

· 临床研究 ·

基于焦亡相关基因构建多囊卵巢综合征诊断模型

陈 讯^{1,2}, 陈雯昕¹, 张 文^{1,2}, 俞池园¹, 许波群^{1*}

¹南京医科大学第二附属医院妇产科, 江苏 南京 210011; ²南京医科大学附属逸夫医院妇产科, 江苏 南京 211100

[摘要] 目的: 探索焦亡相关基因在多囊卵巢综合征(polycystic ovary syndrome, PCOS)发病机制中的作用, 并构建PCOS的精确预测模型。方法: 利用从基因表达公共数据库(Gene Expression Omnibus, GEO)中获取的3个微小RNA(microRNA, mRNA)表达谱, 分析PCOS患者与正常健康女性之间焦亡相关基因(pyroptosis-related gene, PRG)的表达差异。采用广义线性模型(generalized linear model, GLM)、随机森林(random forest, RF)、支持向量机(support vector machine, SVM)和极限梯度提升(extreme gradient boosting, XGB)这4种机器学习算法来识别疾病特征基因。采用实时定量聚合酶链反应(real-time quantitative polymerase chain reaction, RT-qPCR)法检测10例PCOS患者和10例正常健康女性血浆中特征基因的表达量。结果: 建立了基于PRG的PCOS预测模型和列线图。XGB显示出最高的准确性, 决策曲线分析进一步支持了这一结果。一致聚类显示PCOS病例中有两个亚组, 组2比组1表现出更多的免疫浸润。差异表达分析鉴定两个亚型之间的差异表达基因, 并对基因进行富集分析。临床验证结果显示, 含CARD结构域的凋亡相关斑点样蛋白(apoptosis associated speck like protein containing a CARD, 又称PYD and CARD domain containing, PYCARD)、黑素瘤缺乏因子2(absent in melanoma 2, AIM2)、染色质修饰蛋白4B(chromatin modifying protein 4B, CHMP4B)和NOD样受体蛋白2(NOD-like receptor family pyrin domain containing 2, NLRP2)在PCOS患者组中的表达量明显高于正常对照组, 差异有统计学意义, 验证了基于PRG的PCOS预测模型的准确性。结论: 本研究为PCOS与焦亡之间的关系提供了初步见解, 并提出了PCOS的精确预测模型。

[关键词] 多囊卵巢综合征; 细胞焦亡; 免疫浸润; 机器学习; 生物信息学分析

[中图分类号] R711.75

[文献标志码] A

[文章编号] 1007-4368(2025)04-509-14

doi: 10.7655/NYDXBNSN241034

A diagnostic model of polycystic ovary syndrome based on pyroptosis-related genes

CHEN Xun^{1,2}, CHEN Wenxin¹, ZHANG Wen^{1,2}, YU Chiyuan¹, XU Boqun^{1*}

¹Department of Obstetrics and Gynecology, the Second Affiliated Hospital of Nanjing Medical University, Nanjing 210011; ²Department of Obstetrics and Gynecology, the Affiliated Sir Run Run Hospital of Nanjing Medical University, Nanjing 211100, China

[Abstract] **Objective:** To explore the role of genes related to pyroptosis in the pathogenesis of polycystic ovary syndrome (PCOS) and construct an accurate prediction model for PCOS. **Methods:** The differential expression of pyroptosis-related genes (PRG) between PCOS patients and normal healthy women was analyzed by using three microRNA (mRNA) expression profiles obtained from the Gene Expression Omnibus (GEO) database. Four machine learning algorithms, namely the generalized linear model (GLM), random forest (RF), support vector machine (SVM), and extreme gradient boosting (XGB), were employed to identify the gene characteristics of PCOS. Real-time quantitative PCR (RT-qPCR) method was utilized to detect the expression levels of specific genes in the plasma of 10 PCOS patients and 10 normal healthy women. **Results:** A predictive model and a nomogram were established based on PRG to accurately predict PCOS. Among the four machine learning algorithms, the XGB method demonstrated the highest accuracy in validating the performance of the model, which was further supported by decision curve analysis. Consensus clustering revealed two distinct subgroups within PCOS cases, with Cluster 2 exhibiting higher level of immune infiltration compared with Cluster 1. Differential expression analysis was then conducted to identify differentially expressed genes between the two subtypes, followed by pathway enrichment analysis on the genes. Clinical verification showed that the plasma expression levels of the apoptosis associated

[基金项目] 国家自然科学基金(81873820); 江苏省卫健委医学科科研项目(H2023113)

*通信作者(Corresponding author), E-mail: boqun_xu@njmu.edu.cn (ORCID: 0000-0002-3360-528X)

speck like protein containing a CARD(also named PYD and CARD domain containing, PYCARD), the absent in melanoma 2(AIM2), the chromatin modif-yingprotein 4B(CHMP4B) and the NOD-like receptor family pyrin domain containing 2(NLRP2) were significantly higher in the PCOS patients than the healthy controls. This verifies the accuracy of the PCOS prediction model based on PRG. **Conclusion:** This study may offer preliminary insights into the correlation between PCOS and pyroptosis, and provide a precise predictive model for PCOS.

[Key words] polycystic ovary syndrome; pyroptosis; immune infiltration; machine learning; bioinformatics analysis

[J Nanjing Med Univ, 2025, 45(04): 509-522]

多囊卵巢综合征(polycystic ovary syndrome, PCOS)是最常见的内分泌和代谢疾病之一,影响6%~21%的育龄妇女^[1]。临床PCOS以月经稀发或无排卵、高雄激素为临床或生化表现,以卵巢多囊形态、胰岛素抵抗、肥胖、多毛、痤疮等为特征^[2]。PCOS的远期风险包括不孕、心血管疾病、2型糖尿病、血脂异常、代谢综合征和子宫内膜癌等^[3]。虽然对PCOS的研究已有80余年,但其患病率高,发病机制尚不明确,诊断也不统一。PCOS可导致青春期异常子宫出血、育龄期不孕、妊娠并发症以及绝经期心血管疾病、糖尿病、子宫内膜癌等风险增加^[4]。规范PCOS的诊断、治疗和管理非常重要。因此,迫切需要对PCOS的发病机制和早期诊断进行更深入地研究,以改善患者预后。

细胞焦亡又称细胞炎性坏死,在2001年首次被提出,是一种依赖于Caspase 1的炎症细胞死亡模式,与细胞凋亡完全不同^[5]。当细胞持续膨胀,直到细胞膜破裂,导致细胞内容物释放,激活强烈炎症反应时,就会发生这种情况。细胞焦亡广泛参与感染性疾病、神经系统相关疾病及动脉粥样硬化性疾病的发生发展,对细胞焦亡的深入研究有助于为临床防治提供新思路^[6-7]。

研究表明,在PCOS患者中观察到慢性炎症^[8]。一项针对PCOS患者的调查显示,与年龄匹配的非PCOS患者相比,PCOS患者的C反应蛋白水平明显较高^[9]。此外,炎症因子和氧化应激指标在PCOS患者中高表达,而之前研究表明,细胞凋亡不会引发促炎反应,这表明其中存在不同于卵巢细胞凋亡的另一种机制^[10-12]。相反,炎症状态和不同的程序性细胞死亡途径之间有很强的相关性,在细胞焦亡中特别突出^[13]。因此,探讨PCOS中细胞焦亡相关信号通路很有必要。最近研究也表明,PCOS小鼠中的高雄激素血症(hyperandrogenemia, HA)通过激活NOD样受体蛋白3(NOD-like receptor family pyrin domain containing 3, NLRP3)炎症小体引起卵巢焦亡,最终导致卵泡功能障碍^[14]。组蛋白脱乙酰酶1

(histone deacetylase 1, HDAC1)通过组蛋白修饰调控H19/miR-29a-3p/NLRP3介导的颗粒细胞焦亡参与PCOS的发生^[15]。鉴于PCOS患者炎症状态升高,焦亡细胞死亡途径可能在卵巢中过度激活。因此细胞焦亡如何参与PCOS发生和发展的详细调控模式需要进一步研究。

1 资料和方法

1.1 资料

从基因表达公共数据库(Gene Expression Omnibus, GEO)中获得GSE5850、GSE34526和GSE98421数据集的RNA测序(RNA-seq)数据,包括13例对照和17例PCOS患者。为保证数据的一致性,使用“limma”和“sva”R包对数值进行批量归一化处理。将3个数据集的表达谱进行合并。

1.2 方法

1.2.1 细胞焦亡相关基因(pyroptosis-related gene, PRG)差异表达分析

利用limma算法对先前研究发现的47个PRG进行了差异表达分析,旨在鉴定PCOS患者和对照组细胞焦亡相关的差异表达基因(differentially expressed gene, DEG)。基因集变异分析(gene set variation analysis, GSVA)对PCOS患者和对照组的焦亡相关DEG进行富集分析,以评估生物学过程改变和匹配生物学功能。

1.2.2 模型比较、选择和建立

采用广义线性模型(generalized linear model, GLM)、随机森林(random forest, RF)、支持向量机(support vector machine, SVM)和极限梯度提升(extreme gradient boosting, XGB)建立机器学习模型预测PCOS患病率。通过残差逆累积分布、残差箱形图、受试者工作特征(receiver operating characteristic, ROC)曲线分析等多种评估方法评估模型的准确性。在确定表现最佳的机器学习模型后,前5个关键变量被确定为与PCOS相关的预测基因。

