

基于HPLC与GC-MS对五种望都辣椒代谢组学分析

王晓宇，王珍珍，胡梦雅，戴静，沙如意，毛建卫

Metabolomics Analysis of Five Types of Wangdu Chili Peppers Based on HPLC and GC-MS

WANG Xiaoyu, WANG Zhenzhen, HU Mengya, DAI Jing, SHA Ruyi, and MAO Jianwei

在线阅读 View online: <https://doi.org/10.13386/j.issn1002-0306.2024010245>

您可能感兴趣的其他文章

Articles you may be interested in

基于机器学习的代谢组学解析腊八蒜储藏过程中代谢物差异

Metabolomics Analysis of Metabolite Differences during the Storage Process of Laba Garlic Based on Machine Learning

食品工业科技. 2023, 44(8): 26–34 <https://doi.org/10.13386/j.issn1002-0306.2022060332>

基于SPME-GC-MS技术分析不同地理标志辣椒粉香气成分差异

Analysis of Aroma Compounds of Different Geographical Indication Chili Powder by SPME-GC-MS

食品工业科技. 2023, 44(21): 275–284 <https://doi.org/10.13386/j.issn1002-0306.2022110056>

19个苹果品种果实香气成分的GC-MS分析

GC-MS Analysis of Fruit Aroma Components in 19 Apple Varieties

食品工业科技. 2019, 40(14): 227–232 <https://doi.org/10.13386/j.issn1002-0306.2019.14.037>

火锅常用干辣椒品质及挥发性成分研究

Quality and Volatile Components of Dried Peppers Commonly Used in Hot Pot

食品工业科技. 2023, 44(16): 358–366 <https://doi.org/10.13386/j.issn1002-0306.2022100262>

基于HS-SPME-GC-MS分析不同茶树品种晒红茶的香气成分

Analysis of the Aroma Composition of Sun-dried Black Tea Samples Processed by Different Varieties Based on HS-SPME-GC-MS

食品工业科技. 2023, 44(7): 260–268 <https://doi.org/10.13386/j.issn1002-0306.2022040295>

基于低共熔溶剂法超声辅助蒸馏提取胡椒叶精油的工艺优化及其GC-MS分析

Optimization and GC-MS Analysis of Extracting Pepper Leaves Essential Oil by Ultrasonic Assisted Distillation Based on Deep Eutectic Solvents

食品工业科技. 2020, 41(20): 135–141 <https://doi.org/10.13386/j.issn1002-0306.2020.20.022>



关注微信公众号，获得更多资讯信息

王晓宇, 王珍珍, 胡梦雅, 等. 基于 HPLC 与 GC-MS 对五种望都辣椒代谢组学分析 [J]. 食品工业科技, 2024, 45(20): 14–22. doi: 10.13386/j.issn1002-0306.2024010245

WANG Xiaoyu, WANG Zhenzhen, HU Mengya, et al. Metabolomics Analysis of Five Types of Wangdu Chili Peppers Based on HPLC and GC-MS[J]. Science and Technology of Food Industry, 2024, 45(20): 14–22. (in Chinese with English abstract). doi: 10.13386/j.issn1002-0306.2024010245

·特邀主编专栏—组学分析技术在食品品质研究中的应用（客座主编：徐贞贞、张良晓、洪惠）·

基于 HPLC 与 GC-MS 对五种望都辣椒代谢组学分析

王晓宇¹, 王珍珍¹, 胡梦雅², 戴 静¹, 沙如意^{1,*}, 毛建卫¹

(1.浙江科技大学生物与化学工程学院, 浙江省农产品化学与生物加工技术重点实验室,
浙江省农业生物资源生化制造协同创新中心, 浙江杭州 310023;
2.中国生物发酵产业协会, 北京 100005)

