK1	VDAC1	

Table 2. 29 mitophagy-related genes extracted from reactome pathway database.

in each iteration. Additionally, it used the RFE approach to determine the importance and important ranking of each gene. The related genes in the ideal combination with the lowest error rate were the characteristic genes. Meanwhile, the SVM-RFE model was evaluated by calculating the average misclassification rates from their tenfold cross-validations using the R package "e107"[22]. SVM-RFE, a novel technique in machine learning, can avoid overfitting by recursively ranking features[23]. The final feature importance was determined by using the average importance of each feature in every iteration. The genes that fell between the intersections of the three subsets were then chosen as hub genes for further examinations.

Next, we established the ROC curve to assess the discriminatory capacity of the hub genes in the merged training set (GSE95728-GC + GSE155489-GC); the diagnostic performance of each model was measured by the area under the ROC curve (AUROC). The accuracy of these gene predictions would be separately confirmed in the independent validation datasets (GSE168404 [GC, Control vs. PCOS 5: 5] and GSE155489 [Oocyte, Control vs. PCOS 6: 6]). The ROC analyses were completed by the R package "pROC"[24].

Assessment of immune cell infiltration

Based on the principle linear support vector regression, we used CIBERSORT algorithm to estimate the proportion of 22 different types of immune cell infiltration of each oocyte samples in GSE155489[25]. The LM22 matrix file contains 547 genes, which serve as the standard reference leukocyte gene signature. These genes are capable of accurately distinguishing 22 mature human hematological populations that have been isolated from peripheral blood or in vitro cultures[26]. All estimations of immune cell type fractions for each sample added up to 1. To visualize the disparities in immune cell infiltration between the PCOS and control samples, PCA clustering and boxplots were utilized. Correlation analysis and visualization of infiltrating immune cells were performed using the "corrplot" package[27].

Clinical correlation between hub genes and testosterone levels in PCOS patients

According to the expression form of testosterone levels in GSE168404 samples as mean ± standard deviation, we determined that this variable type was in line with normal distribution and homogeneity of variance. Based on the provided values, SPSS software (Rv. Normal function) was applied to generate random numbers that match the mean and standard deviation of testosterone data of individual samples in GSE168404. We then assessed the correlation between hub genes and testosterone levels in patients with PCOS and drew a scatter plot using the "ggplot2" package.

Statistical analysis

R (version 4.3.1) was utilized for all statistical analyses. Group comparisons were undertaken for continuous variables using Student's *t*-test for normally distributed variables or the Mann–Whitney *U*-test for variables with an abnormal distribution. Pearson correlation analyses were applied for the necessary tasks. All statistical analyses were two-sided with P-value < 0.05 were regarded statistically significant.

Results

Panoramic view of MRGs in PCOS

Currently, the role of 29 MRGs has been extensively studied and their interaction was shown in (Fig. 2A). The density distribution of the groups shown in (Fig. 2B) was basically consistent, indicating that the normalized gene expression matrix of each dataset could be applied for subsequent analysis. Next, we merged the GSE95728-GC and GSE155489-GC datasets as a training set (GSE95728-GC + GSE155489-GC) and removed the sequencing batch effects. The PCA cluster plots in (Fig. 2C) showed that the clustering of the two datasets was more obvious after batch removal, indicating that the source of the samples was reliable. Figure 2D showed that 26 MRGs were figured out in the GC samples, of which 12 genes exhibited a marked expression difference between PCOS and the matched controls ($P < 0.05$), including TOMM5 and MTERF3 ($P < 0.0001$).

Volcano in Fig. 3A depicted these 12 DE-MRGs, including 6 up-regulated genes and 6 down-regulated. TOMM5 exhibited the greatest fold-change among these downregulated genes, while MAP1LC3A had the greatest fold-change among those upregulated. The heatmap in Fig. 3B showed the expression of DE-MRGs among GC samples. In the correlation analysis (Fig. 3C), we found that these genes were closely related, indicating that they may work together. The scatterplots of Fig. 3C displayed the genes with the highest positive and negative correlation, specifically, CSNK2A2 and VDAC1 turned out to be the most negative correlation, whereas MTERF3 were most positively correlated with TOMM5.

GO and KEGG analysis of the DE-MRGs

Based on the GO and KEGG databases, we analyzed the functional enrichment of DE-MRGs. Figure 4A showed the 15 highest-ranking GO terms, including organelle disassembly, autophagy of mitochondrion, mitochondrion disassembly, protein targeting to mitochondrion, establishment of protein localization to mitochondrion and macroautophagy (Fig. 4B). The KEGG analysis revealed that the DE-MRGs were involved in the process of neurodegeneration-multiple disorders and the process of mitophagy-animal (Fig. 4C,D).

Identification of hub genes

For a better understanding of the diagnostic potential of DE-MRGs, we then constructed a prediction model for the diagnosis of PCOS applying three different algorithms to distinguish the PCOS patients from healthy controls. The 12 candidate genes were successively submitted into LASSO, RF and SVM-RFE. 9 out of 12 PCOS-related features of non-zero coefficients were filtered by the means of LASSO algorithm (Fig. 5A,B). Next, we identified feature importance using RF and the top 8 genes were selected as diagnostic genes, as shown in (Fig. 5C,D). And then, features were selected and 3 genes were identified as the best candidates for PCOS based on SVM-RFE (Fig. 5E,F). Finally, we crossed the candidate genes obtained from LASSO, RF, and SVM-RFE models and identified 2 hub genes (TOMM5 and MAP1LC3A) for follow-up steps (Fig. 5G).

Performance of hub genes to diagnose PCOS in the training and validation sets

In the training set (GSE95728-GC + GSE155489-GC), MAP1LC3A was significantly overexpressed in PCOS compared with the control ($P < 0.01$, Fig. 6A). The AUROC of MAP1LC3A was 0.860 (95% CI 0.692–1.000), with a sensitivity of 90.9% and a specificity of 81.8% (Fig. 6B). Notably, the AUROC of TOMM5 was 1.000 (95% CI 1.000–1.000), with a specificity of 100.0% and a sensitivity of 100.0% (Fig. S1). The small sample size included in our study may account for this distortion, so we decided that TOMM5 was not suitable for further validation and generalization.