1.2.3 列线图的建立

使用 R 包“rms”(version 6.3.0)构建列线图模型。对模型中的每个预测因子赋予相应评分,将所有预测因子的评分相加,确定“总分”。采用校准曲线评价列线图的准确性和可靠性。通过决策曲线分析(decision curve analysis, DCA)评估列线图在临床决策中的潜在益处。此外利用临床影响曲线来衡量使用列线图做出决策的合理性和优势。

1.2.4 一致性聚类分析

利用差异 PRG 的表达水平将样本分为不同的亚型。使用 R 软件包“Consensus Cluster Plus”对实验组样本进行一致性聚类分析。通过评估 K 值的累积分布函数,确定合适的聚类 K 值,确保聚类分析结果最可靠。对分类组进行免疫细胞浸润分析。利用 GSVA 评估不同代谢通路在未分类样本中的富集情况,该分析用“GSEABase”和“GSVA”两个数据包构建函数和设置参数^[16]。

1.2.5 免疫细胞相关表达分析

为检测不同组中免疫细胞的表达水平,使用专门用于肿瘤免疫浸润分析的 CIBERSORT 软件包。采用箱形图显示不同组间免疫细胞群的变化。

1.2.6 临床试验验证

通过收集 PCOS 患者组和正常对照组血浆(各 10 例),使用 Steadypure 血液血清血浆 Small RNA 提取试剂盒,提取血浆中总 RNA,将得到的总 RNA 通过逆转录反应并进行实时定量聚合酶链反应(real-time quantitative polymerase chain reaction, RT-qPCR)测定特征基因的表达量。反应引物序列见表 1。本研究获得南京医科大学附属逸夫医院伦理委员会批准(伦理审查编号为 2024-SR-04),参与者均签署知情同意书。

1.3 统计学方法

所有统计分析均采用 R 软件 v4.1.2 和 GraphPad Prism10 软件。使用 Wilcoxon 检验比较两组之间差

表 1 引物序列

Table 1 Primer sequences

Gene	Sequence(5'→3')
β-actin(mouse)	Forward: GGAAATCGTGCACATTAAG
	Reverse: CGGCAGTGGCCATCTCTT
Homo-AIM2	Forward: TGGCAAAACGTCTTCAGGAGG
	Reverse: AGCTTGACTTAGTGGCTTTGG
Homo-CHMP4B	Forward: TGCAGAGGAGATTTCAACAGC
	Reverse: TGTTCGGGTCCACTGATTTTC
Homo-NLRC4	Forward: TGCATCATTTGAAGGGGAATCTG
	Reverse: GATTGTGCCAGGTATATCCAGG
Homo-NLRP2	Forward: ACGGTGGTGTGTATGGTC
	Reverse: TCCTCTGCCAGTCTAGCATT
Homo-PYCARD	Forward: TGGATGCTCTGTACGGGAAG
	Reverse: CCAGGCTGGTGTGAAACTGAA

异。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 在 PCOS 中细胞 PRG 的特征

利用 GeneCard 网站收集 PRG 数据集。通过基因本体(gene ontology, GO)和系统分析基因功能、基因组信息数据库(Kyoto encyclopedia of genes and genomes, KEGG)富集分析探讨细胞 PRG 在 PCOS 中的潜在作用^[17]。GO 富集分析显示, PRG 在细胞焦亡、核膜重组、巨自噬相关生物学过程中富集(图 1A)。在 KEGG 富集分析中,焦亡基因在细胞质 DNA 传感通路、NOD 样受体信号通路和 p53 信号通路中尤为

丰富(图 1B)。PPI 分析显示 44 个焦亡基因存在相互作用, Cytoscape 分析显示 8 个核心基因:染色质修饰蛋白(chromatin modifying protein, CHMP) 2A、CHMP2B、CHMP3、CHMP4A、CHMP4B、CHMP4C、CHMP6 和 CHMP7(图 1C)。小提琴图显示 PCOS 患者组与对照组差异的 PRG,结果显示含 CARD 结构域的凋亡相关斑点样蛋白(apoptosis associated speck like protein containing a CARD, 又称 PYD and CARD domain containing, PYCARD)、黑素瘤缺乏因子 2(absent in melanoma 2, AIM2)和核苷酸结合寡聚化结构域蛋白 2(nucleotide-binding oligomerization domain-containing protein 2, NOD2)在 PCOS 患者

组中高表达,且差异有统计学意义(图1D)。

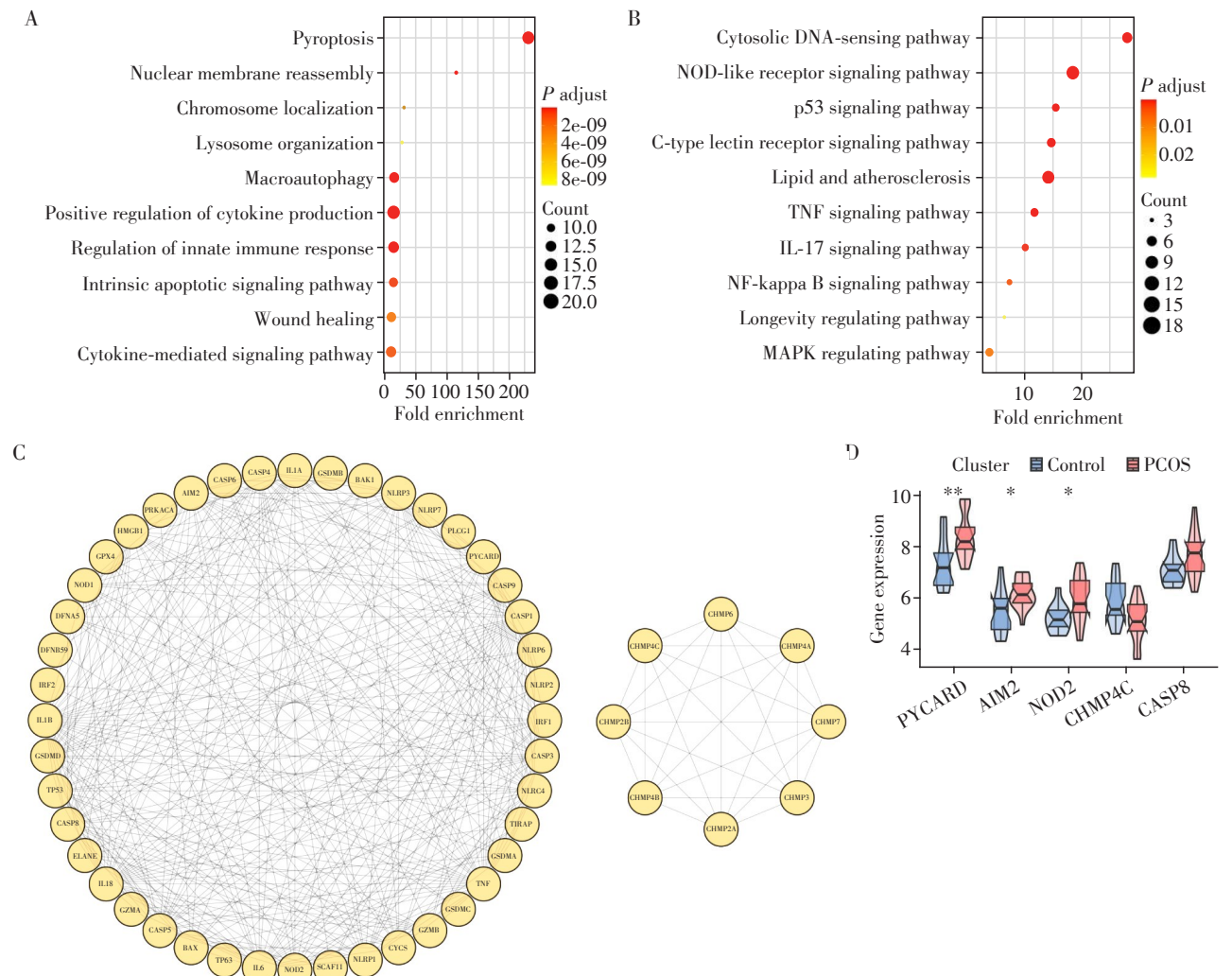
2.2 构建机器学习模型识别PCOS疾病特征基因

使用RF、XGB、GLM和SVM算法建立机器学习模型以识别PCOS疾病特征基因。基于47个关键PRG构建模型,评估PCOS发病风险。残差盒线图(图2A)显示,XGB残差均方根最小,而GLM残差值最大。RF和SVM模型的残差分布在0.25~0.50。反向累积分布图的结果(图2B)与这些发现一致。ROC曲线分析(图2C)显示,XGB模型的曲线下面积为0.867,SVM为0.800,RF为0.733,GLM为0.533。相比之下,GLM模型曲线面积较小,存在过拟合的可能。对所有4种方法进行了基因重要性分析,得到了基因重要性评分(图2D)。在GLM模型中,BCL-2基因家族中的细胞凋亡促进基因(BCL-2-associated X protein, BAX)、谷胱甘肽过氧化物酶4(glutathione

peroxidase 4, GPX4)、CHMP2A、半胱天冬酶8(caspase 8, CASP8)和CHMP4B是得分最高的前5个基因。对于RF模型,NLR家族含CARD结构域4(NLR family CARD domain containing 4, NLRC4)、AIM2、磷脂酶C, γ 1(phospholipase C, gamma 1, PLCG1)、NOD2和蛋白激酶C α (protein kinase C alpha, PRKACA)被认为是最重要的。在SVM模型中,PYCARD、CHMP2A、CASP8、NLRC4和NOD2是显著基因,而在XGB模型中,CHMP4B、PYCARD、NOD样受体蛋白2(NOD like receptor family pyrin domain containing 2, NLRP2)、NLRC4和AIM2是显著基因,准确率为0.867。结合结果,选择精度最高的XGB模型。

2.3 预测列线图的评估

以XGB模型筛选出的前5个基因为疾病特征基



A: GO function analysis of pyroptosis-related genes. B: KEGG pathway enrichment of pyroptosis-related genes. C: PPI network analysis of pyroptosis-related genes. D: Differential expression analysis of pyroptosis-related genes between control (n=13) and PCOS cases (n=17). *P < 0.05, **P < 0.01.

