· 论著 ·

泌尿系结石患者尿有机酸代谢谱特征及关键代谢标志物的筛选

漆子暄[1],边小龙[1],胡浩浦[1],田 聪[1],王辰龙[1],周雨缪[1],徐 涛[1],王 辉[2],曹林林[2],胡 浩[1△]
(北京大学人民医院 1. 泌尿外科,2. 检验科, 北京 100044)

[摘 要] **目的**:系统分析泌尿系结石患者与非结石患者尿有机酸代谢谱的差异,筛选疾病相关特征性代谢标志物及关键信号通路,从代谢层面阐明疾病发生发展的潜在机制,为泌尿系结石的基础研究与干预靶点筛选提供理论依据。**方法**:选择 2024 年 12 月至 2025 年 9 月于北京大学人民医院就诊的泌尿系结石患者 99 例与 57 例非泌尿系结石患者(泌尿外科以外未患有泌尿系结石的就诊者),收集其 24 h 尿,通过气相色谱-质谱(gas chromatography-mass spectrometry, GC-MS)技术行非靶向尿有机酸代谢组学检测。采用主成分分析(principal component analysis, PCA)和正交偏最小二乘判别分析(orthogonal partial least squares discriminant analysis, OPLS-DA)组间代谢谱差异及临床协变量的影响;通过火山图、变量重要性投影(variable importance in the projection, VIP)分析筛选差异有机酸;利用 KEGG(Kyoto Encyclopedia of Genes and Genomes)数据库开展代谢通路富集分析;采用机器学习筛选最具代表性的核心代谢物组合,验证有机酸对疾病状态的区分效能。**结果**:结石组与非结石组患者的尿有机酸代谢呈显著分离趋势(OPLS-DA:$R2X=0.139$、$R2Y=0.48$、$Q2=0.457$),性别是影响代谢谱的重要协变量,糖尿病、高血压及结石复发状态对尿有机酸谱无显著影响。共筛选出 44 种差异有机酸($P<0.05$,6 种上调、38 种下调),VIP 分析确定 10 种核心差异代谢物($P<0.001$,均下调);差异有机酸主要富集于乙醛酸和二羧酸代谢等能量代谢通路及色氨酸代谢等氨基酸代谢通路。机器学习模型中梯度提升机(gradient boosting machine, GBM)、随机森林(random forest, RF)效能最优,平均曲线下面积(area under the curve,AUC)为 0.996,筛选出丙酰甘氨酸、5-羟基吲哚-3-乙酸、犬尿喹啉酸、乳清酸、富马酸 5 种核心代谢物为最佳组合,基于该组合构建的列线图模型 AUC = 0.997,校准曲线校正偏倚后与理想曲线高度吻合,平均绝对误差为 0.022,能够高效区分结石与非结石人群,提示这些代谢物可以作为泌尿系结石机制研究的潜在靶点。**结论**:泌尿系结石患者存在显著尿有机酸代谢紊乱,代谢异常主要涉及能量及氨基酸代谢通路,其紊乱与疾病发病机制密切相关。丙酰甘氨酸、5-羟基吲哚-3-乙酸、犬尿喹啉酸、乳清酸、富马酸 5 种核心代谢物可以有效地区分结石与非结石人群,其可以作为代表性的代谢标志物为后续机制探索与靶向干预提供新靶点。

[关键词] 泌尿系结石;尿有机酸;代谢组学;泌尿外科
[中图分类号] R691.4 [文献标志码] A [文章编号] 1671-167X(2026)04-0738-10
doi:10.19723/j.issn.1671-167X.2026.04.009

Characteristics of urinary organic acid metabolic profile and screening of key metabolic markers in patients with urolithiasis

QI Zixuan[1], BIAN Xiaolong[1], HU Haopu[1], TIAN Cong[1], WANG Chenlong[1], ZHOU Yumou[1], XU Tao[1], WANG Hui[2], CAO Linlin[2], HU Hao[1△]
(1. Department of Urology, 2. Department of Clinical Laboratory, Peking University People's Hospital, Beijing 100044, China)

ABSTRACT **Objective**: To systematically compare the differences in urinary organic acid metabolic profiles between patients with urinary calculi and non-calculi individuals, to screen disease-specific characteristic metabolic biomarkers and key signaling pathways, and to elucidate the potential mechanism underlying the occurrence and progression of urinary calculi at the metabolic level, so as to provide a theoretical basis for basic research and screening of intervention targets for urinary calculi. **Methods**: A total of 99 patients with urinary calculi and 57 non-calculi subjects were enrolled. Twenty-four-hour urine samples were collected and subjected to untargeted urinary organic acid metabolomics detection by gas chromatography-mass spectrometry (GC-MS). Principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) were used to analyze the differences in metabolic profiles between the groups and the effects of clinical covariates. Volcano plot and variable importance in the projection (VIP) analysis were applied to screen differential organic acids. The Kyoto Encyclopedia of

△ Corresponding author's e-mail, huhao@163.com
网络出版时间:2026-05-27 10:27:18 网络出版地址:http://kns.cnki.net/kcms/detail/11.4691.R.20260526.1658.002.html

Genes and Genomes (KEGG) database were used for metabolic pathway enrichment analysis. Machine learning algorithms were adopted to screen the most representative core metabolite combinations and verify the discriminatory efficacy of organic acids for disease status. **Results**: The urinary organic acid metabolic profiles showed a significant separation trend between the calculi group and the non-calculi group (OPLS-DA: $R2X=0.139$, $R2Y=0.48$, $Q2=0.457$). Gender was an important covariate affecting metabolic profiles, while diabetes, hypertension and stone recurrence status exerted no significant influence. A total of 44 differential organic acids were screened ($P<0.05$), including 6 upregulated and 38 downregulated metabolites. Ten core differential metabolites were identified via VIP analysis ($P<0.001$, all downregulated). Differential organic acids were mainly enriched in energy metabolism pathways such as glyoxylate and dicarboxylate metabolism, as well as amino acid metabolism pathways including tryptophan metabolism. Among machine learning models, Gradient Boosting Machine and Random Forest exhibited the optimal efficacy with an average area under the curve (AUC) of 0.996. Five core metabolites including propionylglycine, 5-hydroxyindole-3-acetic acid, kynurenic acid, orotic acid and fumaric acid were identified as the optimal combination. The nomogram model constructed based on this combination yielded an AUC of 0.997; after calibration bias correction, the calibration curve was highly consistent with the ideal curve, with a mean absolute error of 0.022, which could effectively distinguish calculi patients from non-calculi individuals. **Conclusion**: Patients with urinary calculi present obvious disorders in urinary organic acid metabolism, which are mainly involved in energy and amino acid metabolic pathways and closely associated with the pathogenesis of urinary calculi. The five core metabolites (propionylglycine, 5-hydroxyindole-3-acetic acid, kynurenic acid, orotic acid and fumaric acid) can effectively distinguish calculi from non-calculi populations, and can serve as representative metabolic biomarkers to provide novel targets for subsequent mechanism exploration and targeted intervention.