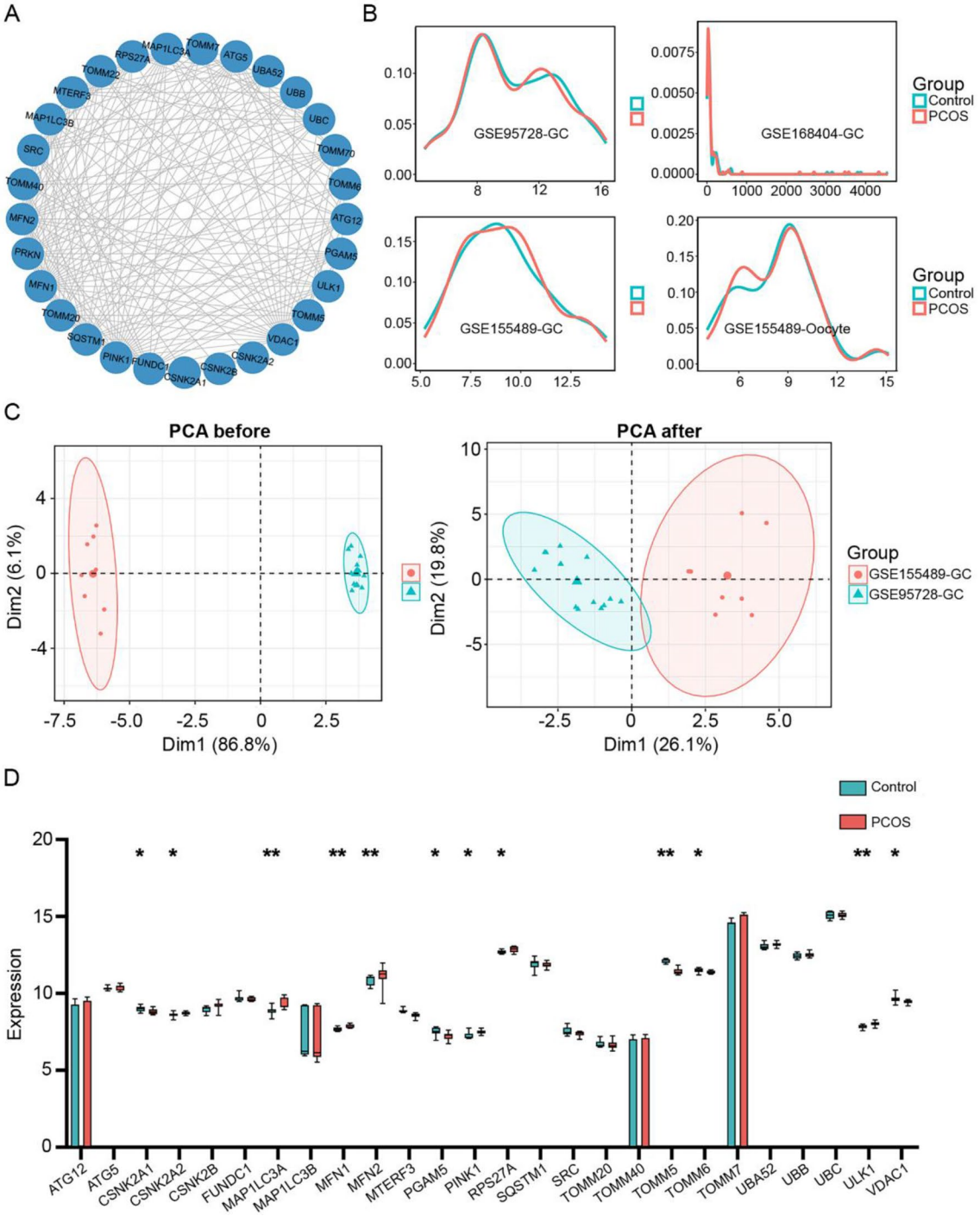


Figure 2. Landscape of mitophagy-related genes in PCOS. (**A**) Protein-protein interaction (PPI) network made up of 29 mitophagy-related genes. (**B**) The density distribution plots of the GSE95728, GSE168404 and GSE155489 datasets in granulosa cell or oocyte samples. (**C**) The PCA plots of the GSE95728 and GSE155489 datasets before and after sample correction in granulosa cell samples. (**D**) Box plots showing the expression levels of 26 mitophagy-related genes in granulosa cell samples of PCOS and the matched control. Differences between groups are represented by "*". *P < 0.05; **P < 0.01. GC, granulosa cell. Data were analyzed by wilcoxon tests.

In the validation set (GSE168404-GC), (Fig. 6C,D) showed the value of MAP1LC3A in the diagnosis of PCOS. The expression of MAP1LC3A was also significantly higher in PCOS groups than in controls ($P < 0.05$, Fig. 6C). The ROC curve demonstrated that MAP1LC3A performed exceptionally well in diagnosing PCOS, with the AUROC of 0.960 (Fig. 6D). Similarly, in another validation set (GSE155489-Oocyte), MAP1LC3A exhibited the excellent diagnostic value with the AUROC of 0.944, as shown in (Fig. 6E–G).