图1 PCOS中焦亡相关基因的特征

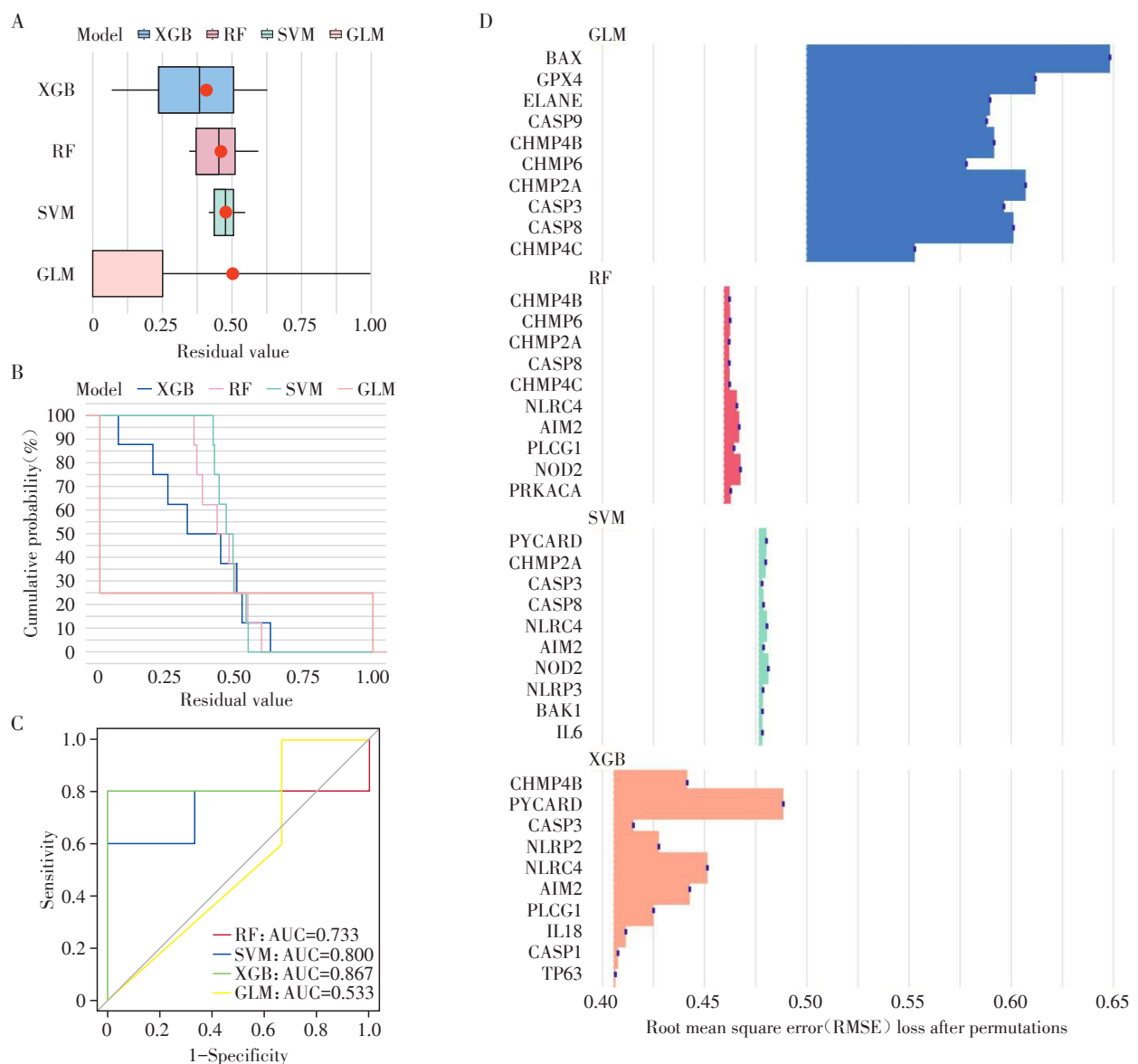
Figure 1 The characteristics of pyroptosis-related genes in PCOS

因,构建预测PCOS发病率的列线图(图3A)。每个特征基因分配一个特定的得分区间,将所有基因的得分汇总,计算最终得分,并与发病率进行比较。通过校准曲线评估列线图的预测准确性,实线和虚线的对齐越紧密,表明模型的准确性越高(图3B)。此外,DCA表明,黄线所代表的模型与所有曲线的距离显著较远,从而表明其在临床决策中的有效性(图3C)。

2.4 基于PRG的新亚型鉴定

使用47个差异表达的PRG,应用共识聚类算法将PCOS患者分为不同的亚型。当Clusters变量设

置为2时,患者被很好地分为两个Cluster(图4A)。从累积分布函数(cumulative distribution function, CDF)曲线在0~0.2波动可以看出,当K=2时,Clusters的稳定性最高(图4B)。另外,图4C说明了K值2~9时的CDF曲线下的面积。值得注意的是,当K=2时,各亚型的一致性得分最大且最接近。因此,评估了PRG在Cluster 1和Cluster 2之间的表达差异,以研究区分集群的分子特征。Cluster 1中CASP3、CASP6和NLRP2的表达水平较高,而Cluster 2中CHMP2A、CHMP4B、白介素(interleukin, IL)1B、CASP9、GPX4、IL6、NLRC4、NLRP1、NLRP3、PYCARD



A: The residual box line diagram of four machine learning algorithms, red dot stands for mean square of residuals. B: Reverse cumulative distribution of residual among different machine learning models. C: ROC curve to verify the accuracy of the model. D: Importance score of feature genes in the model.

图2 构建机器学习模型识别PCOS疾病特征基因

Figure 2 Building a machine learning model to identify the PCOS disease signature genes

和肿瘤坏死因子(tumor necrosis factor, TNF)的表达水平较高。相反,在各基因簇之间,BCL2-拮抗剂/杀伤因子1(BCL2-antagonist/killer 1, BAK1)、BAX、CASP1、环叶素依赖性激酶底物(cyclophlin-dependent kinase substrate, CYCS)和其他基因的表达未观察到显著差异(图4D、E)。GSVA分析图突出显示了亚型之间的差异通路(图4F)。在C2亚组中,丙酸盐代谢、细胞周期、泛酸盐和辅酶a生物合成、半胱氨酸和蛋氨酸代谢以及p53信号通路表现出正调控关系,而在C1亚组中,核糖体、细胞质DNA感应通路和造血细胞谱系表现出正调控关系。