KEY WORDS　Urolithiasis; Urinary organic acids; Metabolomics; Urology

泌尿系结石作为全球高发的泌尿系统疾病,其高发病率与高复发特性已成为临床诊疗及公共卫生领域面临的突出问题。泌尿系结石具有极强的复发倾向,据报道,其5年复发率约为30%,10年复发率更是高达50%[1-3]。长期反复发作不仅易引发泌尿系感染、尿路梗阻等并发症,严重时还可导致肾功能损伤,甚至进展为慢性梗阻性肾病,对患者的健康构成持续性威胁[4]。

泌尿系结石的发生机制复杂,涉及尿液成分代谢紊乱、尿路菌群失衡、尿路炎症反应及遗传因素等多个方面[5-6]。尽管科研工作者已对其开展了多年研究,但目前关于泌尿系结石的具体发病及复发机制仍不明确,如何有效预防结石形成与复发,仍是当前泌尿外科领域面临的重大挑战。

近年来,基因组学、转录组学、蛋白质组学及代谢组学等新兴技术已被广泛应用于泌尿系结石的诊断及发病机制研究中,尿有机酸代谢异常被认为是尿路结石形成的重要诱因之一,某些尿液中有机酸的浓度与种类直接影响结石的形成与生长进程,如尿液中草酸、尿酸浓度升高,通常与草酸钙结石、尿酸结石的形成密切相关[7],而枸橼酸含量减少则可能导致尿液酸性增强,进而促进结石沉积[8]。此外,一些有机酸,比如琥珀酸、龙胆酸等则可以通过抑制结石成核,改善局部炎症等机制间接影响结石的发生[9]。但是,目前明确关键代谢通路及核心分子的调控作用的研究仍较少见。因此,本研究拟系统比较泌尿系结石患者与非结石人群的尿有机酸谱差异,明确与结石发生相关的特征性有机酸标志物,挖掘有机酸与结石相关的代谢通路,从有机酸代谢角度揭示泌尿系结石的发生机制,进一步探究泌尿系结石潜在的干预靶点。

1　资料与方法

1.1　研究对象

本研究为病例对照研究,选择2024年12月至2025年9月于北京大学人民医院就诊的泌尿系结石患者以及非泌尿系结石患者(非结石患者均为北京大学人民医院泌尿外科以外未患有泌尿系结石的就诊者)为研究对象。研究对象均来自北京大学人民医院同期就诊人群,病例组与对照组来源一致,基线资料均衡可比,研究人群具有良好的临床代表性,能够反映目标人群的尿有机酸代谢特征。本研究中泌尿系结石的诊断符合2024年欧洲泌尿外科协会(European Association of Urology, EAU)指南尿石症(EAU Guidelines on Urolithiasis-2024)的诊断标准,即CT平扫可显示患有泌尿系结石。排除标准:(1)尿路上皮性肿瘤或伴有严重的副癌综合征的除尿路上皮肿瘤之外的恶性肿瘤;(2)存在严重肾功能不全;(3)近期接受过其他影响尿液有机酸水平的治疗(如肾移植、长期使用抗生素、长期口服钙剂等);(4)妊娠;(5)年龄小于18岁。本研究开始前已经北京大学人民医院生物医学伦理委员会审查批准(2026PHB424-001),所有参与研究的患者均签署知情同意书。

1.2 尿液样本的处理

收集泌尿系结石患者与非结石患者的24 h尿标本,保存于4 ℃冰箱中,后于离心机中3 500 ~ 4 000 r/min离心10 min,彻底分离上清液中的不溶性代谢物,去除细胞和杂质。离心后立即取上清液,4 ℃冷藏保存。

1.3 仪器和分析条件

气相色谱-三重四级杆质谱仪(PreMed 6000)购自杭州谱聚医疗科技有限公司。色谱条件:RESTEK Rxi-5ms色谱柱(30 m × 0.25 mm ID, 0.25 μm),进样口温度为280 ℃,载气为高纯氦气(≥99.999%),恒定流速为1.0 mL/min,进样量1 μL,分流比为6∶1,程序升温至60 ℃保持2 min,以10 ℃/min的升温速率升至220 ℃,随后以15 ℃/min的速率升至325 ℃,保持1 min。质谱条件离子源为电子轰击离子源,其测定条件为-70 eV,离子源温度为250 ℃,接口温度为280 ℃,锥孔电压为50.00 V,碰撞气流速为0.3 mL/min,数据采集模式为多反应监测。

1.4 数据处理

原始下机数据经质谱分析软件进行峰提取、基线校正、峰对齐及反卷积等预处理;提取的代谢物特征峰通过比对标准数据库进行定性鉴定;为保证后续统计分析的准确性,对提取到的代谢物数据矩阵进行数据清洗;随后,对保留的代谢物峰面积进行归一化处理,并对数据进行以2为底对数转换,以降低极值的影响,使数据符合正态分布或近似正态分布,最终生成用于后续统计及多变量数据分析的标准化表达矩阵。

1.5 统计学分析

使用R Stduio(4.3.1)软件,双侧检验,$P<0.05$认为差异具有统计学意义。

1.5.1 临床基线特征分析 连续型变量若符合正态分布以均数 ± 标准差表示,组间比较采用独立样本 t 检验;若不符合正态分布,则以中位数和四分位间距表示,组间比较采用非参数Mann-Whitney U 检验。分类变量以 n(%)表示,组间差异比较采用卡方检验或Fisher精确检验。

1.5.2 多变量模式识别分析 采用无监督的主成分分析(principal component analysis, PCA)评估结石组与非结石组样本的整体代谢分布趋势及组内聚集特征,进一步构建有监督的正交偏最小二乘判别分析(orthogonal partial least squares discriminant analysis, OPLS-DA)模型。通过模型参数 $R2X$、$R2Y$(解释率)及 $Q2$(预测率)评估模型的拟合优度与预测能力。

1.5.3 差异有机酸筛选与KEGG富集分析 基于OPLS-DA模型提取各代谢物的第一主成分变量投影重要度(variable importance in the projection, VIP)值,结合单变量分析(Wilcoxon 秩和检验)计算 P 值。本研究将VIP > 1且 $P<0.05$ 的代谢物定义为两组间的显著差异有机酸,并利用R包ggplot2绘制火山图。为探究差异代谢物背后的生物学功能与代谢机制,提取所有显著差异有机酸的KEGG ID,借助R包clusterProfiler将其映射至KEGG通路数据库进行富集分析。

1.5.4 机器学习建模与内部验证 为筛选最具临床分类价值的生物标志物组合,构建了10种机器学习预测模型,包括梯度提升机(gradient boosting machine, GBM)、随机森林(random forest, RF)、支持向量机(support vector machine, SVM)、K最近邻(K-nearest neighbors, KNN)、朴素贝叶斯(Naïve Bayes)、弹性网络、LASSO回归、决策树、逻辑回归(Logistic regression, LR)和线性判别分析(linear discriminant analysis, LDA)。通过交叉验证计算各模型在受试者工作特征曲线下的面积,挑选整体表现最优的模型。提取各模型筛选出的特征重要性变量,将至少被3种模型(vote ≥ 3)共同识别的代谢物定义为共识别特征,并绘制小提琴图进行丰度比较。随后,利用R包glmnet对共识别特征进行LASSO回归,以筛选出最优有机酸特征组合。基于该最优组合构建多因素逻辑回归模型,利用R包rms绘制列线图以评估发病风险。最后,通过R包pROC绘制ROC曲线,计算模型的敏感性和特异性,并采用重抽样法(bootstrapping,重抽样1 000次)绘制校准曲线,对模型进行内部交叉验证以评估其准确度。

2 结果

2.1 患者的一般情况

本研究共招募了156例患者,其中结石组99例,非结石组57例。研究对象的临床和人口统计学特征见表1。两组间年龄、高血压、糖尿病等基线特征差异无统计学意义,两组间性别差异有统计学意义,结石组患者男性占比为72.6%,非结石组男性占比为56.1%。

