中山大學 中山眼科中心 ZHONGSHAN OPHTHALMIC CENTER, SUN YAT-SEN UNIVERSITY

全文收录于 Lighthouse Scholar

眼科学报

官网： journal.gzzoc. com



doi: 10.12419/24101607
View this article at: https://dx.doi.org/10.12419/24101607

·眼底影像专栏·

基于机器学习的代谢组学探索 2 型糖尿病及糖尿病视网膜病变进展潜在生物标志物

庞雨卉，罗朝琨，张青箬，张雄泽，廖南英，吉宇莹，米兰，甘雨虹，苏永悦，文峰，陈卉

（中山大学中山眼科中心，眼病防治全国重点实验室，广东省眼科视觉科学重点实验室，广州 510060）

文章亮点

1. 关键发现

- 本研究结合代谢组学和机器学习方法，成功识别出与 2 型糖尿病 (type 2 diabetes mellitus, T2DM) 及糖尿病视网膜病变 (diabetic retinopathy, DR) 进展相关的关键代谢生物标志物。1, 5- 脱水山梨醇、硫酸十四烷基酯和 N, N, N 三甲基 -5- 氨基戊酸在 T2DM 患者中表现出显著差异，亚油酰胺和油酰胺在 DR 不同进展阶段表达差异显著。

2. 已知与发现

- 代谢组学方法已被广泛应用于揭示糖尿病及其并发症的代谢特征，但具体的代谢通路和生物标志物仍有待研究。现有研究大多集中于揭示代谢物与非增殖期 DR 或增殖期 DR 的关联，而对 DR 进展过程中代谢变化的系统性探讨仍显不足。本研究通过对 DR 各阶段的代谢物进行系统分析，发现亚油酰胺和油酰胺在 DR 进展中存在显著差异，为 DR 发病机制的探索提供了新的研究视角和见解。

3. 意义与改变

- 本研究识别了与 T2DM 及 DR 进展相关的关键代谢生物标志物，特别是亚油酰胺和油酰胺，它们在 DR 进展中的显著变化突显了作为潜在生物标志物的重要性，为 DR 的诊断和个体化治疗提供了新的视角。

收稿日期(Received): 2024-10-19; **修订日期(Revised)**: 2025-01-16; **接收日期(Accepted)**: 2025-02-11

通信作者(Corresponding author): 陈卉，Email:chh5413@126.com

基金项目 (Foundation item)：国家自然科学基金面上项目(81970813)，广东省自然科学基金(2023A1515011198)，广州市科学技术计划市校联合项目(SL2022A03J00553)。 This work was supported by the National Natural Science Foundation of China (81970813), Natural Science Foundation of Guangdong Province (2023A1515011198), the Guangzhou Municipal Science and Technology Program (SL2022A03J00553).

作者贡献声明: 构思设计: 陈卉, 庞雨卉; 行政支持: 陈卉, 张雄泽; 提供研究材料: 陈卉, 张雄泽, 廖南英, 吉宇莹, 米兰, 甘雨虹, 苏永悦, 文峰; 数据收集与汇总: 庞雨卉, 罗朝琨, 张青箬; 数据分析与解读: 陈卉, 庞雨卉; 稿件撰写: 陈卉, 庞雨卉; 稿件最终修订与同意: 所有作者

Contributions: Conception and design: CHEN Hui, PANG Yuhui; Administrative support:CHEN Hui, ZHANG Xiongze; Provision of study materials: CHEN Hui, ZHANG Xiongze, LIAO Nanying, JI Yuying, MI Lan, GAN Yuhong, SU Yongyue, WEN Feng; Collection and assembly of data: CHEN Hui, PANG Yuhui; Data analysis and interpretation: PANG Yuhui, LUO Chaokun, ZHANG Qingruo; Manuscript writing: CHEN Hui, PANG Yuhui ; Final approval of manuscript: All authors

电子全文

[摘　要] **目的：**利用超高效液相色谱串联四极杆-静电场轨道阱高分辨质谱(ultra-high performance liquid chromatography tandem quadrupole-electrostatic field orbitrap high resolution mass spectrometer, UHPLC- HRMS)代谢组学技术结合机器学习识别与糖尿病视网膜病变(diabetic retinopathy, DR)进展过程中的房水代谢差异，以寻找DR进展相关生物标志物。**方法：**本研究共纳入78例2型糖尿病(type 2 diabetes mellitus, T2DM)患者以及30名年龄性别匹配健康对照人群。使用UHPLC- HRMS检测所有患者及对照人群房水中的代谢物丰度，结合机器学习筛选T2DM和DR进展相关代谢物标志物并建立预测模型。**结果：**在校正混杂因素后，与健康人群对比1, 5-脱水山梨醇、硫酸十四烷基酯和N,N,N-三甲基-5-氨基戊酸在T2DM患者中表现出显著差异(均P<0.05)；而N-乙酰色氨酸、亚油酰胺、油酰胺、棕榈酰胺、戊酸(游离脂肪酸(5:0)和琥珀酸与DR进展显著相关(均P<0.001)。代谢通路分析表明，“缬氨酸、亮氨酸和异亮氨酸的生物合成”“精氨酸的生物合成”和“半胱氨酸及蛋氨酸代谢”是T2DM差异代谢途径。基于生物标志物的随机森林预测模型显示，差异代谢产物对T2DM和DR进展的预测准确率分别为81.3%和74%。**结论：**代谢组学结合机器学习方法有效揭示了T2DM及与DR进展相关的代谢特征，亚油酰胺和油酰胺有望成为DR进展的生物标志物，为DR的诊断和个体化治疗提供了新的可能性。

[关键词] 2型糖尿病；糖尿病视网膜病变；代谢组学；生物标志物；机器学习

[中图分类号] R587.1，R771.3　　[文献标志码] A

引用本文: 庞雨卉, 罗朝琨, 张青箬, 张雄泽, 廖南英, 吉宇莹, 米兰, 甘雨虹, 苏永悦, 文峰, 陈卉. 基于机器学习的代谢组学探索2型糖尿病及糖尿病视网膜病变进展潜在生物标志物[J]. 眼科学报, 2025, 40(4): 303-314. DOI: 10.12419/24101607.