摘要:为了分析望都辣椒的代谢差异, 本研究采用高效液相色谱 (High-Performance Liquid Chromatography, HPLC) 和气相色谱-质谱技术 (Gas Chromatography-Mass Spectrometry, GC-MS), 对艳椒 (YJ)、辣研 (LY)、国塔 (GT)、研椒 110 (YJA)、热辣 (RL) 等五种不同品种的望都辣椒进行化学成分检测与非靶向代谢组学分析, 并利用机器学习方法对筛选的差异代谢产物进行分类识别。首先利用 HPLC 对五个品种辣椒中辣椒碱、二氢辣椒碱和 V_C 的含量进行测量, 然后利用 GC-MS 对五种辣椒进行非靶向代谢物分析, 对筛选出的代谢产物采用主成分分析 (PCA) 与正交偏最小二乘法判别分析 (OPLS-DA), 得到差异代谢产物及差异代谢通路, 并根据差异代谢产物来进行机器学习识别。在五种辣椒中, 热辣中辣椒碱与二氢辣椒碱含量最高, 分别为 $533.897 \pm 62.187 \mu\text{g/g}$ 和 $264.526 \pm 28.532 \mu\text{g/g}$, V_C 在艳椒中含量最高, 为 $146.9 \pm 0.029 \text{ mg}/100 \text{ g}$ 。由 OPLS-DA 筛选出 16 种差异代谢物, 其中, 奎宁酸和乌头酸等有机酸在艳椒中含量较高; D-山梨醇在辣研中含量最高; 国塔当中柠檬酸、D-果糖、D-甘露糖和乳酸富集程度最高; 研椒 110 中 D-塔格糖与氨基酸含量最高; 而热辣中葡萄糖和肌醇含量在所有品种中占优势。基于京都基因与基因组百科全书 (Kyoto Encyclopedia of Genes and Genomes, KEGG) 通路富集分析结果表明, 差异代谢通路主要为半乳糖代谢、果糖和甘露糖代谢、柠檬酸循环、淀粉和蔗糖代谢、乙醛酸和二羧酸代谢、丙酮酸代谢等。最后, 利用三种机器学习方法随机森林 (Random Forest, RF)、XGBoost 和 BP 神经网络对五种辣椒筛选的差异代谢物进行分类验证, 建立的分类识别模型正确率分别为 100%、92.9% 和 78.6%, 可以用于识别辣椒品种。该研究结果可为望都辣椒的品质评价、品种改良及综合利用提供基础数据。

关键词: 辣椒, 品种, GC-MS, 代谢组学, 机器学习

中图分类号: TS201.1

文献标识码: A

文章编号: 1002-0306(2024)20-0014-09

DOI: 10.13386/j.issn1002-0306.2024010245

本文网刊:



Metabolomics Analysis of Five Types of Wangdu Chili Peppers Based on HPLC and GC-MS

WANG Xiaoyu¹, WANG Zhenzhen¹, HU Mengya², DAI Jing¹, SHA Ruyi^{1,*}, MAO Jianwei¹

(1.Zhejiang Provincial Key Lab for Chem & Bio Processing Technology of Farm Product, Zhejiang Provincial Collaborative Innovation Center of Agricultural Biological Resources Biochemical Manufacturing, School of Biological and Chemical Engineering, Zhejiang University of Science & Technology, Hangzhou 310023, China;
2.China Biofermentation Industry Association, Beijing 100005, China)

Abstract: To analyze the metabolic differences in Wangdu peppers, this study employed high-performance liquid