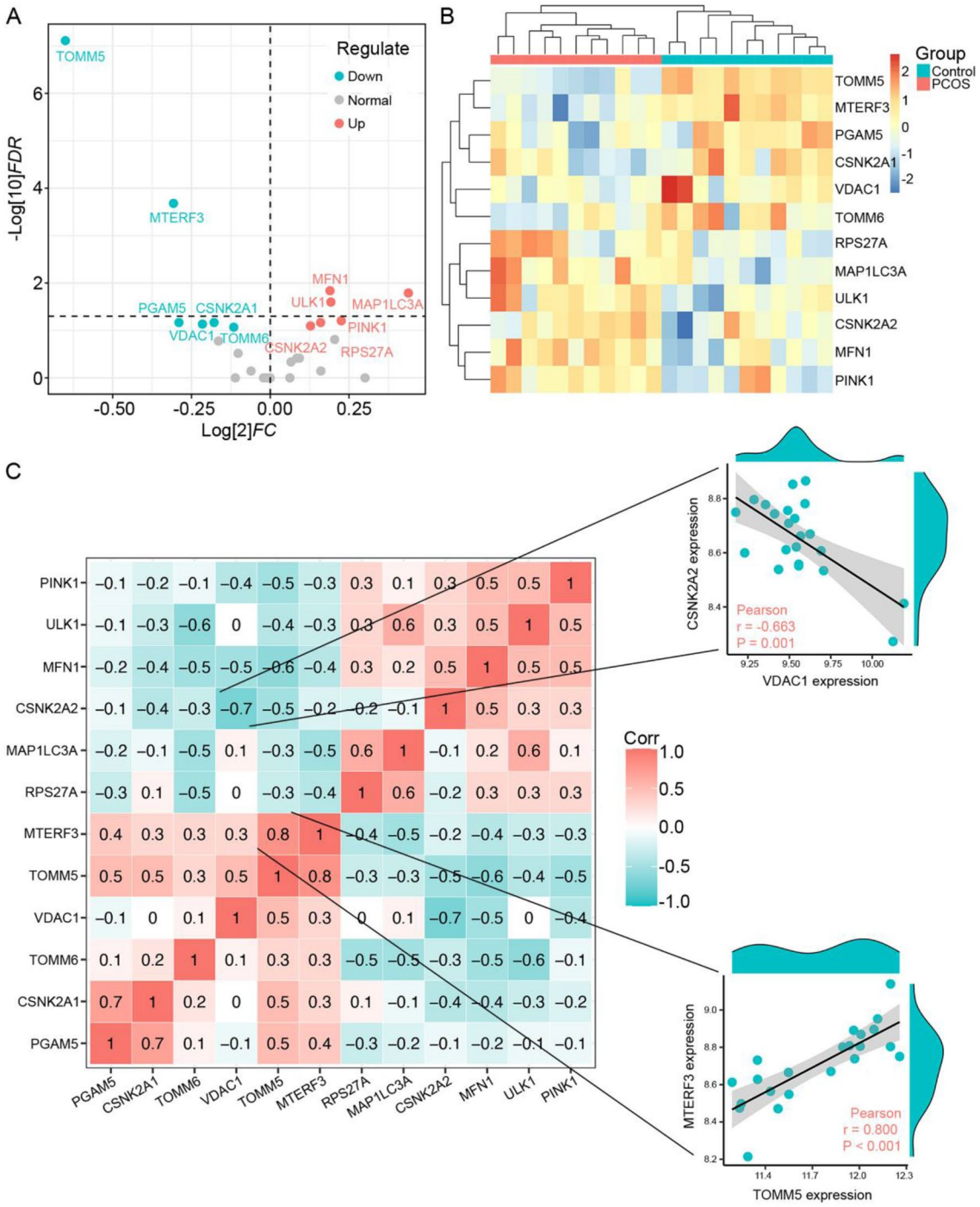


Figure 3. Variance analysis of mitophagy-related genes in PCOS. (**A**) Volcano plot showing a summary of the expression differences of 12 mitophagy-related genes between control and PCOS patients' granulosa cell samples. (**B**) The clustering heatmap exhibiting the expression pattern of 12 PCOS-related DE-MRGs among granulosa cell samples. (**C**) Correlations between DE-MRGs in PCOS granulosa cell samples and the respective scatterplots showing the two pairs of MRGs with the highest correlation. Correlation analyses were assessed using Pearson correlation.

Analysis of immune infiltration

To investigate whether the expression levels of MRGs were related to immunity, the CIBERSORT algorithm was used to evaluate the immune infiltration of PCOS. The analysis of PCA clusters showed that there was a huge distinction between the PCOS and control samples for immune cell infiltration (Fig. 7A). Using the par function, the immune cell percentage was calculated and the stacked histogram was presented (Fig. 7B). Correlation heatmap drawn to assess the correlation among 22 immune cell infiltrations showed that M1 macrophages, CD4 memory resting T cells, and naive B cells had a significant positive relation. Moreover, M1 macrophages and