Metabolomic integration with machine learning identifies potential biomarkers for type 2 diabetes mellitus and diabetic retinopathy progression

PANG Yuhui, LUO Chaokun, ZHANG Qingruo, ZHANG Xiongze, LIAO Nanying, JI Yuying, MI Lan, GAN Yuhong, SU Yongyue , WEN Feng, CHEN Hui

(State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangdong Provincial Key Laboratory of Ophthalmology Visual Science, Guangzhou 510060, China)

Abstract **Objective:** To identify aqueous humor metabolic profiles associated with the progression of type 2 diabetes mellitus (T2DM) and diabetic retinopathy (DR), aiming to discover potential biomarkers for DR progression. Ultra-high performance liquid chromatography tandem quadrupole-electrostatic field orbitrap high-resolution mass spectrometry (UHPLC-HRMS) will be utilized in conjunction with machine learning (ML) for comprehensive analysis. **Methods:** A total of 78 patients with T2DM and 30 age- and gender-matched healthy controls were included. UHPLC-HRMS was used to identify metabolites in the aqueous humor of all participants. ML was employed to screen for metabolites associated with T2DM and DR progression, and predictive models were established. **Results:** After adjusting for covariates, 1,5-anhydroglucitol, tetradecyl sulfate, and n,n,n-trimethyl-5-aminovaleric acid identified as significant indicators for T2DM compared to controls (all P<0.05). N-acetyltryptophan, linoleamide, oleamide, palmitic amide, valeric acid(FFA(5:0), and succinic acid emerged as predictors for DR progression (all P<0.001). Metabolic pathway analysis revealed that "valine, leucine and isoleucine biosynthesis", "arginine biosynthesis," and "cysteine and methionine

metabolism" were the most enriched pathways for T2DM. Predictive models achieved R^2 values of 81.3%, and 74% for T2DM and DR progression, respectively. **Conclusions:** Metabolomic combined with ML effectively uncovered metabolic characteristics associated with T2DM and DR progression. Linoleamide and oleamide represent promising potential biomarkers for DR progression, offering new opportunities for diagnosis and personalized treatment of DR.

Keywords type 2 diabetes mellitus; diabetic retinopathy; metabolomics; biomarkers; machine learning

Cite this article as: PANG Yuhui, LUO Chaokun, ZHANG Qingruo, ZHANG Xiongze, LIAO Nanying, JI Yuying, MI Lan, GAN Yuhong, SU Yongyue , WEN Feng, CHEN Hui. Metabolomic integration with machine learning identifies potential biomarkers for type 2 diabetes mellitus and diabetic retinopathy progression[J]. Yan Ke Xue Bao, 2025, 40(4): 303-314. DOI: 10.12419/24101607.

2型糖尿病(type 2 diabetes mellitus, T2DM)是一种因胰岛素不足引起的以高血糖为特征的代谢性疾病。全球糖尿病患病率逐年上升，根据一项涵盖1 108项群体代表性研究和1.41亿参与者的汇总分析，2022年全世界患有糖尿病的成年人总数高达8.28亿[1]。糖尿病视网膜病变(Diabetic Retinopathy, DR)是T2DM最常见和最严重的眼部微血管并发症[2]。在中国，18～74岁糖尿病患者中，DR的患病率为16.3%，其中严重DR的患病率为3.2%，包括严重的非增殖期DR(non-proliferative diabetic retinopathy, NPDR)、增殖期DR(proliferative diabetic retinopathy, PDR)和临床显著的糖尿病黄斑水肿[3]。全球范围内，DR的患病率约为34.8%，发展中国家的发病率略高于发达国家。DR是导致工作年龄人群视力障碍和失明的主要原因之一[4]。糖化血红蛋白水平虽然常用于监测糖尿病患者的血糖控制，但其与DR进展之间的相关性尚不显著[5]。虽然已有研究提出了一些潜在的生物标志物，能够反映DR的发生与发展，但现有的代谢谱仍存在诸多局限。因此，亟需进一步探索新的生物标志物，以便更有效地预防和控制DR的发生与进展。

代谢组学是对某一既定时期生物体内样本所有小分子代谢物(<1 200 Da)进行定量和定性的分析，并研究这些代谢物在疾病生理条件下的动态变化规律。通过代谢组学分析，我们可以从整体代谢水平研究疾病的动态变化并识别与之相关的潜在生物标志物[6]。DR的进展包括了多个阶段[如糖尿病无糖尿病视网膜病变期(non-diabetic retinopathy, NDR)、NPDR 和PDR]，而各阶段中代谢物水平的变化能够反映出病理过程中细胞、组织和器官的代谢状态。异常的代谢状态通常与炎症反应、氧化应激和微血管功能障碍等病理机制密切相关[7]。脂质和氨基酸代谢物的变化已被发现与视网膜内皮功能障碍相关，这些代谢物的异常水平不仅可以作为提示疾病进展的标志，还可能用于疾病治疗[8]。DR的进展发生伴随着体内多种代谢物丰度水平的变化，近年来，许多研究对DR的代谢特征进行了深入分析，并发现氨基酸代谢紊乱与DR的发生和发展密切相关[9-12]，例如精氨酸生物合成[10, 13-14]、精氨酸-脯氨酸代谢[15-16]、谷氨酸-半胱氨酸[16-17]代谢途径都可能参与DR发病机制。此外，嘌呤代谢途径也与DR的进展密切相关[18]。然而，目前大多数研究集中在血液样本，眼内液(如房水或玻璃体液)的研究较少，且尚未有研究报道代谢物在DR进展中的具体作用。因此，需要进一步扩大样本量，纳入更多眼内液样本和不同病变程度的DR患者，以对DR的眼内微环境代谢特征进行深入分析。

代谢组学结合机器学习的方法在眼科疾病研究中展现出广泛应用潜力。机器学习方法增强了代谢组学数据分析的效率和精准度，能够识别和分类与DR进展相关的代谢物，筛选出潜在的生物标志物和疾病相关代谢通路。特别是随机森林算法在青光眼和视网膜病变等领域的研究中表现尤为突出[19-20]。该算法能够处理高维度和复杂的代谢组学数据，通过集成大量决策树，提高了分析的准确性和稳健性[21-22]。将代谢组学与机器学习相结合的多学科方法，有望显著提升对DR进展机制的理解，并推动新型生物标志物的发现和DR的个体化治疗。

本研究计划使用代谢组学结合机器学习的方法对DR患者的深入分析房水代谢谱，寻找与T2DM及DR进展相关的关键代谢标志物，为探究DR的发病机制及个体化治疗提供理论基础。

1　对象与方法

1.1　研究对象

本研究获得中山大学中山眼科中心伦理委员会批准(批件号: 2023KYPJ292)，并获得所有参与者的同

意。T2DM组纳入标准: 年龄>40岁，符合T2DM的诊断标准，3个月内未接受过侵入性眼科治疗(包括玻璃体切割术、视网膜激光光凝、光动力学治疗等)，符合2型糖尿病的诊断标准。T2DM患者依据视网膜病变程度分为3组：NDR组，NPDR组，PDR组。对照(CON)组：老年性白内障患者，无糖尿病病史及其他眼部病变。排除标准：3个月内有玻璃体腔内注射或全身抗血管内皮生长因子注射治疗，活动性眼部炎症，既往眼底手术史，眼外伤史，其他眼底疾病(包括年龄相关性黄斑变性、视网膜静脉阻塞、青光眼、黄斑前膜等)，全身或局部类固醇治疗和免疫抑制药物治疗病史，伴有糖尿病相关肾病、糖尿病酮症等并发症及传染病的患者。