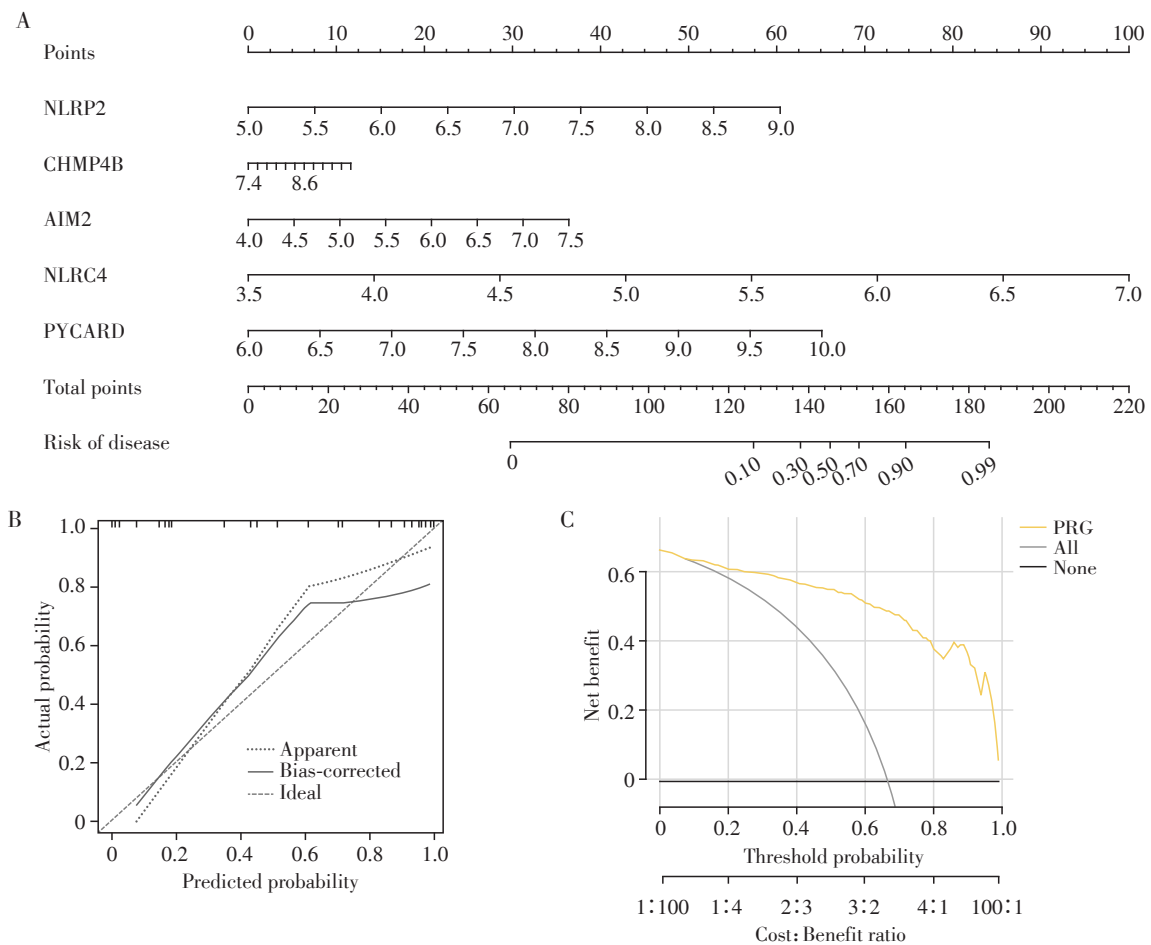
2.5 基于细胞PRG亚型的免疫状况

直方图分析显示不同组免疫细胞浸润模式不同。Cluster 1组的幼稚B细胞、记忆B细胞和中性粒细胞水平较高,而Cluster 2组的静息单核细胞和肥大细胞水平较高。其他免疫细胞含量组间差异相对较小(图5A),但具体原因尚需进一步实验验证。

免疫细胞差异分析图证实了这些发现,在不同亚组之间,幼稚B细胞、调节性T细胞、单核细胞、巨噬细胞M0、活化的肥大细胞和中性粒细胞差异有统计学意义($P < 0.05$,图5B)。此外,为了证实PRG与PCOS之间的相关性,对这两个亚型与细胞因子之间的相关性进行了深入探索。分析显示,细胞因子表达在基因簇之间存在显著差异(图5C)。值得注意的是,白细胞介素1受体拮抗因子(interleukin 1 receptor antagonist, IL1RN)、IL6、IL10、IL16、IL32、转化生长因子B1(transforming growth factor beta 1, TGFB1)、细胞因子信号通路抑制因子3(suppressor of cytokine signaling 3, SOCS3)、CXC基序趋化因子10(CXC chemokine ligand 10, CXCL10)和TNF在cluster 2中的表达水平高于cluster 1。这些发现进一步证实了程序性细胞死亡cluster 2与PCOS的密切相关性。

2.6 特征基因的功能富集分析

对模型基因通过基因集富集分析法(gene set

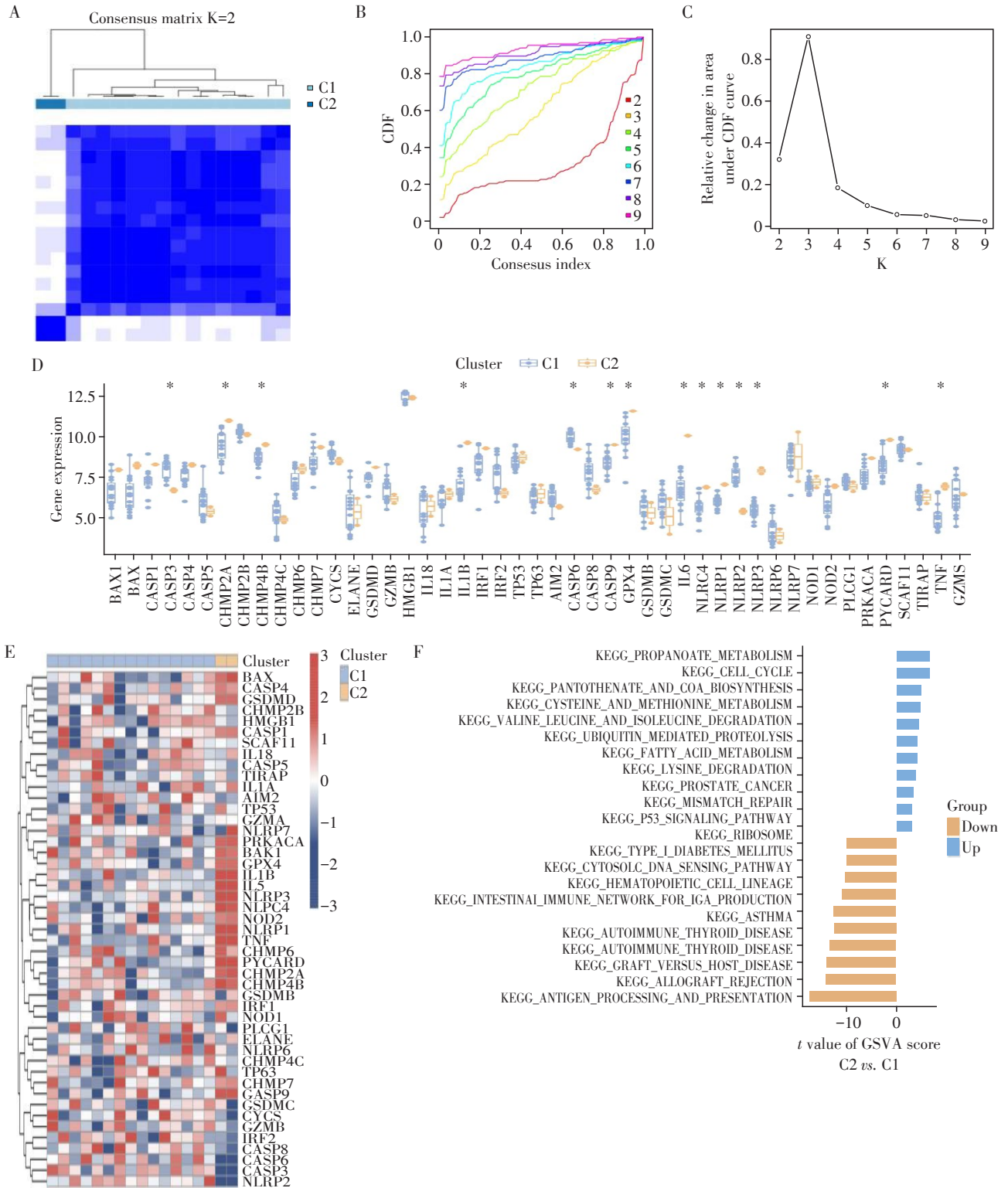


A: Nomogram showing prediction of PCOS prevalence using signature genes. B: Calibration curve illustrating the accuracy of nomogram. C: Decision curve illustrating the accuracy of the model.

图3 预测列线图的评价

Figure 3 Evaluation of a predictive nomogram

enrichment analysis, GSEA)进行通路富集分析。在 GO 功能分析中,CHMP4B 与防御反应、细胞外囊泡 和免疫反应生物学功能呈正相关(图 6A)。NLRC4 与防御反应、生物刺激反应、免疫反应呈负相关(图



A: Consensus matrix heat map defining two clusters(K=2) and their correlation area. B: Representative CDF curves and CDF delta area curves. C: The score of cluster-consensus. D: Boxplot showing the expression patterns of all pyroptosis-related genes between two clusters, *P < 0.05. E: Heatmap of expressions of all pyroptosis-related genes between two clusters. F: Gene set enrichment analysis(GSEA) between the two clusters.