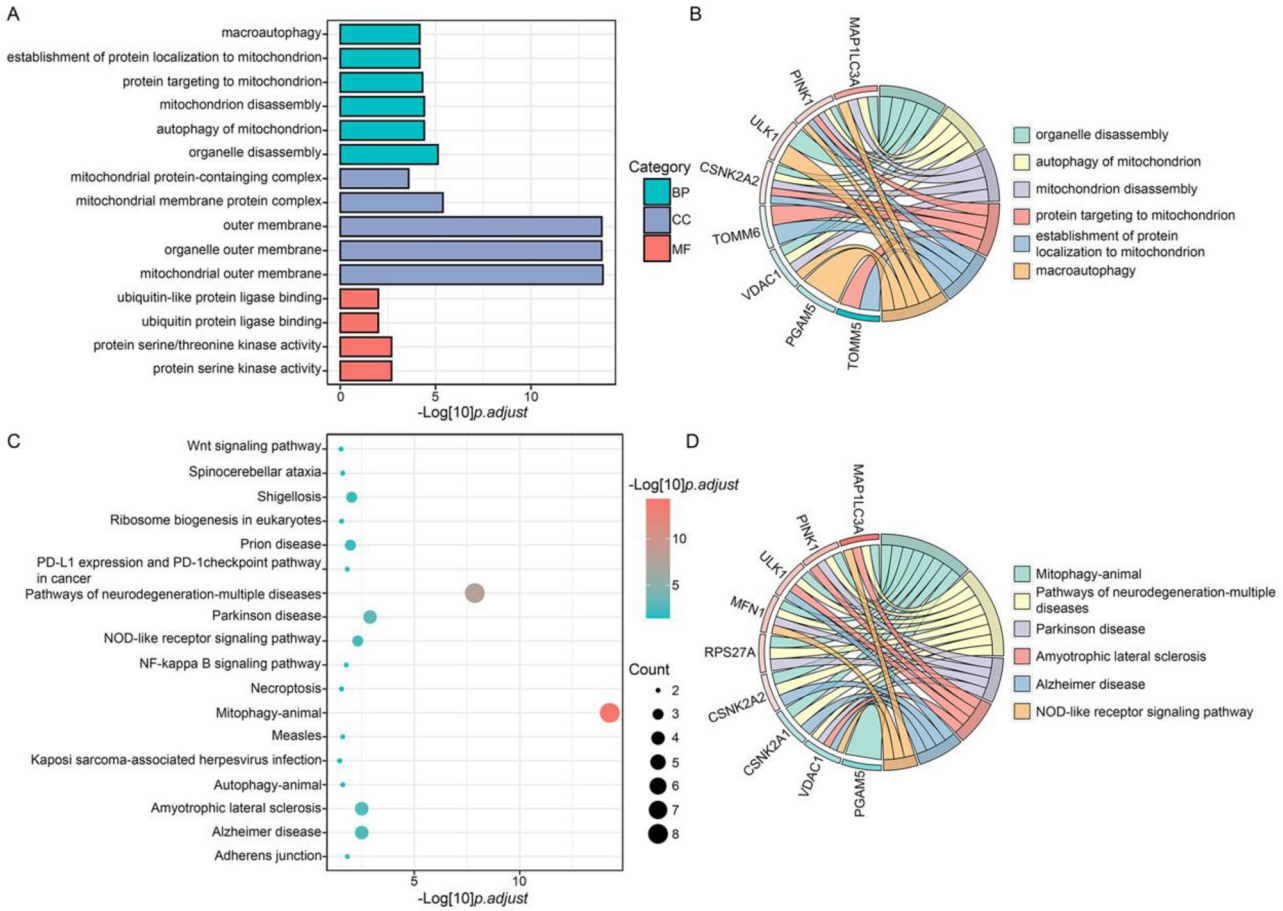


Figure 4. GO and KEGG analysis of 12 PCOS-related DE-MRGs. (**A**) Bar plot of enriched GO terms. (**B**) Chord diagram of enriched GO terms. (**C**) Bubble plot of enriched KEGG terms. (**D**) Chord diagram of enriched KEGG terms. *BP* biological process, *CC* cellular component, *MF* molecular function.

CD4 memory resting T cells also had a positive relation. Activated NK cells, gamma delta T cells, and follicular helper T cells had a significant positive relation. Moreover, activated NK cells and gamma delta T cells also had a positive relation. A positive correlation was also observed between M0 macrophages and memory B cells, and also between CD8 T cells and plasma cells, neutrophils and resting dendritic cells, respectively. CD4 naive T cells had a significant negative correlation with naive B cells (Fig. 7C). Figure 7D showed the difference among 22 immune cell infiltrations, plasma cells in PCOS had a high infiltration compared with control sample. Additionally, it was discovered that the expression of MAP1LC3A was positively related to monocytes ($r = 0.615$, $P = 0.033$) (Fig. 7E).

Clinical correlation of MAP1LC3A with testosterone levels

To further illustrate the status of MRGs in PCOS, correlation analysis between MAP1LC3A and testosterone levels was conducted. MAP1LC3A was positively related to testosterone levels ($r = 0.795$, $P = 0.006$) (Fig. 8), revealing that MAP1LC3A may exert an effect on ovulation disorders in PCOS.

Discussion

Polycystic ovarian syndrome is a common endocrine and metabolic syndrome that accounting for 75% of cases of anovulatory infertility[28]. The development of hyperandrogenemia is the characteristic biochemical feature of the disease and the primary reason behind the majority of PCOS clinical symptoms[29]. Recently, evidence has shown that androgens could impact the cellular metabolic pathways, potentially leading to risks within the mitochondria[30]. This indicates that the mitochondrial dysfunction of follicular cells (granulosa cell and oocyte) caused by hyperandrogenemia may partly account for the PCOS ovulation disorders. Therefore, we proposed a novel proposal for finding potential biomarkers possessing high specificity and sensitivity, capable of delineating the extent of mitochondrial quality control in follicular cells to better understand PCOS pathogenesis.

Mitophagy, which is a form of selective autophagy within mitochondria, serves as a crucial mechanism for maintaining cellular mitochondrial quality and therefore is crucial for sustaining energy production and responding to energy stress. When mitophagy is overstimulated under certain stressful conditions, the essential components for cell survival can be digested and lead to cell dysfunction[31]. So far, a series of studies have shown that excessive mitophagy contributes to the advancement of PCOS. One study observed the autophagy activation in the ovarian tissues of both PCOS individuals and PCOS-liked rats[32]. Furthermore, there was an observed