1.2 房水样本及临床资料收集

记录所有研究对象的临床资料信息，包括年龄、性别、糖尿病病程、实验室检查结果(空腹血糖、尿素氮、总胆固醇和甘油三酯)。DR组患者在我院接受玻璃体腔抗血管内皮生长因子药物(康柏西普)治疗，NDR组及CON组患者在我院接受了白内障手术治疗。所有患者在手术开始时采集房水样本50~150 μL，立即转移到500 μL无菌离心管中，放置在干冰中冷冻，随后转移到−80 ℃冰箱中冻存直至进行代谢组学分析。

1.3 代谢组学样本制备与质谱分析

将冻存房水样本于室温下解冻，取40 μL至EP微型离心管中，加入160 μL甲醇-乙腈混合液(内含同位素内标)，涡旋5 min后，在4℃下13 000 χ g离心20 min，取150 μL上清液转移至另一EP管中并减压离心浓缩干燥。上机检测前，用60 μL 50%甲醇-水复溶，充分溶解后再次离心，取上清液至内衬管中待分析。

代谢组学分析在超高效液相色谱串联四极杆-静电场轨道阱高分辨质谱(ultra-high performance liquid chromatography tandem quadrupole-electrostatic field orbitrap high resolution mass spectrometer, UHPLC- HRMS)平台上进行检测。原始代谢组学数据通过化合物发现软件(Thermo Scientific, USA)进行成分提取，并使用iPhenome™ SMOL高分辨MS/MS库和mzCloud谱库进行代谢物结构注释。

1.4 数据处理与统计分析

临床数据的连续变量以均值 ± 标准差表示，分类变量以频数(%)表示。代谢组学数据经过过滤、归一化、对数变换和规范化处理后，使用SIMCA软件(Sartorius AG Umetrics, Germany)进行主成分分析和正交偏最小二乘判别分析。使用独立*t*检验的方法初步分析得到显著性水平*P*值，并采用Benjamini-Hochberg(BH)方法对显著性水平*P*值进行错误发现率(false discovery rate, FDR)校正，以FDR值<0.05及两组间代谢物变化倍数(foldchange, FC)>1.2或<0.8作为标准初步筛选出差异代谢物。数据分析及可视化使用SPSS(版本26)、R语言(版本4.1.3)和IPOS云平台完成，差异代谢物的Kyoto Encyclopedia of Genes and Genomes(KEGG)通路富集分析使用MetaboAnalyst 6.0进行。采用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)惩罚回归对代谢组学数据进行优化以减少数据冗余，通过十折交叉验证确定最优正则化参数λ(选择交叉验证误差最小的λ值)，保留回归系数非零的代谢物，以进一步筛选最为相关的代谢组学特征子集。使用Logistic回归和多模型线性回归校正混杂因素，*P*<0.05作为筛选标准。使用机器学习中随机森林算法作为模型建立的算法，采用十则交叉验证以避免过拟合，MeanDecreaseGini表示每个变量对分类树上每个节点观测值异质性的影响，值越大表示变量的重要性越大，以MeanDecreaseGini为纵坐标评估标志物对模型的重要性，R^2是一种衡量模型拟合程度的常用指标它表示了模型所能解释的观测数据方差的比例。通过计算每种代谢物的MeanDecreaseGini值以及*P*值(*P*<0.05，标记为“*”)量化代谢物对预测模型的贡献。生物标志物筛选、预测模型构建及图表生成应用R(版本4.1.3)完成。

2 结果

2.1 基线特征

通过对入组患者的一般资料的比较发现(表1)，基线时四组在年龄、性别与甘油三酯均无统计学差异(*P*>0.05)。T2DM组血浆中空腹血糖水平和尿素水平均高于CON组(均*P*<0.001)。T2DM组间比较糖尿病病程也存在统计学差异，PDR组的平均糖尿病病程为(12.14 ± 3.18)年，NPDR组平均糖尿病病程为(8.22 ± 3.20)年，NDR组平均糖尿病病程为(7.50 ± 2.92)年(*P*<0.001)。

2.2 T2DM差异代谢物及相关代谢通路分析

在T2DM组与CON组的比较中共检测到显著差异代谢物143个，使用Log2FC作为横坐标，-Log10FDR作

表 1 研究人群基线特征

Table 1 Baseline characteristics of the study participants

	对照组	2 型糖尿病组	糖尿病无糖尿病视网膜病变组	非增殖期糖尿病视网膜病变组	增殖期糖尿病视网膜病变组	P^a	P^b
n	30	78	30	18	30		
性别 , *n*(%)						0.328	0.386
女性	13(43.33)	42(53.85)	19(63.33)	8(44.44)	15(50.00)		
男性	17(56.67)	36(46.15)	11(36.67)	10(55.56)	15(50.00)		
年龄 / 年	65.07 ± 2.50	64.19 ± 10.1	66.01 ± 2.84	63.05 ± 5.59	64.10 ± 3.08	0.352	0.189
糖尿病病程 / 年	—	9.38 ± 3.75	7.5 ± 2.92	8.22 ± 3.20	12.14 ± 3.18	—	<0.001
空腹血糖 /(mmol/L)	5.5 ± 0.71	7.9 ± 2.90	7.97 ± 2.76	7.99 ± 3.75	7.78 ± 2.53	<0.001	0.982
尿素 /(mmol/L)	4.72 ± 1.20	8.33 ± 4.56	6.85 ± 2.40	9.56 ± 4.51	8.58 ± 6.05	<0.001	0.044
总胆固醇 /(mmol/L)	5.3 ± 1.08	5.07 ± 1.15	4.81 ± 1.19	4.68 ± 0.78	5.46 ± 1.15	0.325	0.040
甘油三酯 /(mmol/L)	1.86 ± 1.31	1.69 ± 0.92	1.62 ± 1.12	1.64 ± 0.74	1.8 ± 0.95	0.953	0.610

P^a表示对照组与2型糖尿病组比较，使用的统计方法为独立t检验；P^b表示三组间（糖尿病无糖尿病视网膜病变组, 非增殖期糖尿病视网膜病变组, 增殖期糖尿病视网膜病变组）比较，使用的统计方法为单因子方差分析。对于分类变量，使用卡方检验计算*P*值。

P^a represents the comparison between the control group and the type 2 diabetes mellitus group, using the independent t-test for statistical analysis. P^b represents the comparison among the three groups (type 2 diabetes mellitus without diabetic retinopathy, non-proliferative diabetic retinopathy, and proliferative diabetic retinopathy), using one-way ANOVA for statistical analysis. For categorical variables, *P*-values were calculated using the chi-square test.