2.2 结石患者与非结石患者尿有机酸代谢谱的组间差异及协变量影响分析

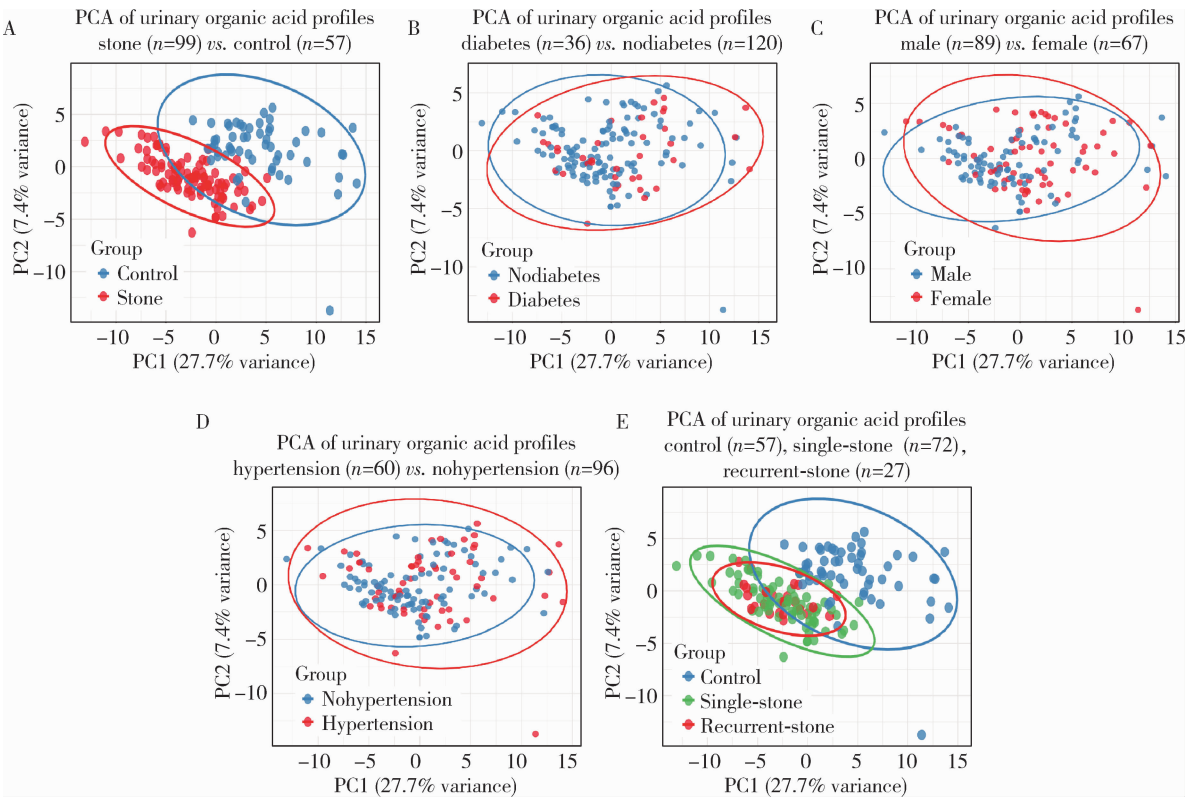
采用PCA对尿液有机酸代谢谱进行全局评估,PCA得分图显示,结石组($n=99$)与非结石患者组($n=57$)在第一主成分(PC1,解释27.7%的方差)

和第二主成分(PC2,解释 7.4% 的方差)维度上呈现出明确的聚类趋势。尽管两组样本存在部分重叠(反映了不同组别间代谢特征的自然变异),但结果仍表明结石患者与非结石人群的尿有机酸代谢谱存在显著差异,其整体代谢特征具有可区分性(图 1)。

表 1　研究对象的临床和人口统计学特征

Table 1　Clinical and demographic characteristics of the study population

Clinical and demographic characteristics	Stone group(n = 99)	Non-stone group(n = 57)	P
Male, n (%)	72(72.7)	21(56.1)	0.027
Age/years, $\bar{x} \pm s$	56.8 ± 12.3	53.4 ± 11.7	0.087
Complicated with diabetes, n (%)	23(23.2)	8(14.0)	0.120
Complicated with hypertension, n (%)	31(31.3)	12(21.1)	0.126
Recurrence	28(28.3)	–	–



A, principal component analysis (PCA) comparing stone patients (n = 99, red) and healthy controls (n = 57, blue), PC1 (27.7%) and PC2 (7.4%) show a clear separation trend, indicating distinct metabolic profiles between stone patients and healthy controls; B, PCA stratified by diabetes status, comparing stone patients with diabetes (n = 36, red) and stone patients without diabetes (n = 120, blue), and the extensive overlap of confidence ellipses suggests that diabetes has minimal impact on the urinary organic acid profile; C, PCA stratified by gender, comparing males (n = 89, blue) and females (n = 67, red), and the clear separation between groups indicates that gender is an important covariate affecting urinary organic acid profiles; D, PCA stratified by hypertension status, comparing stone patients with hypertension (n = 60, red) and stone patients without hypertension (n = 96, blue), and the extensive overlap suggests that hypertension is not a key covariate influencing the urinary organic acid profile; E, PCA comparing healthy controls (n = 57, blue), single-stone patients (n = 72, green), and recurrent-stone patients (n = 27, red), and the controls are clearly separated from both stone groups, while single-stone and recurrent-stone patients show substantial overlap, indicating similar metabolic profiles.

图 1　泌尿系结石患者与非结石患者尿有机酸代谢谱的主成分分析

Figure 1　Principal component analysis of urinary organic acid profiles in patients with urolithiasis and controls

为了最大化组间差异、筛选驱动该分离趋势的潜在生物标志物,并进一步验证组间代谢差异的统计学有效性,构建了 OPLS-DA 模型。OPLS-DA 得分图清晰地展示了结石组(红色)与对照组(蓝色)在预测主成分和正交主成分维度上实现了显著分离,两组样本点的分布区域仅在边缘有少量重叠,提示两组间的代谢谱差异具有统计学意义,存在驱动成分分离的关键代谢物。模型拟合参数为 $R2X$ = 0.139、$R2Y$ = 0.48,预测能力参数 $Q2$ = 0.457,表明该模型具有良好的解释力和预测稳定性,未出现明

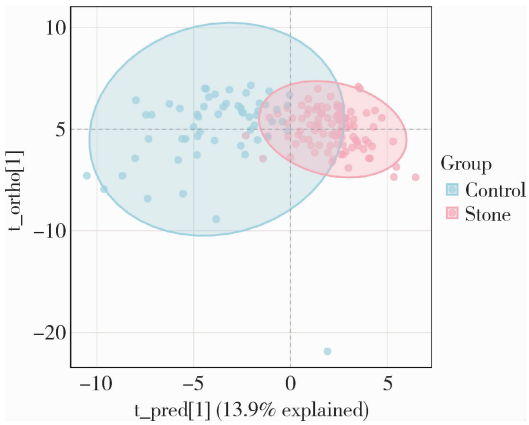
显的过拟合现象（图 2）。

进行分层分析以评估临床协变量对代谢谱的影响，按糖尿病状态分层的 PCA（图 1B）显示，合并糖尿病的结石患者（$n=36$）与未合并糖尿病的结石患者（$n=120$）样本高度重叠，提示糖尿病对尿有机酸代谢谱的影响较小。相比之下，按性别分层的 PCA（图 1C）显示男性（$n=89$）与女性（$n=67$）样本存在明显分离，提示性别是影响代谢谱的重要协变量。按高血压状态分层的 PCA（图 1D）显示，合并高血压的结石患者（$n=60$）与未合并高血压的结石患者（$n=96$）样本高度重叠，提示高血压并非尿有机酸代谢谱的关键影响因素。最后，对比非结石患者（$n=57$）、单发结石患者（$n=72$）与复发结石患者（$n=27$）的 PCA（图 1E）显示，非结石患者组与两类结石患者均明显分离，而单发与复发结石患者代谢谱高度重叠，提示结石是否复发对于尿有机酸代谢特征影响较小。