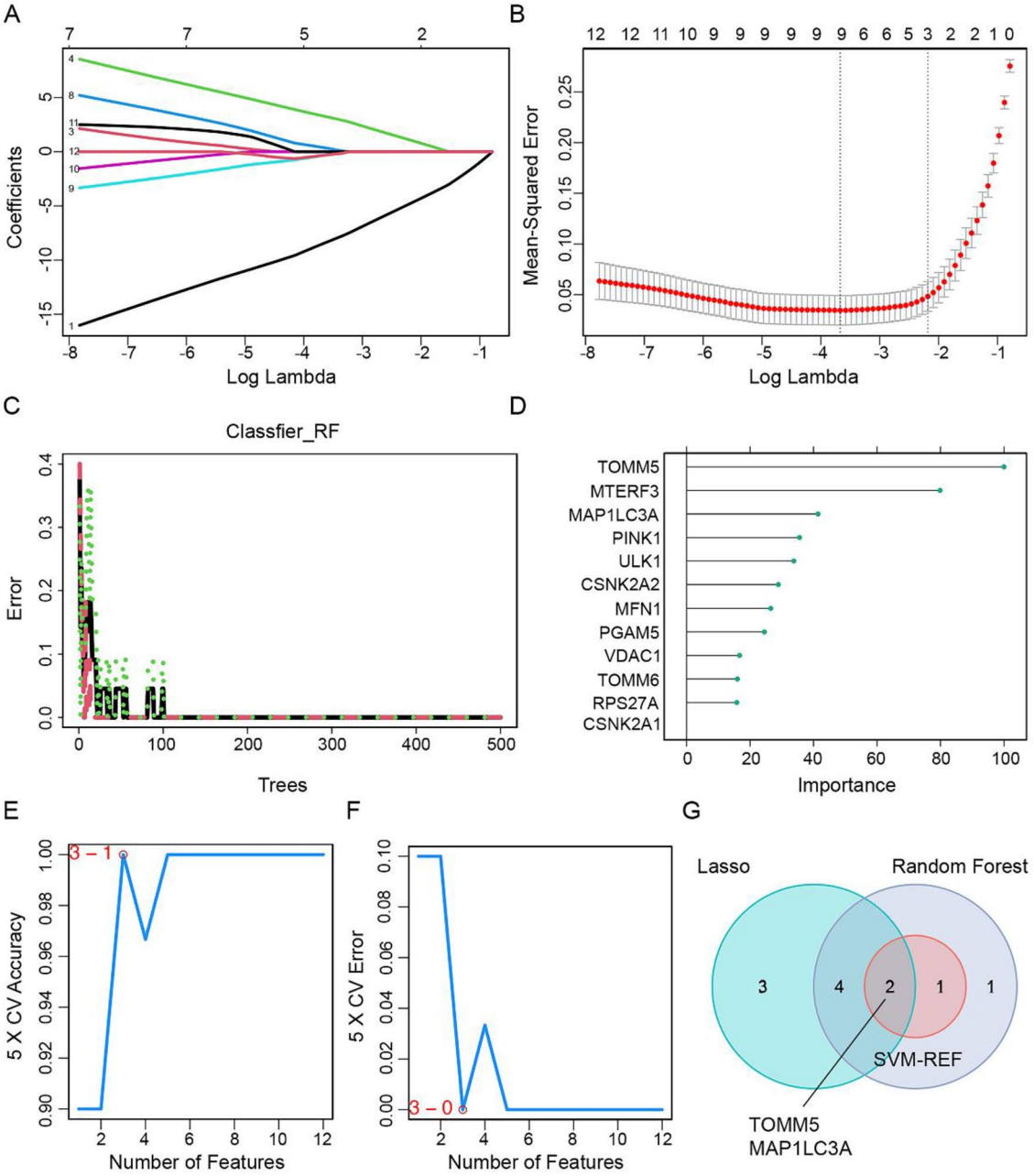
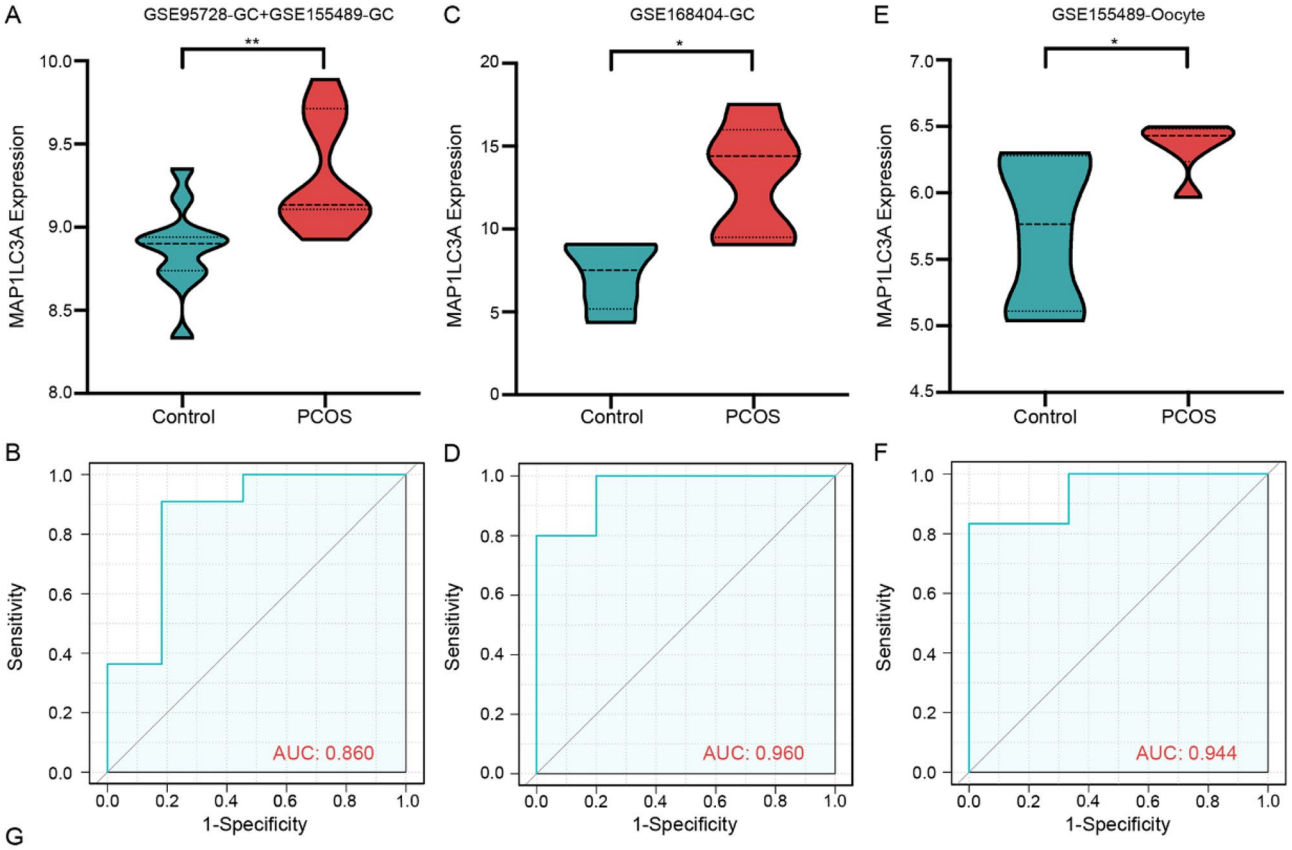


Figure 5. 2 DE-MRGs were identified as potential marker genes for PCOS. (**A**,**B**) Regression coefficient path plot and cross-validation curves in LASSO regression algorithm. (**C**,**D**) The identification of feature importance based on RF algorithm. (**E**,**F**) The curve of change in the true and error value of each gene prediction in SVM-RFE algorithm. (**G**) Venn diagram showing the intersection of selected markers obtained from the three algorithms.

elevation in mitophagy and a higher presence of injured mitochondria in the cumulus cells of individuals with PCOS. Yi et al. also proposed that the granulosa cells of PCOS patients experienced mitochondrial injure due to the excessive activation of PINK1/Parkin-mediated mitophagy[6]. Therefore, it is necessary to speak out here that our finding is consistent with those previous study conclusions. We found an obvious up-regulation of MAP1LC3A (a kind of MRGs) expression in PCOS individuals. Moreover, the positive correlation between MAP1LC3A and testosterone levels in PCOS patients supports the notion that mitophagy contributes significantly to the manifestations of PCOS[4].