为纵坐标绘制差异代谢物火山图，结果见图1 A。使用KEGG数据库鉴定了T2DM的相关代谢通路。结果共富集到13条差异代谢通路(表2)，其中“缬氨酸、亮氨酸和异亮氨酸生物合成”“精氨酸生物合成”“半胱氨酸和蛋氨酸代谢”是T2DM的主要相关通路，以气泡大小及颜色深浅表示富集的代谢通路对T2DM的影响，结果见图1B。

为了进一步筛选出区分能力更强的T2DM差异代谢物，使用LASSO惩罚回归对初步分析得到的代谢数据集进行特征筛选。对比CON组，T2DM组中氨基己二酸、葡萄糖内酯、葡萄糖，*L*-蛋氨酸、N,N,N三甲基-5-氨基戊酸和尿素丰度水平显著上调，而泛醇、硫酸十四烷基酯和1,5-脱水山梨醇的丰度水平显著降低(图2A)。在校正了年龄及性别后，1,5-脱水山梨醇、硫酸十四烷基酯和N,N,N三甲基-5-氨基戊酸被识别为T2DM的生物标志物。1,5-脱水山梨醇的比值比(odds ratio, OR)值为0.893，95%置信区间(confidence interval, CI)为0.851-0.937，硫酸十四烷基酯的OR值为0.809，95%CI为0.712-0.919，N,N,N三甲基-5-氨基戊酸的OR值为1.141，95%CI为1.077-1.210(表3，均$P<0.05$)。95%CI表示，1,5-脱水山梨醇的OR值有95%的可能性落在0.851到0.937之间，硫酸十四烷基酯的OR值有95%的可能性落在0.712 ~ 0.919，而N,N,N三甲基-5-氨基戊酸的OR值有95%的可能性落在1.077到1.210之间。如果置信区间不包含1，表明该代谢物与T2DM之间的关系具有统计学意义，即该代谢物能够区分T2DM和健康对照组。

2.3 DR进展相关差异代谢物

为了探讨与DR进展相关的代谢物，研究分析了T2DM的3组代谢数据(包括NDR、NPDR和PDR)。结果表明，N-乙酰色氨酸、亚油酰胺、油酰胺、棕榈酰胺、戊酸(游离脂肪酸(5:0)和琥珀酸差异显著

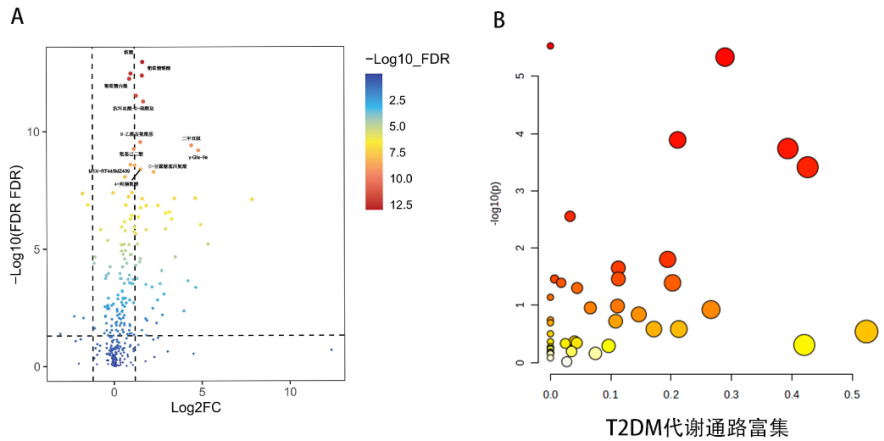


图1 2型糖尿病差异代谢物火山图、代谢通路富集图

Figure 1 Volcano plot and pathway analysis of differential metabolites in type 2 diabetes mellitus

(A) 对照组与2型糖尿病组差异代谢物火山图。(B) 对照组与2型糖尿病组差异代谢物相关代谢通路富集气泡图，红色越深，P值越小。圆圈越大，富集的代谢通路影响越大。

(A) Volcano plot of differential metabolites between the control (CON) and type 2 diabetes mellitus (T2DM) groups. (B) Bubble plot of enriched metabolic pathways related to differential metabolites between the CON and T2DM groups, where darker red indicates more significant P-values. The larger the circle, the greater the impact of the enriched metabolic pathway.

表 2 T2DM 差异代谢物通路富集分析

Table 2 Pathway analysis of differential metabolites in T2DM

代谢通路名称	个数 / 总数	*P* 值	FDR
缬氨酸、亮氨酸和异亮氨酸生物合成	5/8	<0.001*	<0.001*
精氨酸生物合成	6/14	<0.001*	<0.001*
半胱氨酸和蛋氨酸代谢	7/33	<0.001*	0.003*
组氨酸代谢	5/16	<0.001*	0.003*
丙氨酸、天冬氨酸和谷氨酸代谢	6/28	<0.001*	0.006*
缬氨酸、亮氨酸和异亮氨酸降解	6/40	0.003*	0.037*
烟酸和烟酰胺代谢	3/15	0.016*	0.179
赖氨酸降解	4/30	0.022*	0.222
泛酸和辅酶 A 生物合成	3/20	0.035*	0.270
柠檬酸循环	3/20	0.035*	0.270
嘌呤代谢	6/70	0.040*	0.270
精氨酸和脯氨酸代谢	4/36	0.040*	0.270
戊糖磷酸途径	3/23	0.049*	0.306

*表示*P*值或FDR值小于0.05。错误发现率：False discovery rate, FDR。

"*" represents a *P*-value or FDR < 0.05. False discovery rate, FDR.

(FDR<0.05)(表4)。亚油酰胺、油酰胺和棕榈酰胺在NDR、NPDR和PDR组中均表现出逐步升高的趋势，相反，戊酸(游离脂肪酸5:0)和琥珀酸的丰度水平在DR进展过程中逐步降低(图2B)。此外，N-乙酰色氨酸在NPDR组中的丰度水平显著上调，PDR组的N-乙酰色氨酸丰度水平相较于NPDR组显著降低，但相较于NDR组升高。

校正了年龄、性别和糖尿病病程后，我们发现6个差异代谢物仍具有显著差异。如表5所示，N-乙酰色氨酸的回归系数β值为0.632(95%CI: 0.312~ 0.952)，亚油酰胺的回归系数β值为0.230(95%CI: 0.182~0.277)，油酰胺的回归系数β值为0.205(95%CI: 0.157~0.254)，棕榈酰胺的回归系数β值为0.436(95%CI: 0.325~0.547)；戊酸(游离脂肪酸5:0)的回归系数β值为−0.426(95%CI: −0.578~−0.262)，琥珀酸的回归系数β值为−0.613(95%CI: −0.911~−0.315)(均P<0.001)。这些结果表明，N-乙酰色氨酸、亚油酰胺、油酰胺和棕榈酰胺的增加与DR进展呈显著正相关，而戊酸(游离脂肪酸5:0)和琥珀酸的增加则与DR进展呈显著负相关。回归系数β值反映了每个代谢物与DR进展之间的关系强度和方向。N-乙酰色氨酸、亚油酰胺、油酰胺和棕榈酰胺的β值为正，表明代谢物的增加与DR进展呈正相关。而戊酸和琥珀酸的β值为负，说明代谢物的增加与DR进展呈负相关。β值的大小则表示每个代谢物对DR进展的影响程度，β值越大，影响越显著。