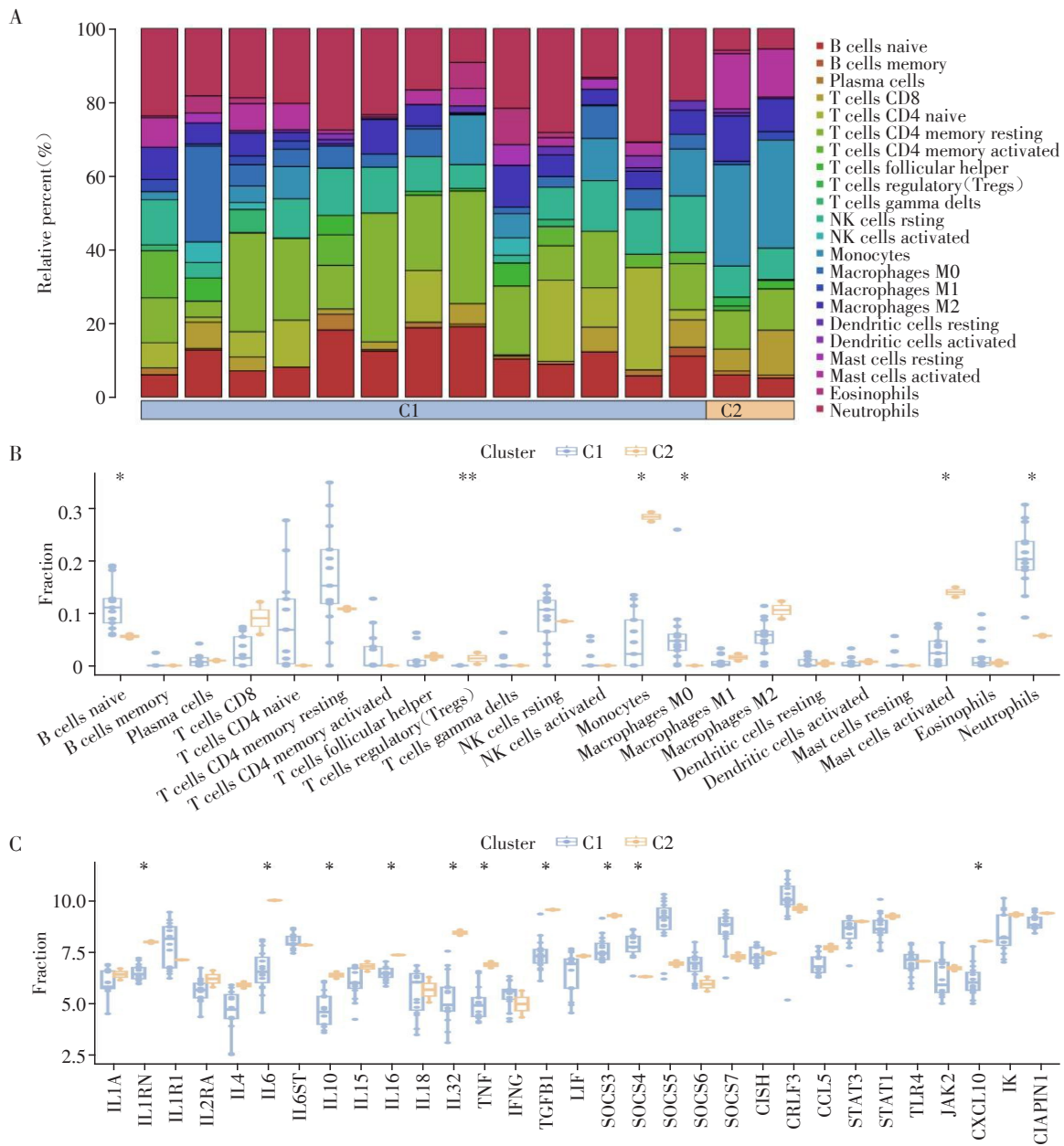
图4 基于焦亡相关基因的新亚型鉴定

Figure 4 Identification of novel subtypes based on pyroptosis-related genes

6B)。NLRP2与细胞间黏附、离子转运、调节白细胞活化呈负相关(图6C)。PYCARD与细胞活化、细胞外泌体、免疫反应、炎症反应、髓系白细胞活化呈负相关(图6D)。在KEGG结果中,CHMP4B与程序性坏死、EB病毒感染、糖尿病心肌病呈正相关(图6E)。NLRP2与代谢通路呈正相关,与造血细胞谱系、细胞因子-细胞因子受体相互作用呈负相关(图6F)。PYCARD与细胞黏附分子、趋化因子信号通路、溶酶体、代谢通路、吞噬体呈正相关(图6G)。

2.7 免疫细胞与特征基因的相关性分析

对模型基因与22种免疫相关细胞类型进行相关性分析。AIM2高表达组与低表达组免疫细胞分布差异无统计学意义(图7A)。在CHMP4B高表达组中,T细胞CD8和活化树突状细胞表达水平较高(图7B)。NLRC4高表达组的活化肥大细胞表达水平升高,NLRC4低表达组的静息肥大细胞表达水平升高(图7C)。NLRP2高表达组具有较高的巨噬细胞M0表达水平,而低表达组具有较高的浆细胞表达水平(图7D)。PYCARD基因表达水平与免疫细



A: Histogram of the expression levels of 22 immunocytes in cluster 1 and cluster 2. B: Differential expressions of immune cell infiltration between two clusters, * $P < 0.05$, ** $P < 0.01$. C: Differential expressions of cytokines between two clusters, * $P < 0.05$.

图5 基于焦亡相关基因亚型的免疫状况

Figure 5 Immune landscape of subtypes of pyroptosis-related genes

胞无显著相关性(图7E)。

2.8 特征基因的细胞因子差异表达分析

细胞因子分析探讨模型基因与细胞因子表达之间的关系。AIM2基因高表达组中,IL10、含细胞因子可诱导SH2蛋白(cytokine inducible SH2-containing protein, CISH)、Toll样受体4(Toll-like receptor 4, TLR4)表达水平较高,而IL4在低表达组中表达水平较高(图8A)。CHMP4B低表达组IL6ST表达水平较高(图8B)。NLRP2高表达组的IL6和TNF表达水平升高(图8C)。NLRP2与IL4表达呈负相关,与SOCS7表达呈正相关(图8D)。TGFB1、CISH与PYCARD的表达呈正相关(图8E)。

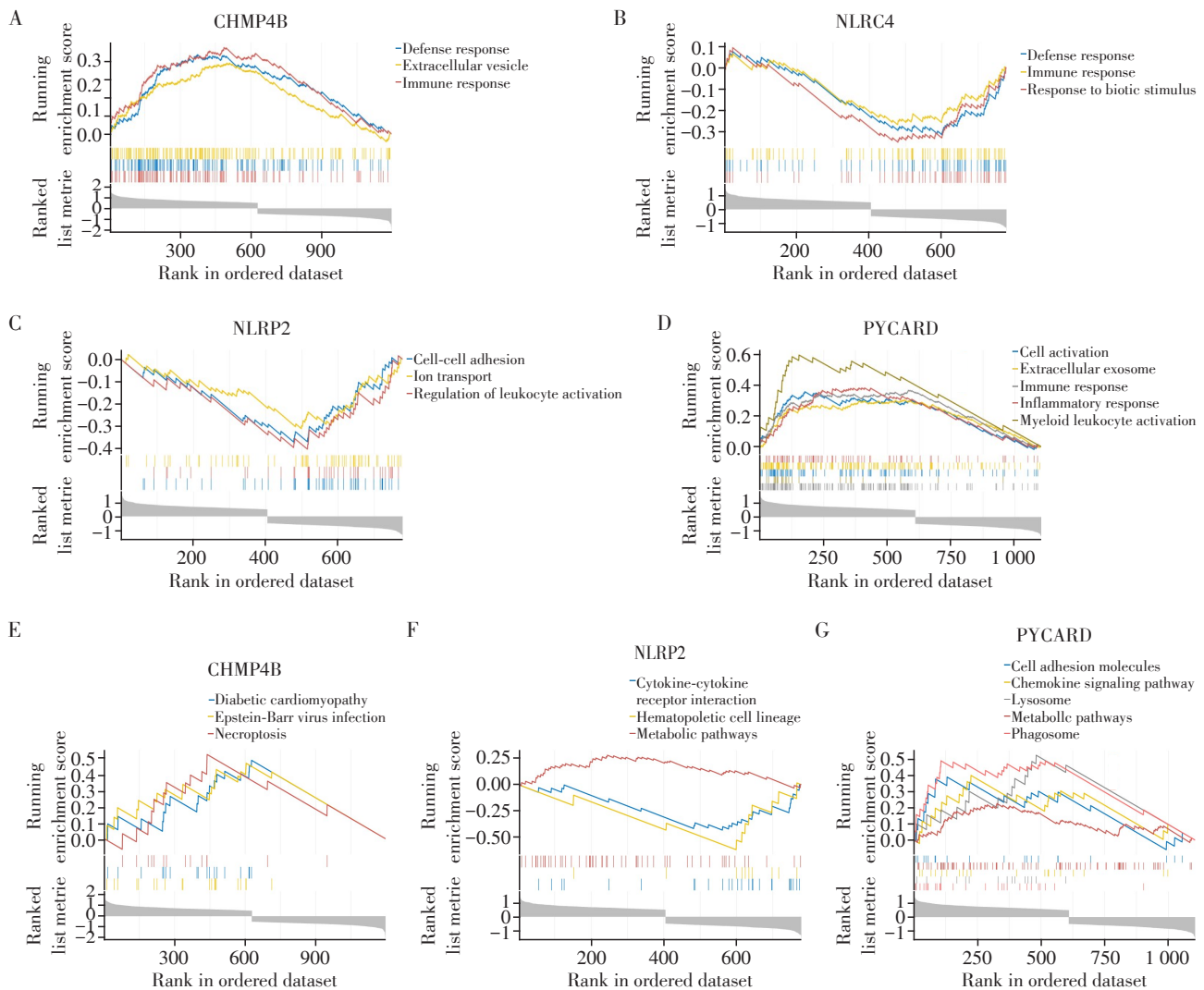
2.9 人血浆中特征基因的表达量

为进一步验证以上生物信息分析结果,收集入

组PCOS患者和正常对照组人群血浆各10例,通过RT-qPCR方法测定血浆中这5种特征基因表达水平。结果显示PYCARD、AIM2、CHMP4B和NLRP2在PCOS患者组中的表达水平高于正常对照组,差异均有统计学意义($P < 0.05$,图9A~D)。虽NLRP4在PCOS患者组中的表达量与正常对照组存在差异,但差异无统计学意义(图9E)。这也进一步验证了XGB建立的预测模型具有可行性。