2.3 结石患者与非结石患者的差异代谢物筛选与核心贡献代谢物分析

基于 OPLS-DA 模型，对结石组与对照组的尿有机酸代谢谱进行了差异代谢物筛选。火山图分析（图 3A）显示，与对照组相比，结石组共鉴定出 44 种具有统计学差异的有机酸（$P<0.001$），其中 6 种表达显著上调，38 种表达显著下调。

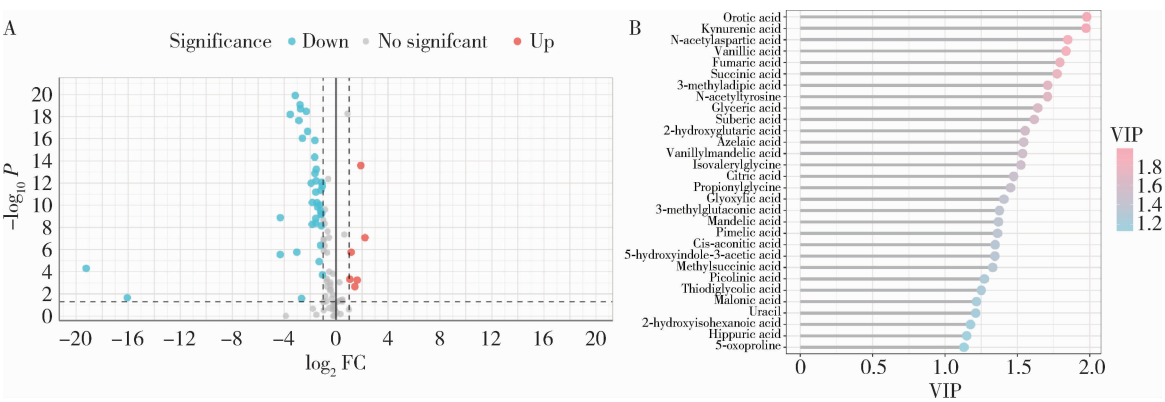
为识别驱动两组代谢分离的关键代谢物，进一步计算了其 VIP 值。图 3B 展示了 VIP 值排名前 30 的有机酸，其中对组间分离贡献度最高的前 10 种核心代谢物均表现为显著下调（$P<0.001$），按 VIP 值从高到低依次为乳清酸、犬尿喹啉酸、N-乙酰天冬氨酸、香草酸、富马酸、琥珀酸、3-甲基己二酸、N-乙酰酪氨酸、甘油酸和辛二酸。



The OPLS-DA model clearly demonstrates the distribution of the stone group (red) and control group (blue) along the predictive principal component (t_pred[1], explaining 13.9% of variance) and the orthogonal principal component (t_ortho[1]); A significant separation trend was observed between the two groups along the t_pred[1] axis, with only minor overlap at the edges, indicating substantial differences in urinary organic acid metabolic profiles; The model fit parameters were $R2X=0.139$, $R2Y=0.48$, and the predictive ability parameter $Q2=0.457$, indicating good explanatory power, predictive stability, and no significant overfitting. OPLS-DA, orthogonal partial least squares discriminant analysis.

图 2 泌尿系结石患者与非结石患者尿有机酸代谢谱的正交偏最小二乘判别分析得分图

Figure 2 Orthogonal partial least squares discriminant analysis score plot of urinary organic acid profiles in urolithiasis patients and controls



A, compared with the control group, a total of 44 organic acids with significant differences were identified in the stone group ($P<0.05$), including 6 significantly up-regulated (red) and 38 significantly down-regulated (blue) metabolites; B, the top 30 organic acids with the highest contribution to metabolic separation between groups are shown; The top 10 key metabolites ranked by VIP value are: Orotic acid, kynurenic acid, N-acetylaspartic acid, vanillic acid, fumaric acid, Succinic acid, 3-methyladipic acid, N-acetyltyrosine, glyceric acid, and suberic acid, which are the critical molecules driving the metabolic differences between the two groups. VIP, variable importance in projection; FC, fold change.

图 3 泌尿系结石患者与非结石患者差异有机酸筛选及核心代谢物贡献度分析

Figure 3 Screening of differentially expressed organic acids and contribution analysis of key metabolites in urolithiasis patients *vs.* controls

2.4 差异有机酸 KEGG 通路富集分析

为了探索结石组与对照组之间差异代谢物背后的生物学功能，利用 KEGG 数据库对筛选出的差异有机酸进行了代谢通路富集分析。

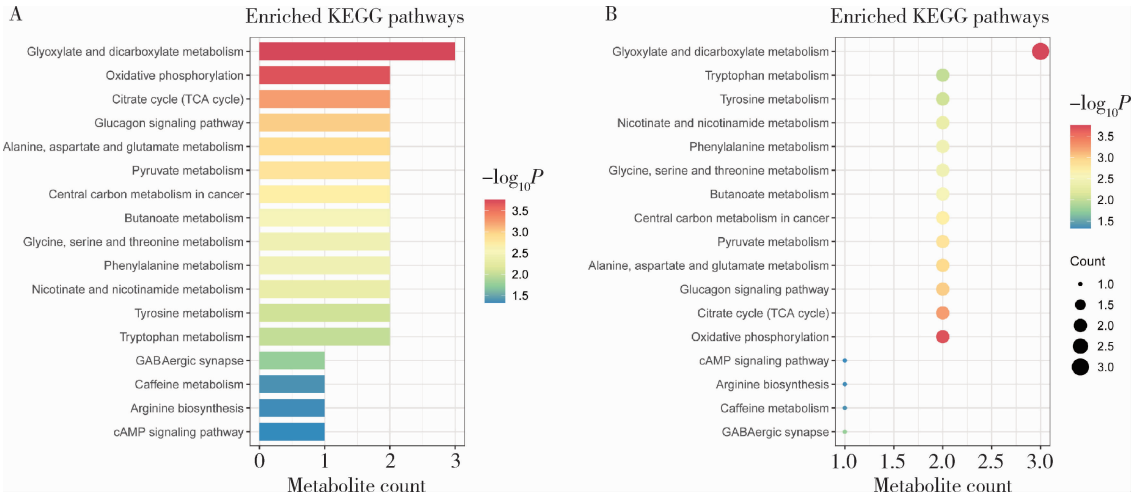
根据统计学显著性排序，展示了富集程度最高的代谢通路（图 4），其中，图 4A 为富集通路的柱状

图,图 4B 为对应的气泡图,两者共同呈现了差异代谢物的功能富集特征。

分析结果表明,泌尿系结石相关的代谢紊乱主要集中在氨基酸代谢和能量代谢相关通路,其中,乙醛酸和二羧酸代谢、氧化磷酸化以及三羧酸循环是富集程度最显著的三条通路,此外,色氨酸代谢、酪氨酸代谢、苯丙氨酸代谢、丙氨酸-天冬氨酸-谷氨酸代谢,以及丁酸代谢也显示出显著改变,提示氨基酸代谢通路的异常在结石形成中同样发挥重要作用。