2.4　预测模型的建立

基于Logistic和多模型线性回归识别出的生物标志

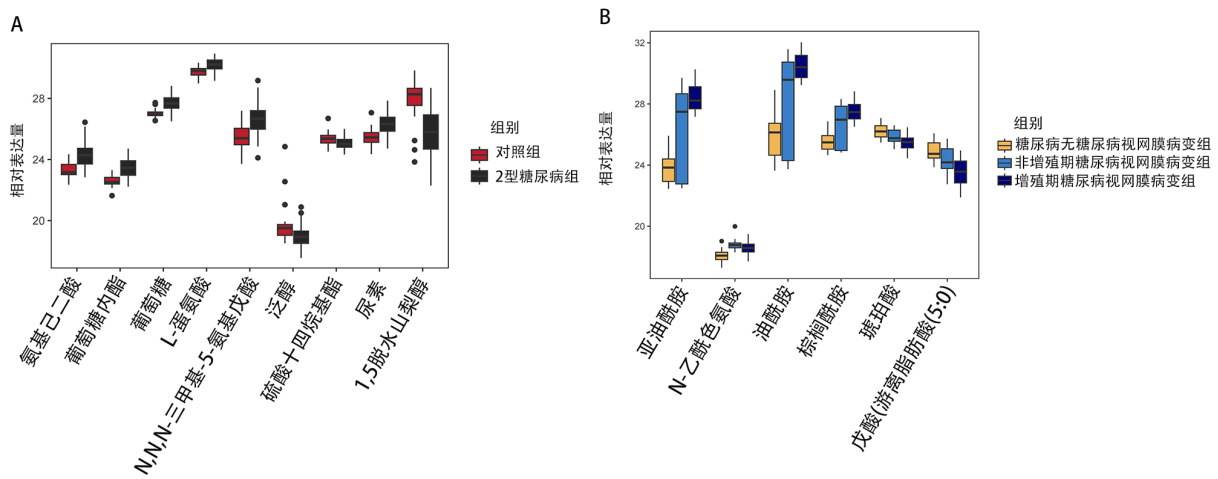


图2　2型糖尿病及糖尿病视网膜病变进展差异代谢物相对表达箱式图

Figure 2　Box plot of relative intensity of differential metabolites in type 2 diabetes mellitus and diabetic retinopathy progression

(A) 2型糖尿病差异代谢物箱式图；(B) 糖尿病视网膜病变进展差异代谢物箱式图。

(A) Box plot of differential metabolites in type 2 diabetes mellitus. (B) Box plot of differential metabolites in diabetic retinopathy progression.

表 3　T2DM 差异代谢物 logistic 回归分析

Table 3　Logistic regression analysis of differential metabolites in T2DM

代谢物	OR 值	95%CI	*P* 值
1,5- 脱水山梨醇	0.893	0.851~0.937	<0.001
硫酸十四烷基酯	0.809	0.712~0.919	0.001
N,N,N 三甲基 -5- 氨基戊酸	1.141	1.077~1.210	<0.001

置信区间：Confidence interval, CI；比值比：Odds ratio, OR。

Confidence interval, CI；Odds ratio, OR.

表 4 DR 进展相关差异代谢物

Table 4 Differential metabolites associated with DR progression

代谢物	糖尿病无糖尿病视网膜病变组(NDR, 均值 ± 标准差，log2)	非增殖期糖尿病视网膜病变组(NPDR, 均值 ± 标准差，log2)	增殖期糖尿病视网膜病变组(PDR, 均值 ± 标准差，log2)	ANOVA *P* 值	FDR	NDR *vs.* NPDR (*P* 值 *)	NPDR *vs.* PDR (*P* 值 *)
N- 乙酰色氨酸	18.05 ± 0.41	18.79 ± 0.37	18.54 ± 0.42	8.66E-08	5.44E-06	6.49E-08	4.66E-02
亚油酰胺	23.81 ± 0.93	26.18 ± 2.89	28.46 ± 0.92	4.62E-16	1.45E-13	7.00E-06	1.60E-05
油酰胺	25.94 ± 1.42	28.01 ± 3.11	30.50 ± 0.83	2.00E-13	3.14E-11	3.69E-04	2.70E-05
棕榈酰胺	25.54 ± 0.60	26.54 ± 1.35	27.56 ± 0.63	1.48E-12	1.55E-10	2.79E-04	1.51E-05
戊酸 (游离脂肪酸 (5:0)	24.93 ± 0.59	24.33 ± 0.85	23.55 ± 0.84	1.08E-08	8.45E-07	1.11E-02	1.15E-03
琥珀酸	26.20 ± 0.41	25.87 ± 0.46	25.48 ± 0.51	6.34E-07	3.32E-05	1.88E-02	7.83E-03

*表示使用最小显著差异（Least Significant Difference, LSD)法进行ANOVA分析后的事后检验，两两比较的*P*值。

"*" represents the *P*-values from post hoc tests after ANOVA analysis using the Least Significant Difference (LSD) method for pairwise comparisons.

表 5 DR 差异代谢物的多模型线性回归分析

Table 5 Multi-model linear regression analysis of differential metabolites in DR

代谢物	模型 1		模型 2		模型 3	
	β 系数 (95%CI)	*P* 值	*β* 系数 (95%CI)	*P* 值	*β* 系数 (95%CI)	*P* 值
N- 乙酰色氨酸	0.742(0.387~1.097)	<0.001	0.732(0.357~1.108)	<0.001	0.632(0.312~0.952)	<0.001
亚油酰胺	0.264(0.217~0.311)	<0.001	0.263(0.218~0.308)	<0.001	0.230(0.182~0.277)	<0.001
油酰胺	0.239(0.190~0.289)	<0.001	0.241(0.193~0.288)	<0.001	0.205(0.157~0.254)	<0.001
棕榈酰胺	0.517(0.405~0.629)	<0.001	0.517(0.409~0.625)	<0.001	0.436(0.325~0.547)	<0.001
戊酸 (游离脂肪酸 5:0)	–0.566(–0,726~–0.405)	<0.001	–0.551(–0.710~–0.392)	<0.001	–0.426(–0.578~–0.262)	<0.001
琥珀酸	–0.884(–1.176~–0.592)	<0.001	–0.849(–1.147~–0.551)	<0.001	–0.613(–0.911~–0.315)	<0.001

模型1：原始模型；模型2：校正年龄和性别；模型3：校正年龄、性别和2型糖尿病病程(年)。置信区间：confidence interval，CI。

Model 1: Crude; Model 2: Adjusted for age and sex; Model 3: Adjusted for age, sex, and duration of type 2 diabetes mellitus (years). Confidence interval, CI.