3 讨论

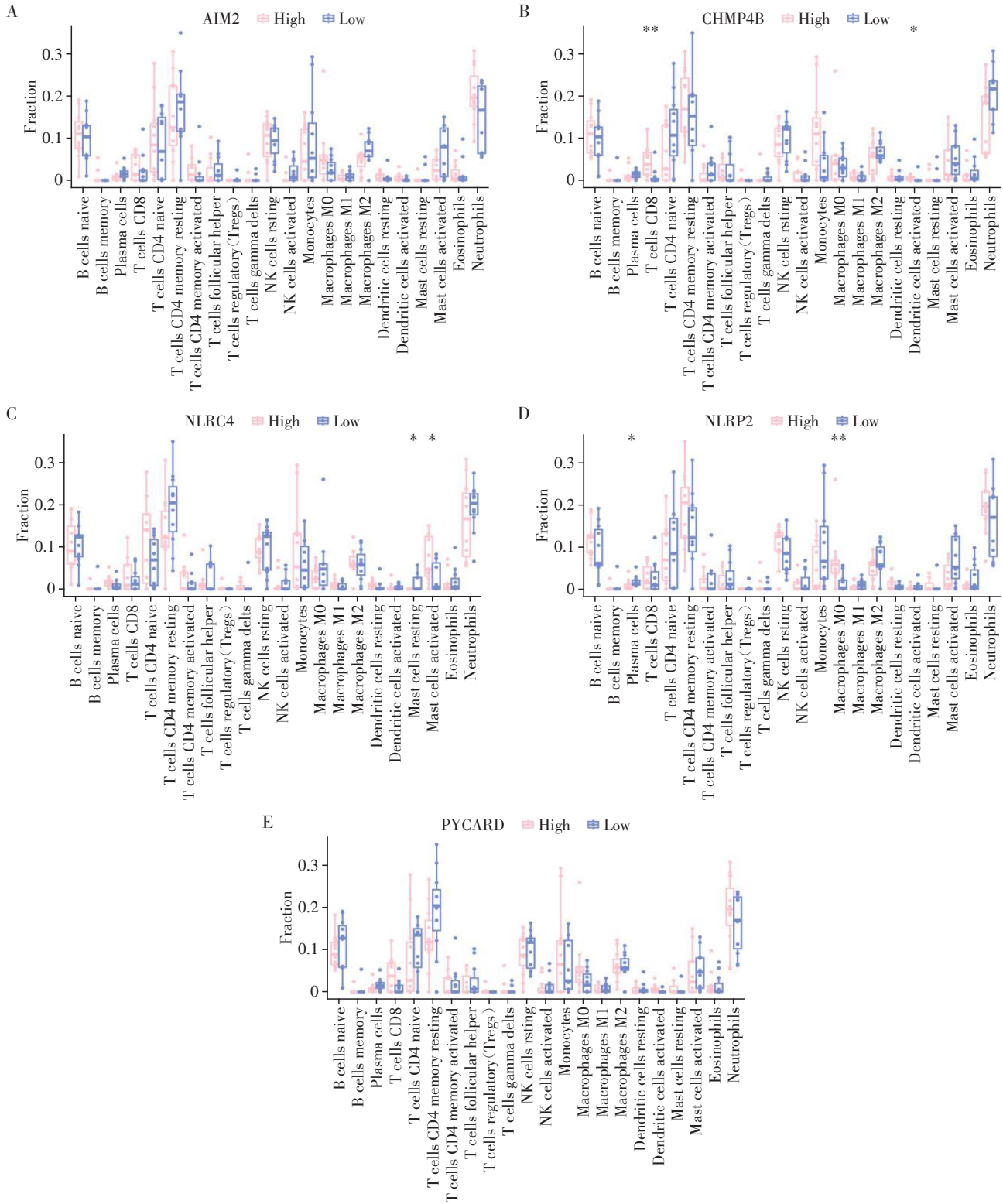
PCOS是育龄妇女最常见的内分泌疾病,其患病率为5%~15%^[1]。PCOS的临床诊断和症状主要包括慢性无排卵和高雄激素血症^[18]。PCOS的病因尚不明确,但研究提示颗粒细胞(granulosa cell, GC)



A-D: GSEA-GO analysis of CHMP4B(A), NLRP4(B), NLRP2(C) and PYCARD(D). E-G: GSEA-KEGG analysis of CHMP4B(E), NLRP2(F), and PYCARD(G).

图6 特征基因功能富集分析

Figure 6 Function enrichment analysis of signature genes



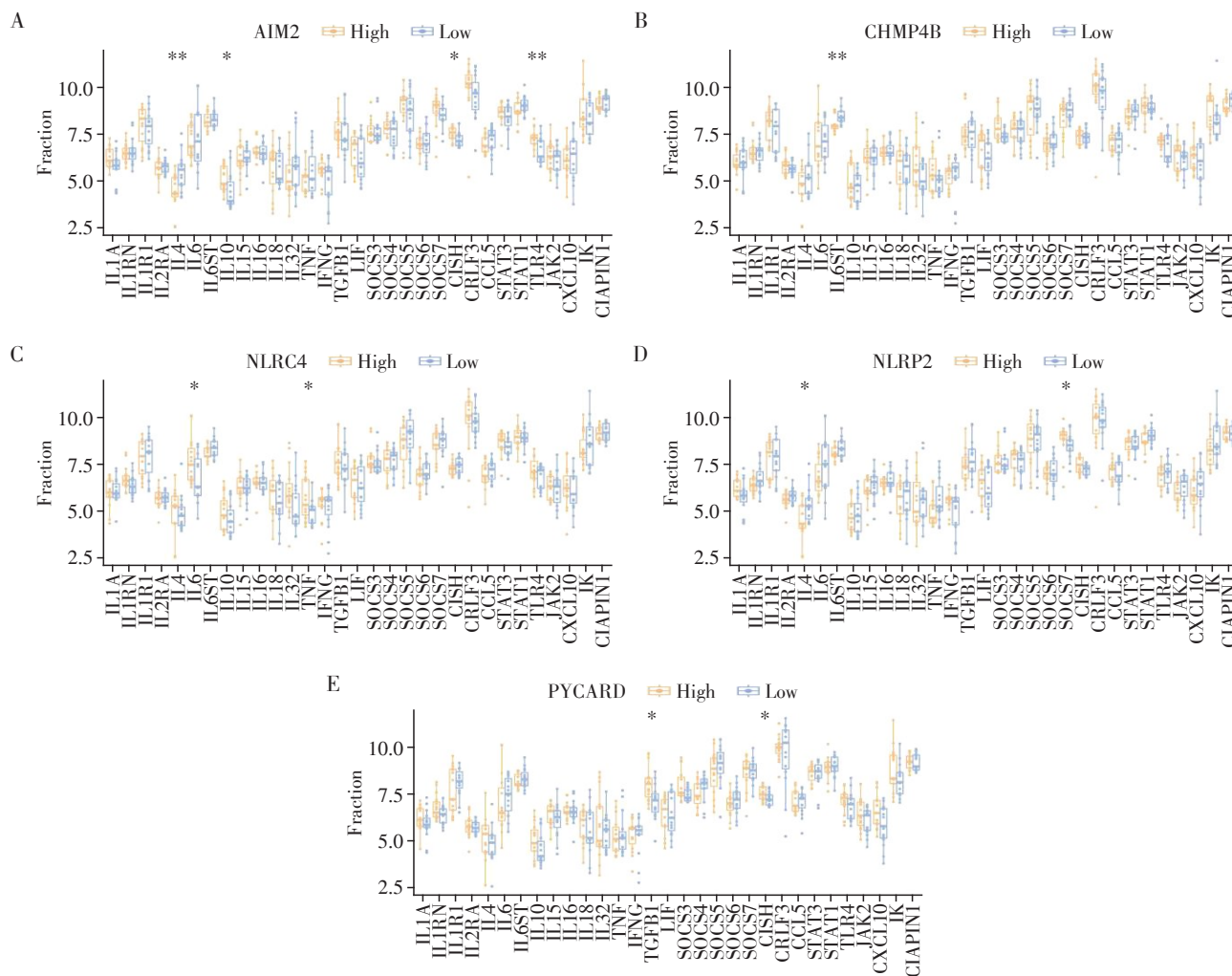
A: Correlation analysis between immunocytes and AIM2. B: Correlation analysis between immunocytes and CHMP4B. C: Correlation analysis between immunocytes and NLR4. D: Correlation analysis between immunocytes and NLRP2. E: Correlation analysis between immunocytes and PYCARD. * $P < 0.05$ and ** $P < 0.01$.