The application of machine learning algorithms in creating decision models that support disease diagnosis and treatment is growing[33]. A total of 2 differential MRGs were identified in our study, namely TOMM5 and MAP1LC3A. In the merged training set (GSE95728-GC + GSE155489-GC), the AUROC value of TOMM5 was 1.0, with a specificity of 100.0% and a sensitivity of 100.0%. We speculated that the model distortion may be due to the limitations of sample size, making it unsuitable for further validation and generalization. MAP1LC3A,



G

dataset	sample	marker	AUROC (95%CI)	Cut-off	Sensitivity	Specificity	PPV	NPV	P Value
GSE95728+ GSE155489	Granulosa Cells_training	MAP1LC3A	0.860 (0.692-1)	8.995	0.909	0.818	0.833	0.900	<0.001
GSE168404	Granulosa Cells_validation	MAP1LC3A	0.960 (0.849-1)	9.060	1.000	0.800	0.833	1.000	<0.001
GSE155489	Oocytes_ validation	MAP1LC3A	0.944 (0.816-1)	6.324	0.833	1.000	1.000	0.857	<0.001

Figure 6. The performance of MAP1LC3A to discriminant PCOS in the training and validation set. (**A**,**C**,**E**) Expression difference of MAP1LC3A in PCOS and control groups. (**B**,**D**,**F**) The ROC curve of MAP1LC3A in PCOS and control groups. (**G**) Diagnostic values of MAP1LC3A for differentiating PCOS from control groups. *GC* granulosa cell, *PPV* positive predictive value, *NPV* negative predictive value, *AUROC* area under the receiver operating characteristics curve, *CI* confidence interval. Data were analyzed by Wilcoxon tests or Student's *t*-tests.

on the other hand, of which the AUROC values were all greater than 0.8 both in the merged training set and validation sets (GSE168404-GC and GSE155489-Oocyte), indicating that MAP1LC3A have the certain accuracy and specificity for distinguishing PCOS from the matched controls. Additionally, it is worth emphasizing that we selected 2 types of follicular cells (granulosa cell and oocyte) for modeling and validation, aiming to compensate for the limitations of the sample size. The quality of oocyte can directly characterize the follicular microenvironment, thereby predicting the ovulation ability of PCOS patients. And the granulosa cells specifically located around the oocyte play an important role in oocyte maturation and ovulation[6]. Abnormal granulosa cell function may indirectly affect follicular development and alter many symptoms of PCOS[34].

Growing proof suggests that the self-clearance of malfunctioning mitochondria is an effective strategy to keep the immune system in check. Mitophagy restricts the secretion of inflammatory cytokines and directly regulates mitochondrial antigen presentation, thereby maintaining the immune cell homeostasis[35]. Moreover, by regulating the adaptive immune response of memory NK cells, CD8 T cells, and dendritic cell-T cell synapses, mitophagy can shield cells against chronic inflammation[36]. Currently, several studies have jointly demonstrated the link between PCOS and low-grade chronic inflammation[37]. Furthermore, the persistent presence of inflammation in PCOS can exacerbate the obstruction of energy supply to oocytes, resulting in ovum quality impairment and subsequently impacting ovulation[38]. In our study, the immune infiltration analysis displayed a positive correlation between the presence of monocytes and MAP1LC3A levels. Based on the consistent trend of changes in MAP1LC3A and testosterone levels mentioned above, this suggested that immune cells and cytokines interact with androgens may result in the disruption of ovarian immune balance in PCOS. As González's findings suggested, the infiltration of monocytes into the ovary could potentially initiate a localized inflammatory response, leading to the stimulation of ovarian androgen synthesis in PCOS women[39]. In addition, the CIBERSORT analysis revealed an increased infiltration of plasma cells in PCOS. In Ewa Rudnicka's review, it