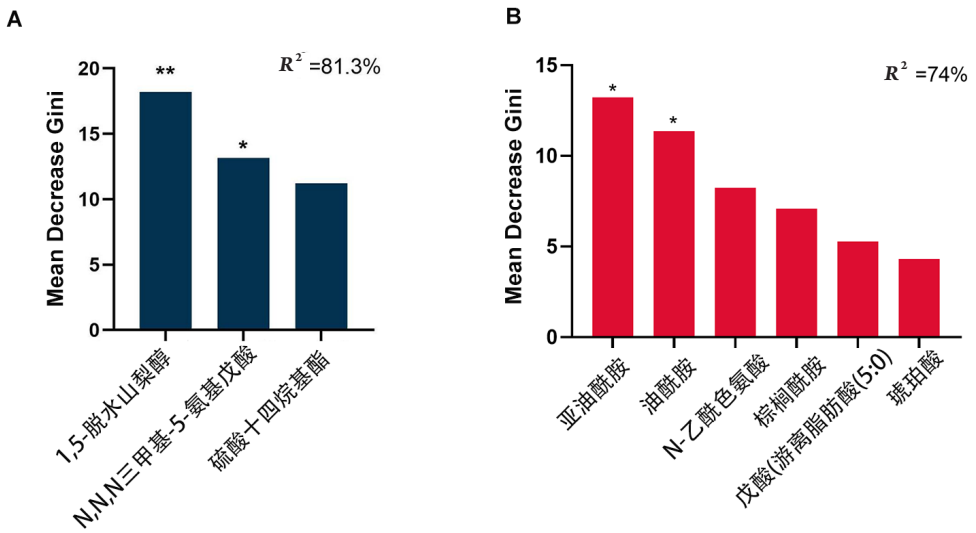


图3　2型糖尿病及糖尿病视网膜病变进展预测模型图

Figure 3　Predictive model for type 2 diabetes mellitus and diabetic retinopathy progression

(A) 2型糖尿病差异联合生物标志物预测模型；(B) 糖尿病视网膜病变进展联合生物标志物预测模型。

(A) Predictive model for type 2 diabetes mellitus; (B) Predictive model for diabetic retinopathy progression.

物将纳入随机森林模型，构建疾病诊断联合生物标志物预测模型。结果显示，T2DM预测模型对疾病的解释度为81.3%，其中1,5-脱水山梨醇对预测模型的贡献最大(图3A)。DR进展的预测模型对疾病的解释度为74%，其中亚油酰胺对预测模型的贡献最大(图3B)。

3　讨论

随着糖尿病患病率的上升，DR的发病率也在不断增加。识别与DR进展相关的生物标志物变得尤为重要。尽管代谢组学在代谢性疾病研究中展现了巨大潜力，但针对DR进展的研究，尤其是眼内液代谢物的分析，仍相对较少。本研究通过UHPLC-HRMS技术对房水中的代谢物进行了全面检测，揭示了DR不同进展阶段(NDR、NPDR、PDR)之间的代谢特征差异。结果表明，T2DM组与CON组及各DR进展阶段患者的房水代谢谱存在显著差异。结合机器学习方法优化高通量代谢数据集，我们筛选出了一系列与T2DM及DR进展相关的生物标志物，并成功构建了预测模型。

人体中1, 5-脱水山梨醇主要来源于食物，经过肾小管重吸收后通过尿液排出。本研究发现，T2DM患者中1, 5-脱水山梨醇的丰度水平有所下降。研究已证实，尿糖可以竞争性抑制1, 5-脱水山梨醇在肾小管中的重吸收，导致血液中的1, 5-脱水山梨醇水平降低[23-24]。在糖尿病患者中，1,5-脱水山梨醇比HbA1C更敏感，能更早检测出血糖的轻微变化。同时，研究也发现高水平的1,5-脱水山梨醇与较低的DR患病率相关[25]，可能通过参与胰岛β细胞功能的损伤过程，从而影响血糖调节[26]。此外本研究显示，硫酸十四烷基酯在T2DM患者中也表达下降，但还未曾有研究报道过其在T2DM中的作用。N, N, N三甲基-5-氨基戊酸是一种肠道微生物代谢产物，被发现对心脏代谢重编程有重要作用，可能导致心力衰竭[27]。我们的研究结果显示，T2DM患者房水中N, N, N三甲基-5-氨基戊酸的丰度水平升高，提示其可能与T2DM相关。

本研究发现亚油酰胺在NDR、NPDR及PDR组中表达水平逐渐上升，表明其可能在DR进展中发挥重要作用。亚油酰胺是一种脂肪酸酰胺类化合物，是由脂肪酸与胺反应生成的羧酸酰胺衍生物，其作为脂肪酰胺类脂质分子，在脂质代谢中发挥关键作用。在DR的发生与发展过程中，视网膜血管内皮细胞内钙离子稳态的紊乱是一个重要特征。研究表明，亚油酰胺通过调节肌浆网/内质网钙ATP酶，从而调控细胞内钙离子浓度，影响线粒体功能和能量代谢，进一步影响视网膜血管内皮细胞的代谢稳态[28]。研究还发现，亚油酰胺能够通过促进内质网钙释放和钙离子内流，显著提高胞质钙浓度，从而调节能量代谢，这一过程独立于传统的肌醇三磷酸信号通路。亚油酰胺对线粒体ATP

生成的调控作用，可能在视网膜血管内皮细胞的能量代谢中发挥关键作用[29]。进一步的研究表明，亚油酰胺通过抑制Ca^{2+}-ATP酶活性，可能在调节视网膜细胞内钙稳态方面起到重要作用，参与DR的病理过程[30]。Janus激酶2/信号转导子及转录激活子3(Janus kinase 2 / signal transducer and activator of transcription 3, Jak2/STAT3)通路可以将各种细胞因子和生长因子的信号传递到细胞核，调节基因表达，引起视网膜神经节细胞的氧化应激损伤[31]。有研究表明，亚油酰胺可以通过调节Jak2/STAT3通路，抑制核因子 κ B p65的活性，抑制促炎细胞因子的产生，可能在DR进展中发挥重要的抗炎和神经保护作用。32这些发现表明，亚油酰胺可能通过多重机制参与DR的代谢调节，进一步研究亚油酰胺在DR中的具体作用机制将有助于揭示其在DR进展中的潜在作用。