收稿日期: 2024-01-23

基金项目: 浙江省重点研发计划项目 (2017C02009)。

作者简介: 王晓宇 (1996-), 男, 硕士研究生, 研究方向: 生物质资源利用技术与工程, E-mail: 382780684@qq.com。

* 通信作者: 沙如意 (1982-), 男, 博士, 副教授, 研究方向: 食品生物技术, E-mail: kevinsha_0204@163.com。

chromatography (HPLC) and gas chromatography-mass spectrometry (GC-MS) to detect chemical components and perform non-targeted metabolomics analysis on five different varieties of Wangdu peppers: Yan Jiao (YJ), La Yan (LY), Guo Ta (GT), Yan Jiao 110 (YJA), and Re La (RL). Machine learning was used to classify and identify the differential metabolites screened. First, HPLC was used to measure the content of capsaicin, dihydrocapsaicin, and Vitamin C (V_C) in the five pepper varieties. Then, GC-MS was used for non-targeted metabolite analysis of the five peppers. Principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) were employed to identify differential metabolites and metabolic pathways. Machine learning methods were used to identify the different pepper varieties based on the differential metabolites. In the five pepper varieties, Re La had the highest content of capsaicin and dihydrocapsaicin, at $533.897 \pm 62.187 \mu\text{g/g}$ and $264.526 \pm 28.532 \mu\text{g/g}$, respectively. Yan Jiao had the highest V_C content at $146.9 \pm 0.029 \text{ mg}/100 \text{ g}$. OPLS-DA identified 16 differential metabolites, including organic acids such as quinic acid and aconitic acid, which were higher in Yan Jiao, D-sorbitol, which was highest in La Yan, citric acid, D-fructose, D-mannose, and lactic acid, which were most enriched in Guo Ta, D-tagatose and amino acids, which were highest in Yan Jiao 110, and glucose and inositol, which were most abundant in Re La. KEGG pathway enrichment analysis indicated that the differential metabolic pathways mainly included galactose metabolism, fructose and mannose metabolism, the citric acid cycle, starch and sucrose metabolism, glyoxylate and dicarboxylate metabolism, and pyruvate metabolism. Finally, three machine learning methods—random forest (RF), XGBoost, and backpropagation (BP) neural networks were used to classify and validate the differential metabolites of the five pepper varieties. The established classification models achieved accuracies of 100%, 92.9%, and 78.6%, respectively, demonstrating their utility in identifying pepper varieties. These results would provide fundamental data for the quality evaluation, variety improvement, and comprehensive utilization of Wangdu peppers.

Key words: chili pepper; varieties; GC-MS; metabolomics; machine learning

辣椒(*Capsicum annuum* L.)隶属茄科辣椒属, 是一种一年生或多年生草本植物, 原产于中南美洲, 约于 16 世纪后期传入中国, 是一种重要的蔬菜作物^[1]。我国辣椒产量居世界第一^[2], 贵州、四川、湖南、云南、河南和河北等省份是重要的辣椒生产区^[3], 其中望都辣椒是我国国家地理标志产品, 拥有几十种辣椒制品。辣椒中富含胡萝卜素、柠檬酸、蛋白质、 V_C 等营养成分^[4] 以及钙、磷、铁等微量元素^[5], 其中辣椒的 V_C 含量在蔬菜中占首位, 但由于 V_C 是一种水溶性代谢物, 在高温下易分解; 辣椒碱是辣椒中的一种生物活性成分, 赋予辣椒特有的辛辣味, 是影响辣椒品质的另一项重要指标。辣椒的果实和根茎有抗氧化、抗肿瘤、减肥等药理作用^[6], 具有较高的药用价值。然而, 辣椒的营养组分、功效作用与其品种密切相关, 此外辣椒的深加工产品制备工艺也受产品品种的影响。因此, 建立辣椒品种识别模型对于辣椒品种选育和深加工具有重要意义。

基于液相色谱-质谱(LC-MS)、气相色谱-质谱(GC-MS)等的非靶向代谢组学技术(Untargeted Metabolomics)可用于筛选样品中已知和未知代谢物的定量信息, 近年来, 也被广泛应用于食品、药材等的代谢物组分变化识别, 从而实现产品品种和产地来源的鉴别^[7]。辣椒代谢物主要是挥发性化合物, 例如挥发性有机酸、酯类、醛类等, 使用 GC-MS 更为合适, 并且 GC-MS 的运行成本通常比 LC-MS 低, 分析速度快, 这对于大量样品的高通量分析是非常重要的。杨创创等^[8] 对 13 份干辣椒的挥发性代谢物进行了检测, 得到主要代谢物是柠檬烯、香橙烯、乙酸乙酯。Elvia 等^[9] 基于核磁共振(NMR)的代谢指纹图谱确定对两个不同地区种植的辣椒的代谢物水平,

发现两者的差异代谢物为氨基酸、有机酸和糖。以上研究分析了不同品种辣椒的代谢差异, 但并未对辣椒的品种进行识别验证, 因此, 本研究对辣椒的品种进行了识别验证。目前, 机器学习与代谢组学的联用, 为差异代谢物识别和预测模型建立提供了新的思路。