2.5　机器学习筛选核心代谢物及区分效能验证

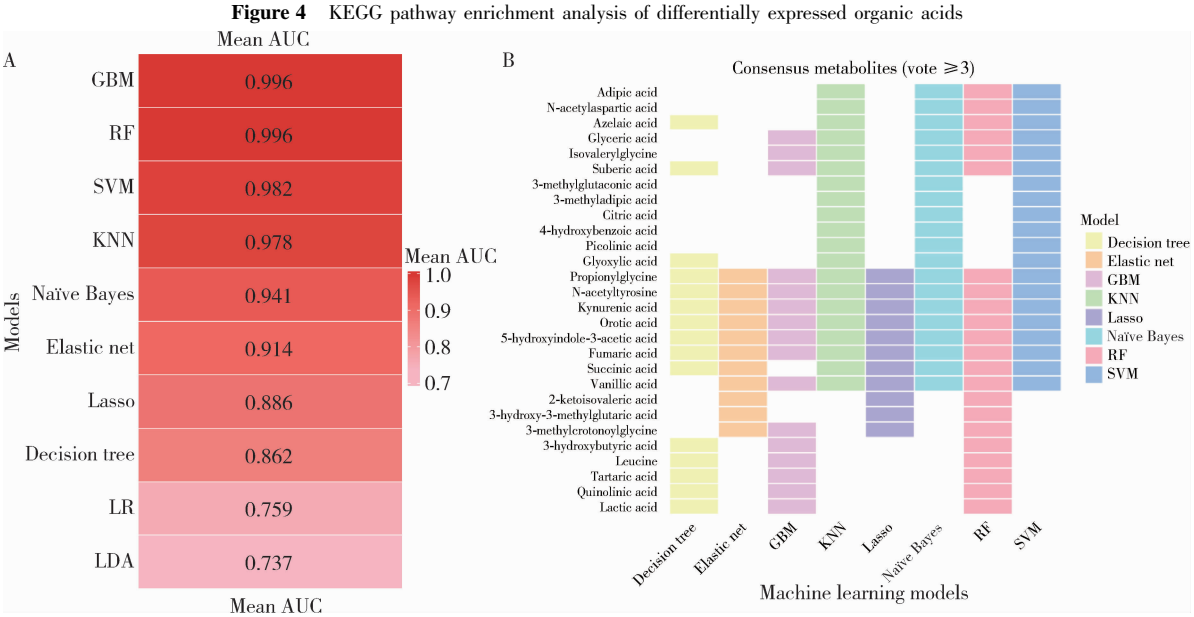
为评估差异有机酸作为泌尿系结石标志物的潜力,本研究构建了 10 种机器学习模型并比较其效能(图 5)。GBM 与 RF 的平均 AUC 值最高(图 5A),均为 0.996;其次为 SVM(AUC 值为 0.982)和 KNN(AUC 值为 0.978),提示基于尿有机酸代谢物可高效区分结石患者与非结石患者。



A, bar chart showing the KEGG pathways significantly enriched by the differentially expressed organic acids, the results show that glyoxylate and dicarboxylate metabolism, oxidative phosphorylation, and the citrate cycle (TCA cycle) are the most significantly enriched pathways, suggesting that the development of urolithiasis is closely related to disturbances in energy metabolism and central carbon metabolism; B, bubble plot further illustrating the characteristics of the enriched pathways, the glyoxylate and dicarboxylate metabolism pathway has the highest number of metabolites and the strongest significance, while oxidative phosphorylation and the citrate cycle also show significant enrichment, further confirming the central role of energy and central carbon metabolism in the pathogenesis of urolithiasis. KEGG, kyoto encyclopedia of genes and genomes; TCA, tricarboxylic acid cycle; GABAergic, gamma-aminobutyric acid-ergic.

图 4　差异有机酸富集的 KEGG 通路分析

Figure 4　KEGG pathway enrichment analysis of differentially expressed organic acids



A, bar chart of mean AUC values for different machine learning models, GBM and RF achieved the highest mean AUC of 0.996, followed by SVM, and KNN, indicating that urinary organic acid metabolites can effectively distinguish stone patients from healthy controls; B, heatmap of consensus metabolites (vote ⩾ 3); Key metabolites consistently selected by at least 3 models are shown, including orotic acid, kynurenic acid, succinic acid, and fumaric acid, which serve as core biomarkers for the diagnosis of urolithiasis. VIP, variable importance in projection; FC, fold change; AUC, area under the curve; GBM, gradient boosting machine; RF, fandom forest; SVM, support vector machine; KNN, K-nearest neighbors; LR, Logistic regression; LDA, linear discriminant analysis.

图 5　基于差异有机酸的泌尿系结石机器学习模型分析

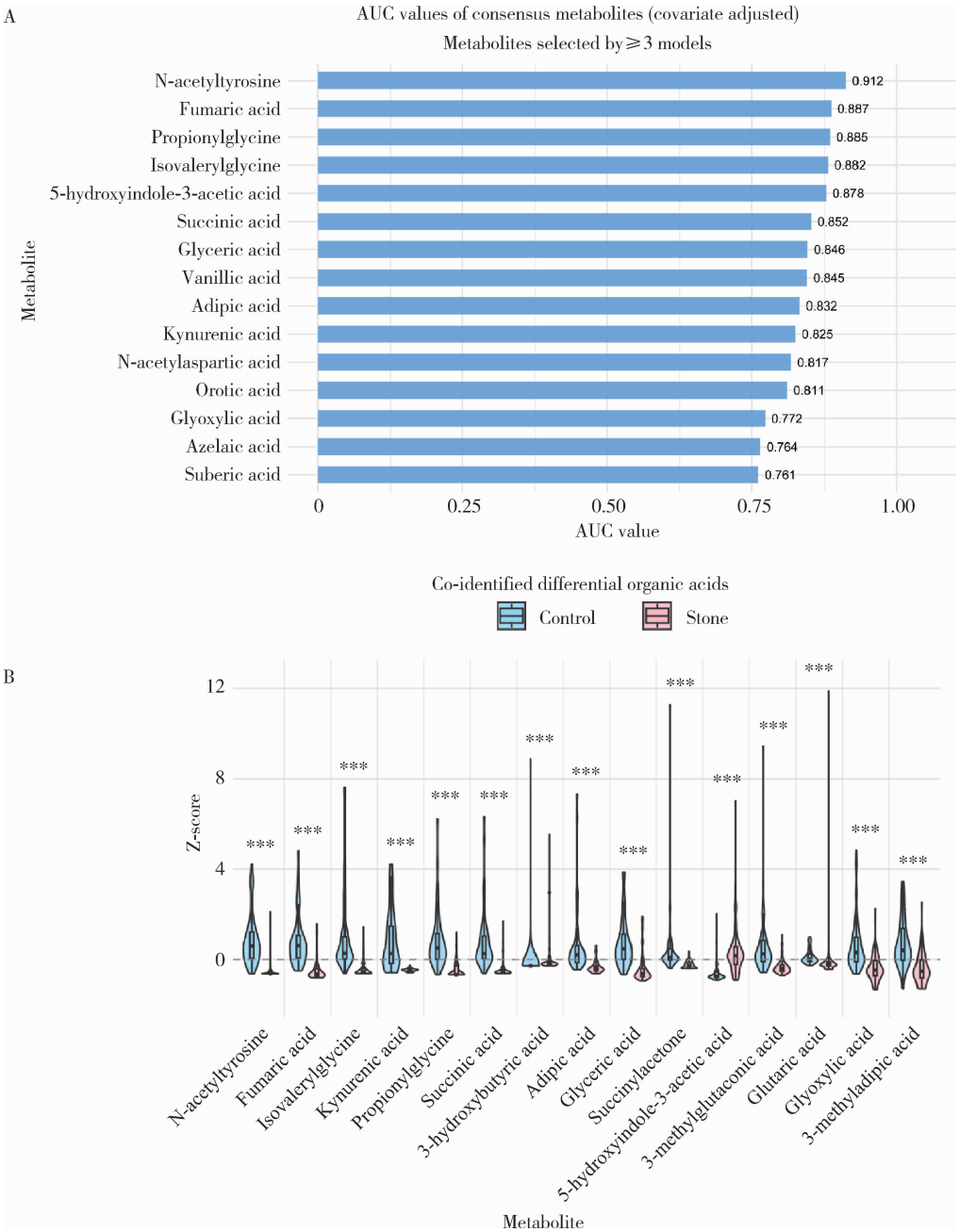
Figure 5　Machine learning analysis of urinary calculi using differential organic acids