油酰胺是一种由油酸转化而成的脂肪酸酰胺，是天然存在于动物体内的内源性物质。由于其结构与内源性大麻素花生四烯乙醇胺相似，油酰胺能够作为完全激动剂结合大麻素1型受体。通过与大麻素1型受体结合，油酰胺调节胰岛素分泌和葡萄糖代谢，从而可能影响DM中的代谢紊乱[33-34]。Kim等[35]的研究发现，长期饮食干预能够改变DM患者外周血单核细胞中的油酰胺水平，油酰胺水平的变化与脂蛋白相关磷脂酶A2(lipoprotein-associated phospholipase A2, Lp-PLA2)活性的变化呈正相关，这提示油酰胺可能通过调节Lp-PLA2的活性，在调节糖尿病炎症反应中发挥重要作用。Mahali等[36]发现，油酰胺能够激活过氧化物酶体增殖因子激活受体 γ (peroxisome proliferator-activated receptor gamma, PPAR γ)，从而提高其在不同细胞中的活性，而晚期糖基化终产物(advanced glycation end products, AGEs)能够抑制油酰胺对PPAR γ 的激活作用。AGEs通过促使PPAR γ 从细胞核转移到细胞质，并激活细胞外信号调节激酶(extracellular signal-regulated kinase, ERK)信号通路，从而抑制PPAR γ 的活性。这一机制揭示了AGEs可能通过调节油酰胺与PPAR γ 的相互作用，进一步影响视网膜细胞的代谢稳态，进而促进DR的发生发展。此外，有研究发现油酰胺作为糖尿病引起的动脉粥样硬化性病变的重要生物标志物，可能与血管内皮细胞的钙化过程密切相关[37]。本研究结果提示，油酰胺在DR进展中表达水平升高，可能在DR的发生和发展中发挥重要作用，未来还需要进一步探究油酰胺在DR进展中的具体作用。

本研究中，与DR进展相关的生物标志物，如亚油酰胺和油酰胺，均属于脂肪酸酰胺类化合物，这表明脂肪酸酰胺在DR进展中发挥着关键作用。脂肪酸酰胺通过调节其降解酶——脂肪酸酰胺酶(fatty acid amide hydrolase, FAAH)，影响高糖条件下的细胞反应。在高糖环境中，FAAH 1的表达下调会导致脂肪酸酰胺的积累，这些脂肪酸酰胺通过与大麻素1型受体结合，引发氧化应激的增加和细胞凋亡。相反，FAAH 1的过表达能够缓解这些不良效应，进一步证明了脂肪酸酰胺在DR病理过程中的重要调节作用。38脂肪酸乙醇酰胺可以通过激活G蛋白偶联受体119(G protein-coupled receptor 119, GPR119)来增加细胞内环磷酸腺苷水平，从而促进胰岛素分泌和改善葡萄糖代谢，然而此激活作用可能受时间及配体间的竞争影响[39]。这些证据表明，脂肪酸酰胺类化合物通过调节G蛋白偶联受体及其他相关受体，改变细胞的代谢状态，是导致DR代谢紊乱的关键因素。

综上所述，本研究通过代谢组学结合机器学习分析，揭示了T2DM及DR患者的代谢变化，特别是脂肪酸酰胺类化合物在识别DR进展中的重要作用。结果显示，亚油酰胺和油酰胺在DR各阶段的表达存在显著差异，提示它们具有作为DR进展标志物的潜力。然而，本研究的样本量相对较少，差异代谢物仍需在大样本人群中进一步验证，其具体作用机制仍需进一步研究。

声明

在论文撰写中无使用生成式人工智能。论文撰写中的所有内容均由作者独立完成，并对出版物的真实性和准确性承担全部责任。

利益冲突 所有作者均声明不存在利益冲突。

开放获取声明

本文适用于知识共享许可协议(Creative Commons)，允许第三方用户按照署名(BY)-非商业性使用(NC)-禁止演绎(ND)(CC BY-NC-ND)的方式共享，即允许第三方对本刊发表的文章进行复制、发行、展览、表演、放映、广播或通过信息网络向公众传播，但在这些过程中必须保留作者署名、仅限于非商业性目的、不得进行演绎创作。

参考文献

1. Risk Factor Collaboration (NCD-RisC) NCD. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. Lancet, 2024, 404(10467): 2077-2093. DOI: 10.1016/S0140-6736(24)02317-1.
2. Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. Lancet, 2010, 376(9735): 124-136. DOI: 10.1016/S0140-6736(09)62124-3.
3. Hou X, Wang L, Zhu D, et al. Prevalence of diabetic retinopathy and vision-threatening diabetic retinopathy in adults with diabetes in China. Nat Commun, 2023, 14(1): 4296. DOI: 10.1038/s41467-023-39864-w.
4. Chai YH, Zhang YP, Qiao YS, et al. Association between diabetic retinopathy, brain structural abnormalities, and cognitive impairment for accumulated evidence in observational studies. Am J Ophthalmol, 2022, 239: 37-53. DOI: 10.1016/j.ajo.2022.01.011.
5. Kim HU, Park SP, Kim YK. Long-term HbA1c variability and the development and progression of diabetic retinopathy in subjects with type 2 diabetes. Sci Rep, 2021, 11(1): 4731. DOI: 10.1038/s41598-021-84150-8.
6. González-Peña D, Brennan L. Recent advances in the application of metabolomics for nutrition and health. Annu Rev Food Sci Technol, 2019, 10: 479-519. DOI: 10.1146/annurev-food-032818-121715.
7. Zhang J, Zhang J, Zhang C, et al. Diabetic macular edema: current understanding, molecular mechanisms and therapeutic implications. Cells, 2022, 11(21): 3362. DOI: 10.3390/cells11213362.
8. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol, 2016, 17(7): 451-459. DOI: 10.1038/nrm.2016.25.
9. Peters KS, Rivera E, Warden C, et al. Plasma arginine and citrulline are elevated in diabetic retinopathy. Am J Ophthalmol, 2022, 235: 154-162. DOI: 10.1016/j.ajo.2021.09.021.
10. Sumarriva K, Uppal K, Ma C, et al. Arginine and carnitine metabolites are altered in diabetic retinopathy. Invest Ophthalmol Vis Sci, 2019, 60(8): 3119-3126. DOI: 10.1167/iovs.19-27321.
11. Rhee SY, Jung ES, Park HM, et al. Plasma glutamine and glutamic acid are potential biomarkers for predicting diabetic retinopathy. Metabolomics, 2018, 14(7): 89. DOI: 10.1007/s11306-018-1383-3.
12. Sun Y, Zou H, Li X, et al. Plasma metabolomics reveals metabolic profiling for diabetic retinopathy and disease progression. Front Endocrinol (Lausanne), 2021, 12: 757088. DOI: 10.3389/fendo.2021.757088.
13. Wang Z, Tang J, Jin E, et al. Serum untargeted metabolomics reveal potential biomarkers of progression of diabetic retinopathy in asians. Front Mol Biosci, 2022, 9: 871291. DOI: 10.3389/fmolb.2022.871291.
14. Wang X, Yang S, Yang G, et al. Novel risk score model for non-proliferative diabetic retinopathy based on untargeted metabolomics of venous blood. Front Endocrinol (Lausanne), 2023, 14: 1180415. DOI: 10.3389/fendo.2023.1180415.
15. Zhu XR, Yang FY, Lu J, et al. Plasma metabolomic profiling of proliferative diabetic retinopathy. Nutr Metab (Lond), 2019, 16: 37. DOI: 10.1186/s12986-019-0358-3.
16. Guo C, Jiang D, Xu Y, et al. High-coverage serum metabolomics reveals metabolic pathway dysregulation in diabetic retinopathy: a propensity score-matched study. Front Mol Biosci, 2022, 9: 822647. DOI: 10.3389/fmolb.2022.822647.
17. Zuo J, Lan Y, Hu H, et al. Metabolomics-based multidimensional network biomarkers for diabetic retinopathy identification in patients with type 2 diabetes mellitus. BMJ Open Diabetes Res Care, 2021, 9(1): e001443. DOI: 10.1136/bmjdrc-2020-001443.
18. Han X, Zhang L, Kong L, et al. Comprehensive metabolic profiling of diabetic retinopathy. Exp Eye Res, 2023, 233: 109538. DOI: 10.1016/j.exer.2023.109538.
19. Liu H, Li J, Wong L. A comparative study on feature selection and classification methods using gene expression profiles and proteomic patterns. Genome Inform, 2002, 13: 51-60.
20. Strobl C, Boulesteix AL, Kneib T, et al. Conditional variable importance for random forests. BMC Bioinformatics, 2008, 9: 307. DOI: 10.1186/1471-2105-9-307.
21. Sammut SJ, Crispin-Ortuzar M, Chin SF, et al. Multi-omic machine learning predictor of breast cancer therapy response. Nature, 2022, 601(7894): 623-629. DOI: 10.1038/s41586-021-04278-5.
22. Lewis JE, Kemp ML. Integration of machine learning and genome-scale metabolic modeling identifies multi-omics biomarkers for radiation resistance. Nat Commun, 2021, 12(1): 2700. DOI: 10.1038/s41467-021-22989-1.
23. Yamanouchi T, Ogata N, Tagaya T, et al. Clinical usefulness of serum 1, 5-anhydroglucitol in monitoring glycaemic control. Lancet, 1996, 347(9014): 1514-1518. DOI: 10.1016/s0140-6736(96)90672-8.
24. Dungan KM. 1, 5-anhydroglucitol (GlycoMark) as a marker of short-term glycemic control and glycemic excursions. Expert Rev Mol Diagn, 2008, 8(1): 9-19. DOI: 10.1586/14737159.8.1.9.
25. Selvin E, Rawlings AM, Grams M, et al. Association of 1, 5-anhydroglucitol with diabetes and microvascular conditions. Clin Chem, 2014, 60(11): 1409-1418. DOI: 10.1373/