河北省望都辣椒作为我国国家地理标志产品, 因色泽纯正、肉质肥厚、辣香浓郁而负盛名, 但长期以来没有专业的育种技术, 使传统望都辣椒在纯度和品质上受到一定影响, 因此比较不同品种辣椒品质与差异代谢物可以为辣椒品种改良、品种选育奠定基础。本研究以望都县的艳椒、辣研、热辣、研椒 110 和国塔五种外形相近的辣椒为研究对象, 利用 HPLC 测定样品中 V_C 和辣椒碱类物质的含量, 采用 GC-MS 非靶向代谢组学方法, 对辣椒进行代谢物的差异分析, 并结合机器学习来进行品种鉴别, 探究品种差异对辣椒代谢物形成的影响, 以期为辣椒种类识别和产品加工提供理论支撑。

1 材料与方法

1.1 材料与仪器

红辣椒: 艳椒(YJ)、辣研(LY)、国塔(GT)、研椒 110(YJA)、热辣(RL) 产自河北望都(图 1); 甲醇 色谱纯, 德国默克公司; N-甲基-N-(三甲基硅烷)三氟乙酰胺(MSTFA)(纯度≥99%)、吡啶(色谱纯)、甲氧胺盐酸盐(纯度≥99%)、抗坏血酸标品(纯度≥99%)、辣椒碱标品(纯度≥99%)、二氢辣椒碱标品(纯度≥99%) 上海阿拉丁生化科技股份有限公司; 核糖醇 纯度≥98%, 上海源叶生物科技有限公司。

CNWB-15ZK 电磁波真空干燥机 杭州三特医

药化工设备有限公司; Allegra X-12R 冷冻离心机 美国贝克曼库尔特有限公司; JT-DCY-12Y 水浴氮吹仪 杭州聚同电子有限公司; CM-5 台式分光测色计 日本 Konica Minolta 公司; CHA-2 恒温气浴箱 苏州威尔实验用品有限公司; Agilent 7890-5975C 气相色谱串联质谱仪 安捷伦科技有限公司; e2695 高效液相色谱仪 美国沃特世公司。



图 1 不同品种望都辣椒图

Fig.1 Different varieties of Wangdu chili peppers

注: 从左到右依次为艳椒(YJ)、辣研(LY)、国塔(GT)、研椒110(YJA)、热辣(RL)。

1.2 实验方法

1.2.1 色差测定 取五种辣椒样品, 使用分光测色计对其色度进行测定, 记录 L^* 、 a^* 、 b^* 值。其中, L^* 表示样品色泽的亮度, a^* 值表示红度或绿度, b^* 值表示黄度或蓝度。每组样品平行测定 3 次^[10]。

1.2.2 样品处理 取辣椒样品, 40 °C 真空干燥, 粉碎过 40 目筛, 干燥保存待用。

1.2.3 辣椒碱与二氢辣椒碱的测定 称取 50 mg 样品, 加入 1.5 mL 75% 的乙醇, 超声处理(400 W、40 kHz、60 °C、50 min), 离心分离 10 min(5000 r/min、4 °C), 取上清液, 加入 1.5 mL 75% 的乙醇, 混匀, 重复操作 2 次, 定容至 5.0 mL, 过 0.22 μm 微孔滤膜, 待测。

色谱条件: 参考刘录等^[11]的方法上机, 色谱柱为 xBridge® C₁₈(5 μm, 4.6×250 mm), 甲醇: 3% 磷酸水溶液梯度洗脱, 流速 1 mL/min, 检测波长 280 nm, 进样量 2 μL, 柱温 40 °C。

1.2.4 V_C 含量的测定 取 1.2.2 制备的样品 1 g, 按 1:20 固液比加入水, 15000 r/min 均质提取 2 min, 离心 3 min(4000 r/min、4 °C)后, 将上清液移至 50 mL 容量瓶中并定容, 按上述方法重复提取 1 次。参考蒋依辉等^[12]的上机方法。