图7 免疫细胞与特征基因的相关性分析

Figure 7 Correlation analysis between immunocytes and signature genes

活性异常可能与PCOS的发病密切相关^[19]。细胞焦亡作为一种新的促炎程序性细胞死亡方式,在形态

和机制上与细胞凋亡有差异,与阿尔茨海默病和2型糖尿病相关^[7]。然而,细胞焦亡调控PCOS的确切机



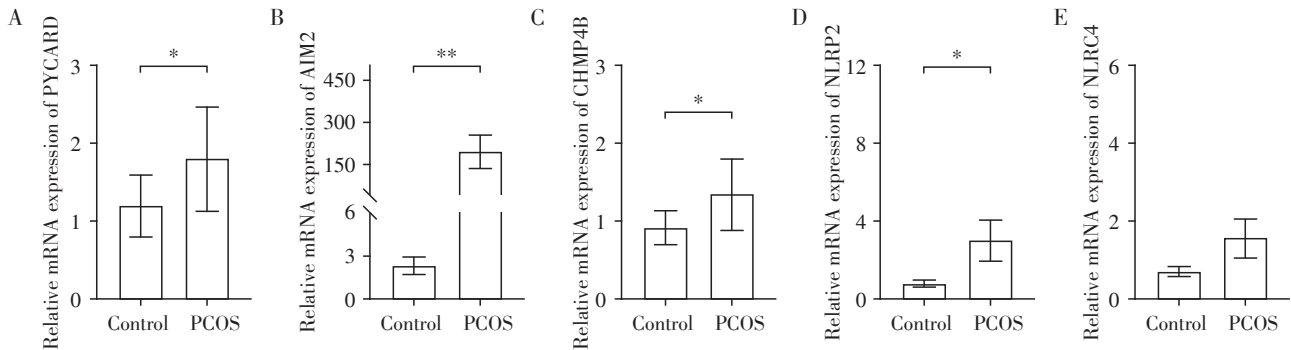
A: Differential expression analysis of cytokines for AIM2. B: Differential expression analysis of cytokines for CHMP4B. C: Differential expression analysis of cytokines for NLR4. D: Differential expression analysis of cytokines for NLRP2. E: Differential expression analysis of cytokines for PYCARD, * $P < 0.05$ and ** $P < 0.01$.

图8 特征基因的细胞因子差异表达分析

Figure 8 Differential expression analysis of cytokines for signature genes

制尚不明确。因此,本研究试图通过生物信息学分析来阐明PRG在PCOS和免疫微环境中的作用。此外,使用与细胞焦亡相关的遗传标记来预测PCOS亚型。首先,对PRG在PCOS患者和正常对照中的表达和通路进行了分析。在PCOS患者中,观察到PYCARD、AIM2和NOD2基因表达水平显著上调,表明这些PRG可能参与PCOS的发生发展^[20-22]。这提示细胞焦亡可能在PCOS的发病机制中起重要作用。随后,利用这些PRG建立了4个预测模型,以进一步研究细胞焦亡和PCOS之间的相互作用。在这些模型中,选择XGB作为最好的机器学习模型。XGB是一个并行预测模型,它使用多棵树,并不断迭代,在每次迭代时生成更新的树^[23],使预测值与真实值密切一致,成为本研究的有效模型。

本研究最初选择的5个基因是CHMP4B、PYCARD、NLRP2、NLR4和AIM2。随后,通过共识聚类,确定了PCOS的两个不同亚型,称为Cluster 1和Cluster 2。进一步分析了两个亚型之间的差异表达的PRG,以识别具有显著表达差异的基因簇。与Cluster 1相比,Cluster 2中PRG的表达水平显著升高,强调了这一区别的关键作用。研究表明,PCOS的特征是激素异常,过度刺激免疫系统,导致慢性低水平炎症,进而导致卵巢功能障碍和代谢紊乱^[24-25]。鉴于慢性低水平炎症在PCOS中的突出作用,本研究比较了不同组间的免疫细胞浸润和细胞因子水平。值得注意的是,Cluster 2表现出较高水平的免疫细胞浸润,这与之前的研究一致,并且在各基因组亚型中均观察到这些浸润趋势。此外,



A: PYCARD expression levels in human plasma. B: AIM2 expression levels in human plasma. C: CHMP4B expression levels in human plasma. D: NLRP2 expression levels in human plasma. E: NLRC4 expression levels in human plasma, * $P < 0.05$ and ** $P < 0.01$ ($n=10$).

图9 人血浆中特征基因的表达量

Figure 9 The expression levels of specific genes in human plasma

与 Cluster 1 相比, Cluster 2 中的细胞因子水平显著升高,尤其是 TNF、TGFB1 和 IL6,这与现有研究一致^[26-28]。因此,这些发现提示 Cluster 2 可能表现出更典型的 PCOS 特征,预示着更高的风险和潜在较差的临床结局。

NLRC4 在 GO 分析中与防御反应、生物刺激反应、免疫反应呈负相关。这表明 NLRC4 可能在免疫调控过程中起负反馈作用。NLRC4 作为一种炎症小体相关基因,其抑制可能减弱焦亡途径的活性,进而影响 PCOS 患者的慢性炎症状态^[29-30]。通过调控免疫反应, NLRC4 可能影响 PCOS 的炎症水平。NLRP2 在 GO 分析中与细胞间黏附、离子转运、白细胞活化呈负相关。其在 KEGG 分析中与代谢通路呈正相关,提示 NLRP2 可能参与调控代谢和免疫过程。已有研究表明, NLRP2 与细胞焦亡和胰岛素抵抗相关,而胰岛素抵抗是 PCOS 的重要代谢特征^[31]。因此 NLRP2 可能通过影响免疫调节以及调控代谢,在 PCOS 的发病过程中发挥重要作用。PYCARD 在 GO 分析中与细胞活化、细胞外泌体、免疫反应和炎症反应呈正相关,在 KEGG 分析中与细胞黏附分子、趋化因子信号通路和溶酶体通路呈正相关。这些功能表明 PYCARD 可能通过调控炎症因子释放和免疫细胞活化,参与 PCOS 相关的慢性炎症和代谢紊乱^[32]。PYCARD 的高表达可能加剧 PCOS 患者的炎症状态,并与内分泌失调和代谢异常密切相关。这些特征基因通过调控免疫反应、细胞间信号传递、代谢平衡等多种途径,与 PCOS 的病理机制高度相关。

细胞焦亡在 PCOS 发生发展中的临床意义和潜在机制尚不完全清楚,本研究有助于更好地理解细胞焦亡在 PCOS 发病机制中的作用。此外, PRG 的

鉴定为 PCOS 患者的精准基因靶向治疗开辟了新途径。然而,本研究有一定局限性。第一,本研究主要基于生物信息学方法,虽进行了临床试验验证了预测模型,但血浆测定出特征基因 PYCARD、AIM2、CHMP4B 和 NLRP2 有统计学意义,而 NLRC4 有差异但无统计学意义 ($P=0.2475$),可能是样本量较少,影响模型的泛化性和统计学显著性。第二,本研究聚焦于焦亡基因,可能忽略了 PCOS 相关的其他关键分子机制,仅依赖 PRG 可能限制了模型的诊断覆盖范围。第三,缺少对 PCOS 患者的随访数据,无法评估预测模型对疾病进展的预测能力,需要进一步的临床随访和实验研究来证实 PCOS 中不同 PRG 参与细胞焦亡的调控作用。之后还需要更多 PCOS 样本来确定 PRG 簇的准确性。第四, PRG 与 PCOS 免疫细胞浸润的潜在相关性值得进一步探索。第五,需要考虑更全面的临床特征来评估预测模型的性能。

综上所述,本研究强调了 PRG 和 PCOS 之间的显著关联,以及在 PCOS 中由这些 PRG 调节的免疫细胞浸润模式。此外,本研究还发现了 PCOS 中两种细胞群之间的免疫异质性。值得注意的是,本研究结果证明了机器学习模型,尤其是 5 基因 XGB 模型在准确评估 PCOS 亚型方面的潜力。本研究揭示了细胞焦亡与 PCOS 之间的新关系,为进一步探索两者之间的相互作用提供了新思路。此外,本研究还建立了一个有前景的预测模型,可以有效地评估不同 PCOS 亚型相关的风险。

利益冲突声明:

所有作者声明无利益冲突。

Conflict of Interests:

All authors declare no conflict of interests.

作者贡献声明:

陈讯提出研究问题和概念框架,主导研究设计与规划;

负责数据收集、分析与整理;撰写论文初稿,修订论文。张文协助研究设计,临床样本的收集与保存。陈雯昕和俞池园完善实验流程及参与实验,参与部分数据收集工作。许波群负责研究指导与规划,数据质量审核,论文修改与审查。

Author's Contributions:

CHEN Xun proposed research questions and conceptual frameworks, led the design and planning of the research, was responsible for data collection, analysis, and organization, wrote the initial draft of the paper, and participated in revision. ZHANG Wen assisted in research and design of the study, participated in collection and preservation of clinical samples. CHEN Wenxin and YU Chiyuan participated in the experimental process, and took part in the data collection. XU Boqun provided guidance on the research and planning of the study, verified and data quality, reviewed and revised the article.