随后的共识代谢物分析（图 5B）筛选出在至少 3 种模型中被一致识别的关键代谢物，包括乳清酸、犬尿喹啉酸、琥珀酸，富马酸、N-乙酰酪氨酸等，这些代谢物可以作为区分结石患者与非结石患者的核心生物标志物。

2.6 共识代谢物的区分效能及表达验证

为验证共识代谢物在区分结石患者上的价值，校正了性别等协变量后，进一步评估了其 AUC 值及组间表达差异。如图 6A 所示，N-乙酰酪氨酸（AUC = 0.912）、富马酸（AUC = 0.887）、丙酰甘氨酸（AUC = 0.885）、异戊酰甘氨酸（AUC = 0.882）、5-羟基吲哚-3-乙酸（AUC = 0.878）等均表现出较高的区分效能，提示这些代谢物可作为泌尿系结石机制研究的潜在靶点。

小提琴图分析（图 6B）进一步证实，所有共识别的差异有机酸（包括 N-乙酰酪氨酸、富马酸、犬尿喹啉酸、琥珀酸等）在结石组与对照组间均存在显著表达差异（$P < 0.05$），进一步验证了其作为泌尿系结石的生物标志物的可靠性。



A, bar chart of AUC values for consensus metabolites after covariate adjustment, N-acetyltyrosine achieved the highest AUC value of 0.912, while fumaric acid (0.887) and propionylglycine (0.885) also exhibited high diagnostic efficacy, indicating their potential as independent diagnostic biomarkers for urolithiasis; B, violin plot of co-identified differential organic acids, blue represents the control group, and pink represents the stone group, with Z-score as the normalized expression level, all displayed metabolites showed significant differences between the two groups ($P < 0.001$), further validating their reliability as diagnostic biomarkers for urolithiasis. ***$P < 0.001$. AUC, area under the curve.

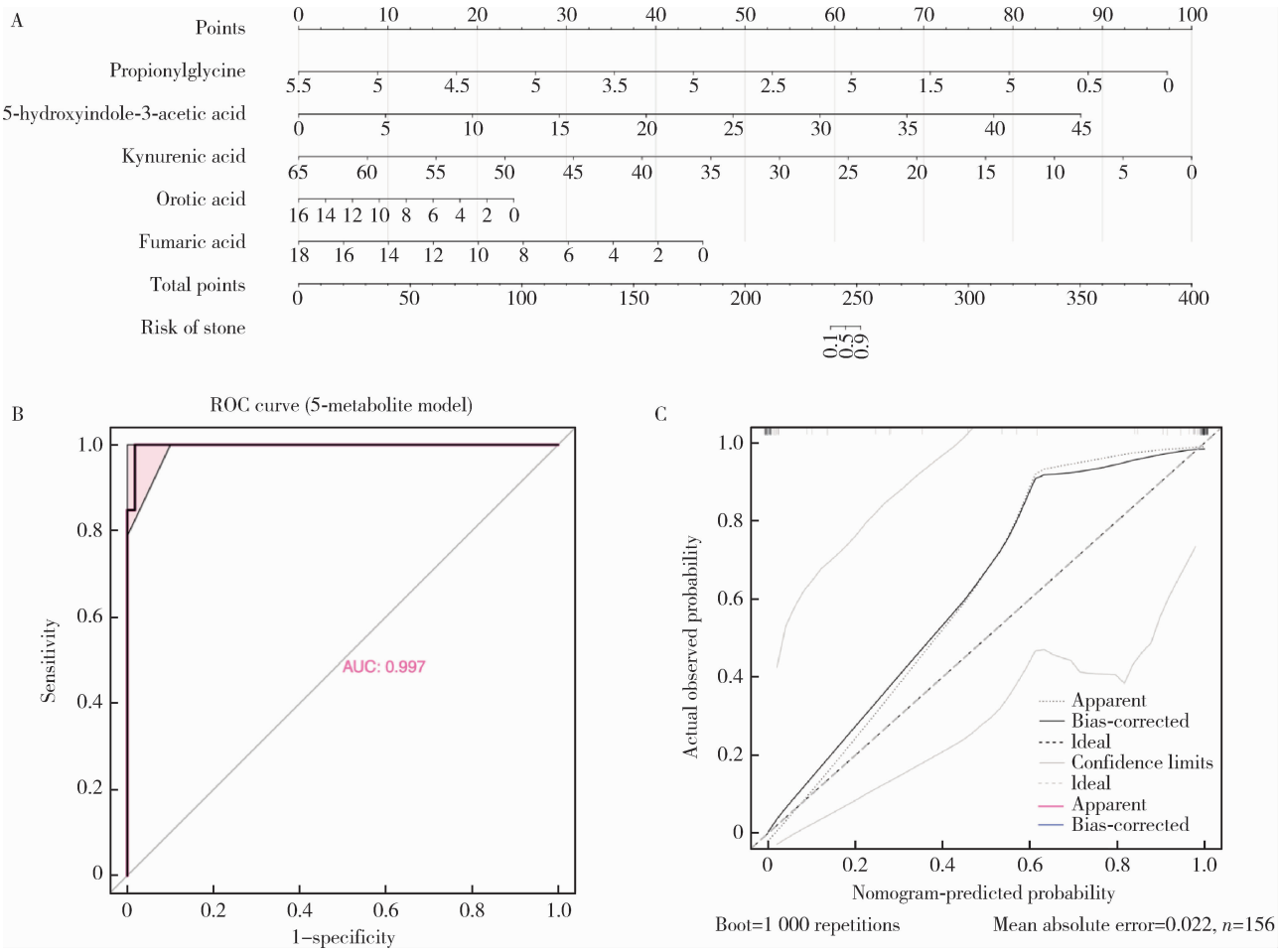
图 6 共识代谢物的诊断效能及组间表达差异分析

Figure 6 AUC plot and violin plot of consensus urinary organic acids

2.7　核心代谢物组合的筛选与效能验证

为进一步从差异代谢物中筛选出生物学意义最突出的核心代谢标志物,以更好地阐释泌尿系结石相关的代谢紊乱特征,本课题组从前期筛选出的共识代谢物中,筛选出丙酰甘氨酸、5-羟基吲哚-3-乙酸、犬尿喹啉酸、乳清酸和富马酸 5 种核心代谢物,构建了整合模型。经多组代谢物组合的效能对比验证,该 5 种核心代谢物组合表现出最优的区分效能,其综合区分能力显著优于单一代谢物及其他代谢物组合,可作为反映结石相关代谢异常的核心分子组合。图 7A 所示的列线图整合了上述 5 种代谢物的表达水平,每个代谢物对应相应的得分区间,通过将 5 种代谢物的单项得分累加得到总分,进而量化发生泌尿系结石的可能。ROC 曲线分析显示,该 5 种代谢物构建的列线图模型的 AUC 高达 0.997,提示其能够有效区分结石患者与非结石患者(图 7B)。