clinchem.2014.229427.

26. Shen Y, Si Y, Lu J, et al. Association between 1, 5-anhydroglucitol and acute C peptide response to arginine among patients with type 2 diabetes. J Diabetes Res, 2020, 2020: 4243053. DOI: 10.1155/2020/4243053.
27. Zhao M, Wei H, Li C, et al. Gut microbiota production of trimethyl-5-aminovaleric acid reduces fatty acid oxidation and accelerates cardiac hypertrophy. Nat Commun, 2022, 13(1): 1757. DOI: 10.1038/s41467-022-29060-7.
28. Yamamoto S, Takehara M, Ushimaru M. Inhibitory action of linoleamide and oleamide toward sarco/endoplasmic reticulum Ca2+-ATPase. Biochim Biophys Acta Gen Subj, 2017, 1861(1 Pt A): 3399-3405. DOI: 10.1016/j.bbagen.2016.09.001.
29. Huang JK, Jan CR. Linoleamide, a brain lipid that induces sleep, increases cytosolic Ca2+ levels in MDCK renal tubular cells. Life Sci, 2001, 68(9): 997-1004. DOI: 10.1016/s0024-3205(00)01002-x.
30. Yamamoto S, Takehara M, Kabashima Y, et al. Identification of novel inhibitors of human SPCA2[J]. Biochem Biophys Res Commun, 2016, 477(2): 266-270. DOI: 10.1016/j.bbrc.2016.06.055.
31. Yu H, Zhou D, Wang W, et al. Protective effect of baicalin on oxidative stress injury in retinal ganglion cells through the JAK/STAT signaling pathway in vitro and in vivo. Front Pharmacol, 2024, 15: 1443472. DOI: 10.3389/fphar.2024.1443472.
32. Ibrahim WW, Sayed RH, Abdelhameed MF, et al. Neuroprotective potential of Erigeron bonariensis ethanolic extract against ovariectomized/D-galactose-induced memory impairments in female rats in relation to its metabolite fingerprint as revealed using UPLC/MS. Inflammopharmacology, 2024, 32(2): 1091-1112. DOI: 10.1007/s10787-023-01418-3.
33. Leggett JD, Aspley S, Beckett SG, et al. Oleamide is a selective endogenous agonist of rat and human CB1 cannabinoid receptors. Br J Pharmacol, 2004, 141(2): 253-262. DOI: 10.1038/sj.bjp.0705607.
34. Pawar HD, Patil Y, Patil A, et al. Cardioprotective effect of CB1 receptor antagonist AM251 against β receptor-stimulated myocardial infarction via modulation of NF-kB signaling pathway in diabetic mice. Heliyon, 2024, 10(15): e35138. DOI: 10.1016/j.heliyon.2024.e35138.
35. Kim M, Song G, Kang M, et al. Replacing carbohydrate with protein and fat in prediabetes or type-2 diabetes: greater effect on metabolites in PBMC than plasma. Nutr Metab (Lond), 2016, 13: 3. DOI: 10.1186/s12986-016-0063-4.
36. Mahali SK, Manna SK. Beta-D-glucoside protects against advanced glycation end products (AGEs)-mediated diabetic responses by suppressing ERK and inducing PPAR gamma DNA binding. Biochem Pharmacol, 2012, 84(12): 1681-1690. DOI: 10.1016/j.bcp.2012.09.033.
37. Ding X, Qiu Y, Wu G, et al. L-thyroxine attenuates extracellular Hsp90 α -induced vascular endothelial calcification in diabetes mellitus, as revealed by parallel metabolic profiles. Atherosclerosis, 2024, 392: 117527. DOI: 10.1016/j.atherosclerosis.2024.117527.
38. Lim SK, Park MJ, Lim JC, et al. Hyperglycemia induces apoptosis via CB1 activation through the decrease of FAAH 1 in retinal pigment epithelial cell. J Cell Physiol, 2012, 227(2): 569-577. DOI: 10.1002/jcp.22756.
39. Kumar P, Kumar A, Song ZH. Structure-activity relationships of fatty acid amide ligands in activating and desensitizing G protein-coupled receptor 119. Eur J Pharmacol, 2014, 723: 465-472. DOI: 10.1016/j.ejphar.2013.10.044.

(**责任编委**：罗明杰；**编辑**：殷秋霞)