[参考文献]

- [1] WALTER K. What is polycystic ovary syndrome? [J]. *JAMA*, 2022, 327(3): 294
- [2] SADEGHI H M, ADELI I, CALINA D, et al. Polycystic ovary syndrome: a comprehensive review of pathogenesis, management, and drug repurposing [J]. *Int J Mol Sci*, 2022, 23(2): 583
- [3] GOMEZ J M D, VANHISE K, STACHENFELD N, et al. Subclinical cardiovascular disease and polycystic ovary syndrome[J]. *Fertil Steril*, 2022, 117(5): 912-923
- [4] GLUECK C J, GOLDENBERG N. Characteristics of obesity in polycystic ovary syndrome: etiology, treatment, and genetics[J]. *Metabolism*, 2019, 92: 108-120
- [5] ATABAKI R , KHALEGHZADEH-AHANGAR H , ESMAEILI N, et al. Role of pyroptosis, a pro-inflammatory programmed cell death, in epilepsy[J]. *Cell Mol Neurobiol*, 2023, 43(3): 1049-1059
- [6] RAO Z P, ZHU Y T, YANG P, et al. Pyroptosis in inflammatory diseases and cancer [J]. *Theranostics*, 2022, 12(9): 4310-4329
- [7] HSU S K, LI C Y, LIN I L, et al. Inflammation-related pyroptosis, a novel programmed cell death pathway, and its crosstalk with immune therapy in cancer treatment [J]. *Theranostics*, 2021, 11(18): 8813-8835
- [8] WANG D J, WENG Y J, ZHANG Y L, et al. Exposure to hyperandrogen drives ovarian dysfunction and fibrosis by activating the NLRP3 inflammasome in mice [J]. *Sci Total Environ*, 2020, 745: 141049
- [9] GÖZÜKÜÇÜK M, GÜRSOY A Y, DESTEĞÜL E, et al. Homocysteine and C-reactive protein levels in women with polycystic ovary syndrome [J]. *Gynecol Minim Invasive Ther*, 2021, 10(4): 210-214
- [10] STRINGER J M, ALESI L R, WINSHIP A L, et al. Beyond apoptosis: evidence of other regulated cell death pathways in the ovary throughout development and life [J]. *Hum Reprod Update*, 2023, 29(4): 434-456
- [11] AMSTERDAM A, SASSON R, KEREN-TAL I, et al. Alternative pathways of ovarian apoptosis: death for life [J]. *Biochem Pharmacol*, 2003, 66(8): 1355-1362
- [12] MURRI M, LUQUE-RAMÍREZ M, INSENSER M, et al. Circulating markers of oxidative stress and polycystic ovary syndrome (PCOS): a systematic review and meta-analysis [J]. *Hum Reprod Update*, 2013, 19(3): 268-288
- [13] CAI Z W, HE S J, LIU R J, et al. Plumbagin rescues the granulosa cell's pyroptosis by reducing WTAP-mediated N6-methylation in polycystic ovary syndrome [J]. *J Ovarian Res*, 2022, 15(1): 126
- [14] ZHANG Y, XIE X G, MA Y B, et al. Cyproterone acetate mediates IRE1 α signaling pathway to alleviate pyroptosis of ovarian granulosa cells induced by hyperandrogen [J]. *Biology*, 2022, 11(12): 1761
- [15] CHEN J Y, ZHU Z Y, XU S, et al. HDAC1 participates in polycystic ovary syndrome through histone modification by regulating H19/miR-29a-3p/NLRP3-mediated granulosa cell pyroptosis [J]. *Mol Cell Endocrinol*, 2023, 573: 111950
- [16] PEI S B, ZHANG P P, YANG L L, et al. Exploring the role of sphingolipid-related genes in clinical outcomes of breast cancer [J]. *Front Immunol*, 2023, 14: 1116839
- [17] PEI S B, ZHANG P P, CHEN H L, et al. Integrating single-cell RNA-seq and bulk RNA-seq to construct prognostic signatures to explore the role of glutamine metabolism in breast cancer [J]. *Front Endocrinol*, 2023, 14: 1135297
- [18] SIDDIQUI S, MATEEN S, AHMAD R, et al. A brief insight into the etiology, genetics, and immunology of polycystic ovarian syndrome (PCOS) [J]. *J Assist Reprod Genet*, 2022, 39(11): 2439-2473
- [19] ZHOU J H, CHEN Z J, LI J M, et al. Expression of PGRMC1 in patients with polycystic ovary syndrome and its molecular mechanism for regulating ovarian granulosa cell apoptosis and glucolipid metabolism [J]. *J Cent South Univ Med Sci*, 2023, 48(4): 538-549
- [20] LAMMERT C R, FROST E L, BELLINGER C E, et al. AIM2 inflammasome surveillance of DNA damage shapes neurodevelopment [J]. *Nature*, 2020, 580(7805): 647-652
- [21] SHI C X, WANG Y, CHEN Q, et al. Extracellular histone H3 induces pyroptosis during sepsis and may act through NOD2 and VSIG4/NLRP3 pathways [J]. *Front Cell Infect Microbiol*, 2020, 10: 196
- [22] KHARE S, DORFLEUTNER A, BRYAN N B, et al. An NLRP7-containing inflammasome mediates recognition of

- microbial lipopeptides in human macrophages[J]. *Immunity*, 2012, 36(3):464-476
- [23] WANG K, TIAN J, ZHENG C, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP[J]. *Comput Biol Med*, 2021, 137: 104813
- [24] BURCHALL G, LINDEN M D, TEEDE H, et al. Hemostatic abnormalities and relationships to metabolic and hormonal status in polycystic ovarian syndrome [J]. *Trends Cardiovasc Med*, 2011, 21(1):6-14
- [25] LI Y, CHEN C Y, MA Y, et al. Multi-system reproductive metabolic disorder: significance for the pathogenesis and therapy of polycystic ovary syndrome (PCOS) [J]. *Life Sci*, 2019, 228: 167-175
- [26] SADEGHI A, FADAEI R, MORADI N, et al. Circulating levels of C1q/TNF- α -related protein 6 (CTRP6) in polycystic ovary syndrome [J]. *IUBMB Life*, 2020, 72(7): 1449-1459
- [27] AZUMAH R, LIU M H, HUMMITZSCH K, et al. Candidate genes for polycystic ovary syndrome are regulated by TGF β in the bovine foetal ovary [J]. *Hum Reprod*, 2022, 37(6):1244-1254
- [28] RUDNICKA E, SUCHTA K, GRYMOWICZ M, et al. Chronic low grade inflammation in pathogenesis of PCOS [J]. *Int J Mol Sci*, 2021, 22(7): 3789
- [29] OH S, LEE J, OH J, et al. Integrated NLRP3, AIM2, NLRC4, Pyrin inflammasome activation and assembly drive PANoptosis [J]. *Cell Mol Immunol*, 2023, 20(12): 1513-1526
- [30] ROSTAMTABAR M, ESMAEILZADEH S, KARKHAH A, et al. Elevated expression of IL-18 but not IL-1 β gene is associated with NALP3 and AIM2 inflammasome in polycystic ovary syndrome [J]. *Gene*, 2020, 731: 144352
- [31] LI C, LIU Q, XIE L Q. Suppressing NLRP2 expression accelerates hepatic steatosis: a mechanism involving inflammation and oxidative stress [J]. *Biochem Biophys Res Commun*, 2018, 507(1/2/3/4):22-29
- [32] STIENSTRA R, VAN-DIEPEN J A, TACK C J, et al. Inflammasome is a central player in the induction of obesity and insulin resistance [J]. *Proc Natl Acad Sci U S A*, 2011, 108(37): 15324-15329
- [收稿日期] 2024-09-29
(本文编辑: 陈汐敏)

(上接第486页)

- Analytical methodology to measure periodontal bone morphometry following orthodontic tooth movement in mice [J]. *Eur J Orthod*, 2021, 43(6): 665-671
- [33] KONDO T, HOTOKEZAKA H, HAMANAKA R, et al. Types of tooth movement, bodily or tipping, do not affect the displacement of the tooth's center of resistance but do affect the alveolar bone resorption [J]. *Angle Orthod*, 2017, 87(4): 563-569
- [34] SMITH R J, BURSTONE C J. Mechanics of tooth movement [J]. *Am J Orthod*, 1984, 85(4): 294-307
- [35] BAXTER S J, SYDORAK I, MA P X, et al. Impact of pharmacologic inhibition of tooth movement on periodontal and tooth root tissues during orthodontic force application [J]. *Orthod Craniofac Res*, 2020, 23(1): 35-43
- [36] MUSTAFY T, BENOIT A, LONDONO I, et al. Can repeated *in vivo* micro-CT irradiation during adolescence alter bone microstructure, histomorphometry and longitudinal growth in a rodent model? [J]. *PLoS One*, 2018, 13(11): e0207323
- [收稿日期] 2024-11-18
(本文编辑: 陈汐敏)