为进一步验证其可靠性,本研究构建了该模型的校准曲线,校正偏倚后的曲线与理想曲线高度吻合,平均绝对误差仅为 0.022(图 7C),表明该列线图模型具有良好的校准度与准确性,能够较为准确地区分结石患者与非结石患者,证实上述 5 种代谢物的组合模式稳定可靠,可作为后续机制研究与靶点验证的关键代谢分子。



A, metabolite combination panel: Integrates the expression levels of five metabolites, propionylglycine, 5-Hydroxyindole-3-acetic acid, kynurenic acid, orotic acid, and fumaric acid to calculate a total score for reflecting metabolic abnormality associated with urolithiasis; B, ROC curve of the 5-metabolite model, the area under the curve is 0.997, indicating excellent performance in distinguishing patients with urolithiasis from controls; C, calibration curve, the bias-corrected curve closely aligns with the ideal curve, with a mean absolute error of 0.022, demonstrating good calibration and stability of the metabolite combination model. ROC, receiver operating characteristic; AUC, area under the curve.

图 7　核心代谢物组合对泌尿系结石的区分效能分析

Figure 7　Discrimination efficiency of core metabolite panel for urolithiasis

3　讨论

泌尿系结石是泌尿外科常见疾病,其发病机制复杂,与遗传、环境、代谢等多种因素密切相关。目前已有研究表明,结石的生成与尿代谢物谱的改变有关,但是没有明确的研究关注结石的发生与尿有机酸代谢改变以及其所处代谢通路变化的关联[10-12],并且目前临床缺乏结石相关的精准的早期

诊断标志物及便捷的风险预测工具。本研究通过对泌尿系结石患者与非结石患者的尿有机酸代谢谱进行系统分析,结合无监督及有监督统计分析、机器学习、通路富集分析及列线图建模,全面揭示了结石患者的有机酸特征,筛选出具有诊断价值的核心有机酸,并通过组合效能比较证实其具有最强的组间区分能力与代谢代表性,可作为后续机制研究与靶点验证的关键代谢分子组合。

本研究首先通过 PCA 分析发现,泌尿系结石患者与非结石患者的尿有机酸代谢谱存在显著差异,这一结果证实结石患者存在明显的尿有机酸代谢异常,与既往代谢组学相关研究结论相一致[13]。为明确临床协变量对尿有机酸代谢谱的影响,本研究进一步开展分层 PCA 分析,结果显示性别是影响代谢谱的重要因素,男女群体的尿有机酸代谢谱呈现明显的分离趋势;而本研究中结石组男女比例为 3∶1、非结石组为 1∶1 的人群分布特征,与泌尿系结石的流行病学研究结果相符[2-3],这提示不同性别间的尿有机酸代谢谱差异,可能是导致结石发病存在性别不均衡的重要原因之一,也为后续泌尿系结石代谢标志物的筛选研究提供了关键参考,即需重点校正性别这一混杂因素,以此提升标志物的特异性与可靠性,避免性别差异造成的研究偏差。

基于 OPLS-DA 模型筛选了差异代谢物,本研究共鉴定出 44 种具有统计学意义的差异有机酸($P < 0.05$),其中 6 种在结石患者中表达上调,38 种表达下调,提示泌尿系结石患者存在广泛的尿有机酸代谢异常,且以代谢物表达下调为主要特征。经 VIP 分析进一步筛选出 10 种 VIP 值最高的核心差异代谢物,包括乳清酸、犬尿喹啉酸、N-乙酰天冬氨酸、香草酸、富马酸、琥珀酸等,且该 10 种核心代谢物在结石患者中均呈下调趋势,其表达异常可能通过多种途径参与结石的形成过程。其中,犬尿喹啉酸作为色氨酸代谢的重要产物,已有研究证实其可调节组织炎症反应、保护肾小球功能[14-16],且肾癌、肾移植及慢性肾病患者中均存在犬尿氨酸代谢异常的现象[17],推测其代谢紊乱可能通过调控炎症通路参与泌尿系结石的发生发展。琥珀酸则被证实可通过抑制结石成核,减少肾钙沉积与肾组织损伤,从而降低结石发病风险[9];富马酸作为三羧酸循环的关键中间产物,可通过影响琥珀酸、枸橼酸等有机酸的体内浓度,间接调控结石形成过程。

为深入探究差异有机酸背后的生物学机制,本研究对筛选出的差异代谢物开展 KEGG 通路富集分析,结果显示,差异有机酸主要显著富集于乙醛酸和二羧酸代谢、氧化磷酸化及三羧酸循环(tricarboxylic acid cycle, TCA)等能量代谢通路,同时酪氨酸代谢、苯丙氨酸代谢、丙氨酸-天冬氨酸-谷氨酸代谢等氨基酸代谢通路也呈现明显的异常表达。其中,乙醛酸和二羧酸代谢通路可直接调控尿液中草酸浓度,现有研究已证实尿液草酸浓度与结石形成密切相关,草酸浓度升高会显著增加结石发病风险[18-20]。氧化磷酸化与三羧酸循环是细胞能量代谢的核心通路,其代谢紊乱可能与肾小管上皮细胞损伤相关,其一方面会促进结石结晶的黏附与沉积,另一方面还会诱发局部炎症反应,双重作用下进一步增加结石形成风险;同时该通路中间产物的表达异常也会直接影响结石生成,如枸橼酸作为尿液中重要的结石形成抑制剂,提高其浓度可有效降低草酸钙、尿酸结石的发病风险[21],本研究中结石组尿液枸橼酸浓度呈显著降低趋势,与现有研究结论高度一致。而琥珀酸、富马酸等本研究发现的核心差异代谢物,均为该通路的关键中间产物。此外,氨基酸代谢通路的异常同样参与结石的发生发展过程,如色氨酸代谢通路可通过调控犬尿喹啉酸的表达,经炎症途径影响结石形成[22]。上述通路分析结果,进一步完善了泌尿系结石的代谢发病机制,也为后续从代谢调控角度开发结石防治靶点提供了新的研究思路。

本研究构建了 10 种机器学习模型并进行效能对比,基于这 10 种模型,筛选出丙酰甘氨酸、5-羟基吲哚-3-乙酸、犬尿喹啉酸、乳清酸和富马酸 5 种核心代谢物,该 5 种代谢物组合的区分效能显著优于单一代谢物及其他组合,能够较好地代表泌尿系结石相关的特征性代谢紊乱,为后续结石发病机制的深入探索、关键通路验证及潜在干预靶点筛选提供了可靠且高效的代谢分子靶标。以上 5 种代谢物分别参与色氨酸代谢、能量代谢、乙醛酸代谢等结石相关通路,其协同异常提示泌尿系结石的发生并非单一代谢的改变,而是多通路、多环节共同驱动的复杂疾病。未来可基于本研究筛选出的核心代谢物,进一步开展体内外实验,验证其在草酸钙结晶形成、肾小管上皮损伤、肾炎症反应及氧化应激中的具体作用与调控机制,明确其在结石发生中的生物学功能。同时,可围绕关键代谢物开展靶向干预研究,探索代谢调控策略在预防结石形成、降低复发风险中的应用潜力,为开发新型代谢靶向防治方案提供实验基础与理论支撑。