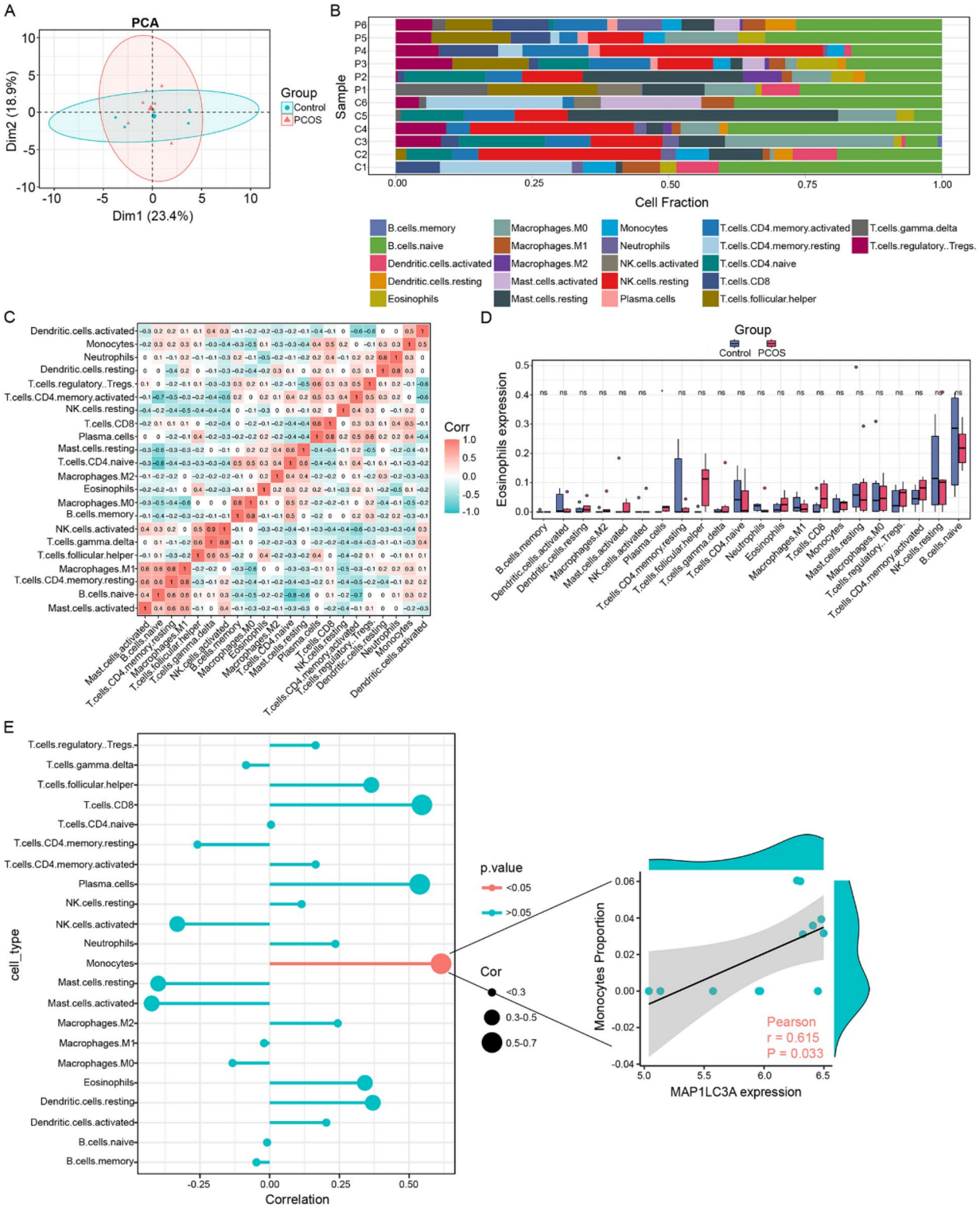


Figure 7. Evaluation and visualization of immune cell infiltration. (**A**) The PCA plot showing immune cell infiltration between PCOS and control samples. (**B**) Stacked histogram comparing PCOS and control samples for the immune cell proportion. (**C**) Correlation heatmap of 22 types of immune cells. (**D**) Boxplots showing 22 types of immune cells in proportion. (**E**) Lollipop diagram showing the correlation between MAP1LC3A and infiltrating immune cells. Scatter diagram indicating the correlation between MAP1LC3A expression and Monocytes. Data were analyzed by Wilcoxon tests; Correlation analyses were assessed using Pearson correlation.

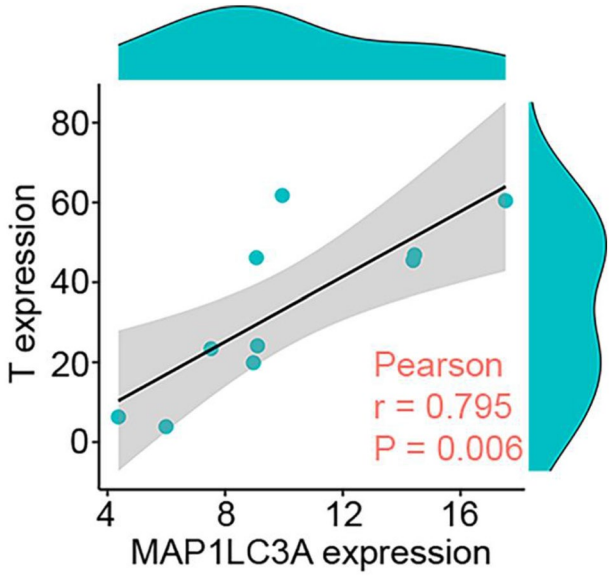


Figure 8. Scatter diagram indicating the relationship between MAP1LC3A expression and testosterone levels. Correlation analysis was assessed using Pearson correlation.

was also noted that PCOS women display higher serum concentration of TNF and C-reactive protein (CRP) as well as monocyte and lymphocyte circulating levels[37]. To summarize, infiltrating immune cells contributes to the initiation and advancement of PCOS, and targeting MAP1LC3A may help correct this aberrant immunological status in the coming times.

In our study, the MRGs were acquired from the Reactome database, an up-and-coming resource that has been extensively utilized by numerous studies[40]. However, several specific mitophagy receptors, including BNIP3, p62, OPTN, etc., were absent from it. Thus, it may be preferable to combine the Reactome database with other databases such as KEGG to obtain more thorough MRGs for future studies. Furthermore, the absence of experimental verification for the samples poses a constraint on our study. The limited sample size of the three datasets in our study necessitates selecting more datasets and confirming our findings in a larger PCOS cohort.

Conclusions

Our study identified the mitophagy-related gene MAP1LC3A as a promising biomarker in PCOS. Additionally, we discussed the possible correlation between MAP1LC3A and infiltrating immune cells, shedding new perspective on its significant contribution to the progress of PCOS. This offers a new insight into the prevention and treatment of PCOS.

Data availability

The data presented in this study are publicly available in the NCBI Gene Expression Omnibus (GEO) repository (GSE95728, GSE168404 and GSE155489).

Received: 17 March 2024; Accepted: 18 July 2024
Published online: 23 July 2024

References

1. Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil. Steril.* **81**, 19–25 (2004).
2. Kumariya, S., Ubba, V., Jha, R. K. & Gayen, J. R. Autophagy in ovary and polycystic ovary syndrome: Role, dispute and future perspective. *Autophagy* **17**, 2706–2733 (2021).
3. Kim, I., Rodriguez-Enriquez, S. & Lemasters, J. J. Selective degradation of mitochondria by mitophagy. *Arch. Biochem. Biophys.* **462**, 245–253 (2007).
4. Zhang, J., Bao, Y., Zhou, X. & Zheng, L. Polycystic ovary syndrome and mitochondrial dysfunction. *Reprod. Biol. Endocrinol.* **17**, 67 (2019).
5. Nguyen, T. N., Padman, B. S. & Lazarou, M. Deciphering the molecular signals of PINK1/Parkin mitophagy. *Trends Cell Biol.* **26**, 733–744 (2016).
6. Yi, S. *et al.* Melatonin ameliorates excessive PINK1/Parkin-mediated mitophagy by enhancing SIRT1 expression in granulosa cells of PCOS. *Am. J. Physiol. Endocrinol. Metab.* **319**, E91-101 (2020).
7. Pei, C. Z., Jin, L. & Baek, K. H. Pathogenetic analysis of polycystic ovary syndrome from the perspective of omics. *Biomed. Pharmacother.* **142**, 112031 (2021).
8. Sagvekar, P., Kumar, P., Mangoli, V., Desai, S. & Mukherjee, S. DNA methylome profiling of granulosa cells reveals altered methylation in genes regulating vital ovarian functions in polycystic ovary syndrome. *Clin. Epigenet.* **11**, 61 (2019).
9. Jiang, D. *et al.* Identification and validation of senescence-related genes in polycystic ovary syndrome. *J. Ovarian Res.* **17**, 7 (2024).