本研究作为病例对照研究,虽明确了尿有机酸谱的差异及相关代谢通路,但仍存在一定局限性:首

先,本研究为单中心研究,样本量相对有限(结石组99例、对照组57例),可能存在选择偏倚,后续需扩大样本量、开展多中心研究,纳入不同地区、不同人群的样本,进一步验证核心代谢物及列线图模型的有效性与适用性;其次,本研究仅聚焦于尿有机酸代谢谱,未结合其他代谢组学(如脂质组学、糖代谢组学)、转录组学及蛋白质组学数据,未能全面揭示泌尿系结石发病的代谢调控网络,后续可通过多组学联合分析,深入探索代谢物异常背后的分子机制;最后,本研究未对结石成分(如草酸钙、尿酸、胱氨酸等)进行细分,不同成分结石的有机酸代谢特征可能存在差异,后续需按结石类型分层分析。

综上所述,本研究通过系统的尿有机酸代谢组学分析,证实泌尿系结石患者存在显著的尿有机酸代谢紊乱,性别是影响代谢谱的重要协变量,而糖尿病、高血压及结石复发状态对代谢谱影响较小;研究筛选出44种差异有机酸及10种核心差异代谢物,明确差异有机酸主要富集于能量代谢及氨基酸代谢相关通路,揭示了代谢紊乱与泌尿系结石发病机制的密切关联;通过机器学习模型筛选出5种核心代谢物(丙酰甘氨酸、5-羟基吲哚-3-乙酸、犬尿喹啉酸、乳清酸、富马酸),该组合可高效、稳定地区分结石患者与非结石人群,模型评估显示其具有优异的区分效能(AUC = 0.997)与可靠性(平均绝对误差 = 0.022),可作为反映结石相关代谢异常的核心分子组合。本研究从代谢组学视角系统解析了泌尿系结石的代谢特征与潜在机制,为深入阐明疾病发病机制、挖掘关键代谢调控靶点及开展后续机制研究提供了理论依据与研究方向。

利益冲突　所有作者均声明不存在利益冲突。

作者贡献声明　漆子暄:设计实验方案,收集、整理数据,撰写论文;边小龙:设计研究方案,分析、整理数据,撰写论文;胡浩浦、田聪、王辰龙、周雨缪:收集整理数据;徐涛、王辉:设计试验方案;曹林林:提出研究思路,审定论文;胡浩:提出研究思路,总体把关和审定论文。所有作者均参与论文修改,并对最终文稿进行审读和确认。

参考文献

[1] Shoag J, Tasian GE, Goldfarb DS, et al. The new epidemiology of nephrolithiasis[J]. Adv Chronic Kidney Dis, 2015, 22(4): 273 – 278.

[2] Ziemba JB, Matlaga BR. Epidemiology and economics of nephrolithiasis[J]. Investig Clin Urol, 2017, 58(5): 299.

[3] Sorokin I, Mamoulakis C, Miyazawa K, et al. Epidemiology of stone disease across the world[J]. World J Urol, 2017, 35(9): 1301 – 1320.

[4] Tang X, Lieske JC. Acute and chronic kidney injury in nephrolithiasis[J]. Curr Opin Nephrol Hypertens, 2014, 23(4): 385 – 390.

[5] Bargagli M, Scoglio M, Howles SA, et al. Kidney stone disease: Risk factors, pathophysiology and management[J]. Nat Rev Nephrol, 2025, 21(11): 794 – 808.

[6] Kim HN, Kim JH, Chang Y, et al. Gut microbiota and the prevalence and incidence of renal stones[J]. Sci Rep, 2022, 12(1): 3732.

[7] Lai Y, Zheng H, Sun X, et al. The advances of calcium oxalate calculi associated drugs and targets[J]. Eur J Pharmacol, 2022, 935: 175324.

[8] Peerapen P, Thongboonkerd V. Kidney stone prevention[J]. Adv Nutr, 2023, 14(3): 555 – 569.

[9] Zhang XZ, Lei XX, Jiang YL, et al. Application of metabolomics in urolithiasis: The discovery and usage of succinate[J]. Sig Transduct Target Ther, 2023, 8(1): 41.

[10] Duan X, Zhang T, Ou L, et al. ^{1}H NMR-based metabolomic study of metabolic profiling for the urine of kidney stone patients[J]. Urolithiasis, 2020, 48(1): 27 – 35.

[11] Kaneko K, Kobayashi R, Yasuda M, et al. Comparison of matrix proteins in different types of urinary stone by proteomic analysis using liquid chromatography-tandem mass spectrometry[J]. Int J Urol, 2012, 19(8): 765 – 772.

[12] Boonla C, Tosukhowong P, Spittau B, et al. Inflammatory and fibrotic proteins proteomically identified as key protein constituents in urine and stone matrix of patients with kidney calculi[J]. Clin Chim Acta, 2014, 429: 81 – 89.

[13] Wang Z, Zhang Y, Zhang J, et al. Recent advances on the mechanisms of kidney stone formation (Review)[J]. Int J Mol Med, 2021, 48(2): 149.

[14] Bolanos-Palmieri P, Schenk H, Bähre H, et al. Kynurenine pathway dysregulation impairs podocyte morphology and bioenergetics *in vitro* and leads to glomerular dysfunction[J]. FASEB J, 2025, 39(22): e71228.

[15] Liu JR, Miao H, Deng DQ, et al. Gut microbiota-derived tryptophan metabolism mediates renal fibrosis by aryl hydrocarbon receptor signaling activation[J]. Cell Mol Life Sci, 2021, 78(3): 909 – 922.

[16] Agudelo LZ, Ferreira DMS, Cervenka I, et al. Kynurenic acid and Gpr35 regulate adipose tissue energy homeostasis and inflammation[J]. Cell Metab, 2018, 27(2): 378 – 392. e5.

[17] Zakrocka I, Załuska W. Kynurenine pathway in kidney diseases[J]. Pharmacol Rep, 2022, 74(1): 27 – 39.

[18] Sadaf H, Raza SI, Hassan SW. Role of gut microbiota against calcium oxalate[J]. Microb Pathog, 2017, 109: 287 – 291.

[19] Song Q, Song C, Chen X, et al. Oxalate regulates crystal-cell adhesion and macrophage metabolism via JPT2/PI3K/AKT signaling to promote the progression of kidney stones[J]. J Pharm Anal, 2024, 14(6): 100956.

[20] Tavasoli S, Alebouyeh M, Naji M, et al. Association of intestinal oxalate-degrading bacteria with recurrent calcium kidney stone formation and hyperoxaluria: A case-control study[J]. BJU Int, 2020, 125(1): 133 – 143.

[21] Cao C, Li F, Ding Q, et al. Potassium sodium hydrogen citrate intervention on gut microbiota and clinical features in uric acid stone patients[J]. Appl Microbiol Biotechnol, 2024, 108(1): 51.

[22] Gao H, Lin J, Xiong F, et al. Urinary microbial and metabolomic profiles in kidney stone disease[J]. Front Cell Infect Microbiol, 2022, 12: 953392.

(2026-03-02 收稿)
(本文编辑:王　蕾)