10. Huang, J. *et al.* Polycystic ovary syndrome: Identification of novel and hub biomarkers in the autophagy-associated mRNA-miRNA-lncRNA network. *Front. Endocrinol.* **13**, 1032064 (2022).
11. Fabris, F., Palmer, D., de Magalhães, J. P. & Freitas, A. A. Comparing enrichment analysis and machine learning for identifying gene properties that discriminate between gene classes. *Br. Bioinform.* **21**, 803–814 (2020).
12. Barrett, T. *et al.* NCBI GEO: Archive for functional genomics data sets–update. *Nucleic Acids Res.* **41**, D991-995 (2013).
13. Liu, Y.-D. *et al.* Long noncoding RNAs: Potential regulators involved in the pathogenesis of polycystic ovary syndrome. *Endocrinology* **158**, 3890–3899 (2017).
14. Zhao, R., Jiang, Y., Zhao, S. & Zhao, H. Multiomics analysis reveals molecular abnormalities in granulosa cells of women with polycystic ovary syndrome. *Front. Genet.* **12**, 648701 (2021).
15. Li, J. *et al.* Molecular features of polycystic ovary syndrome revealed by transcriptome analysis of oocytes and cumulus cells. *Front. Cell Dev. Biol.* **9**, 735684 (2021).
16. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).
17. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
18. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).
19. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. *Gene Ontol. Consort. Nat. Genet.* **25**, 25–29 (2000).
20. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
21. Engebretsen, S. & Bohlin, J. Statistical predictions with glmnet. *Clin. Epigenet.* **11**, 123 (2019).
22. Sundermann, B. *et al.* Support vector machine analysis of functional magnetic resonance imaging of interoception does not reliably predict individual outcomes of cognitive behavioral therapy in panic disorder with agoraphobia. *Front. Psychiatr.* **8**, 99 (2017).
23. Zhao, E., Xie, H. & Zhang, Y. Predicting diagnostic gene biomarkers associated with immune infiltration in patients with acute myocardial infarction. *Front. Cardiovasc. Med.* **7**, 586871 (2020).
24. Robin, X. *et al.* pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
25. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).
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. Friendly, M. Corrgrams: Exploratory displays for correlation matrices. *Am. Stat.* **56**(4), 316–324. https://doi.org/10.1198/000313002533 (2002).
28. Ibáñez, L. *et al.* An International consortium update: Pathophysiology, diagnosis, and treatment of polycystic ovarian syndrome in adolescence. *Horm. Res. Paediatr.* **88**, 371–395 (2017).
29. Diamanti-Kandarakis, E., Papailiou, J. & Palimeri, S. Hyperandrogenemia: Pathophysiology and its role in ovulatory dysfunction in PCOS. *Pediatr. Endocrinol. Rev.* **3**(1), 198–204 (2006).
30. Chappell, N. R., Gibbons, W. E. & Blesson, C. S. Pathology of hyperandrogenemia in the oocyte of polycystic ovary syndrome. *Steroids* **180**, 108989 (2022).
31. Levine, B. & Kroemer, G. Autophagy in the pathogenesis of disease. *Cell* **132**, 27–42 (2008).
32. Li, D. *et al.* Autophagy is activated in the ovarian tissue of polycystic ovary syndrome. *Reproduction* **155**, 85–92 (2018).
33. Yang, Y. *et al.* Biomarkers prediction and immune landscape in ulcerative colitis: Findings based on bioinformatics and machine learning. *Comput. Biol. Med.* **168**, 107778 (2024).
34. Chen, M. *et al.* Resveratrol ameliorates polycystic ovary syndrome via transzonal projections within oocyte-granulosa cell communication. *Theranostics*. **12**, 782–95 (2022).
35. Xu, Y., Shen, J. & Ran, Z. Emerging views of mitophagy in immunity and autoimmune diseases. *Autophagy* **16**, 3–17 (2020).
36. Moossavi, M., Parsamanesh, N., Bahrami, A., Atkin, S. L. & Sahebkar, A. Role of the NLRP3 inflammasome in cancer. *Mol. Cancer* **17**, 158 (2018).
37. Rudnicka, E. *et al.* Chronic low grade inflammation in pathogenesis of PCOS. *Int. J. Mol. Sci.* **22**, 3789 (2021).
38. Repaci, A., Gambineri, A. & Pasquali, R. The role of low-grade inflammation in the polycystic ovary syndrome. *Mol. Cell Endocrinol.* **335**, 30–41 (2011).
39. González, F. Inflammation in polycystic ovary syndrome: underpinning of insulin resistance and ovarian dysfunction. *Steroids* **77**, 300–305 (2012).
40. Chen, H., Zhang, J., Sun, X., Wang, Y. & Qian, Y. Mitophagy-mediated molecular subtypes depict the hallmarks of the tumour metabolism and guide precision chemotherapy in pancreatic adenocarcinoma. *Front. Cell Dev. Biol.* **10**, 901207 (2022).

Acknowledgements

We would like to thank the GEO Database and the Reactome database for providing us with a shareable dataset.

Author contributions

YZ Yang and XJ Chen are the co-lead authors of this work. YZ Yang performed the experiments and collected the data, analyzed and interpreted the data, and draft the manuscript. XJ Chen performed the data analysis and interpreted the data. XH Liao, WW Jiang and Y Zhou provided critical comments and revised the manuscript. BH Zheng and Y Sun helped critically revised and approved the final manuscript. All authors contributed to the article and approved the final version of the manuscript.

Funding

Grants were received from Natural Science Foundation of Fujian Province, China (Grant Number 2023J011229), Natural Science Foundation of Fujian Province, China (Grant Number 2023J05232), Science and technology innovation start-up fund, Fujian Maternity and Child health Hospital (YCXY 23-03), the Joint funds for the innovation of science and technology, Fujian province (Grant No. 2020Y9137), Major Scientific Research Program for Young and Middle-aged Health Professionals of Fujian Province, China (Grant No. 2022ZQNZD010), Key Project on the Integration of Industry, Education and Research Collaborative Innovation of Fujian Province (Grant No. 2021YZ034011) and Key Project on Science and Technology Program of Fujian Health Commission (Grant No. 2021ZD01002).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-67969-9.

Correspondence and requests for materials should be addressed to Y.S. or B.Z.

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.