色谱条件: 色谱柱, Atiantis T3 柱(5 μm, 4.6×250 mm); 检测波长 210 nm, 流动相为甲醇: 0.1% 的磷酸=2.5:97.5(V/V)梯度洗脱, 流速 1 mL/min, 柱温: 35 °C。

1.2.5 代谢产物的测定

1.2.5.1 样品预处理 取 50 mg 辣椒样品到 EP 管中, 每种辣椒取 9 份, 加入 1 mL 水(含 25 μg 核糖醇), 混匀, 室温下 500 W 超声提取 15 min, 10000 r/min 离心 15 min。将上清液转移到新的 EP 管中, 取 1 mL 甲醇对残渣进行同样的处理。将两种上清液混合进行衍生化(胺化)反应, 取 500 μL 的混合上清液于 EP 管中, 用氮吹仪吹干, 加入 100 μL 甲氧胺盐试剂(甲氧胺盐酸盐溶于吡啶 20 mg/mL), 混匀后, 室温下将混合物 500 W 超声处理 5 min, 80 °C 孵育 30 min; 加入 100 μL 双(三甲基硅基)三氟乙酰胺, 70 °C 孵育 1.5 h, 冷却至室温; 过 0.22 μm 微孔滤膜, 上机检测。每个待测样本提取液等体积混合后进行衍生化反应, 制成 QC 样品, 在上机时等间隔插入 QC 样品进样。

1.2.5.2 GC 条件 色谱柱: HP-5MS 毛细管柱(30 m×250 μm×0.25 μm), 载气: 氮气; 进样模式: 不分流; 进样量: 1 μL; 升温程序: 起始温度 80 °C, 保持 2 min, 以 5 °C/min 的速度升温至 300 °C 并保持 12 min。

1.2.5.3 MS 条件 质谱条件: 全扫描方式; 离子源温度 240 °C; 溶剂延迟 5 min; 质量扫描范围 m/z 50~650。界面温度 280 °C; 电压 1.2 kV; 电子能量 70 eV; 扫描速度 5 scans/s。采用 NIST14 数据库进行定性分析, 最终得到代谢物鉴定结果。

1.3 数据处理

试验结果为 3 次重复的平均值, 以平均值±标准偏差表示, 采用 Excel 2019 进行数据整理、标准偏差计算, 采用 SPSS 22 统计软件进行差异显著性分析, 采用 Origin 2022 软件进行相关分析和绘图。

采用 SIMCA14.1 软件进行主成分分析(Principal Component Analysis, PCA)、正交偏最小二乘分析(Orthogonal Partial Least Squares Discriminant Analysis, OPLS-DA)、多元统计分析。首先通过主成分分析(PCA)对数据进行降维处理, 并对整体的分布趋势进行预览, 然后采取正交偏最小二乘-判别分析(OPLS-DA), 以 $P<0.05$ 和 $VIP\geq 1$ 作为截断标准, 筛选差异代谢物, 利用 Metaboanalyst 6.0(<http://www.metaboanalyst.ca>)对差异代谢物进行聚类热图分析, 用 KEGG 通路对差异代谢物进行分析, 并用显著性分析(P value)确定相关通路, 最后利用 SPSSPRO 中三种机器学习方法(随机森林、XGBoost 和 BP 神经网络)进行建模, 通过差异代谢物来对品种分类验证。

2 结果与分析

2.1 不同品种辣椒色泽分析

对艳椒(YJ)、辣研(LY)、国塔(GT)、研椒110(YJA)、热辣(RL)等五种不同品种的辣椒进行色差分析。结果如图 2 所示, 不同辣椒样品的亮度、红度、黄度均有差别, 其中 RL 的 L^* 、 a^* 、 b^* 值最大, 为 37.99、35.81、23.67, 且与其余样品间存在显著性差

异($P<0.05$)。辣椒果皮中的辣椒红素、玉米黄质、 β -胡萝卜素、叶绿素 a、叶绿素 b 和飞燕草素, 是影响果色的主要物质^[13], 说明 RL 的辣椒红素等类胡萝卜素累计最多, 而其余品种辣椒色差度并不显著, 因此不能通过色泽来进行区分识别。

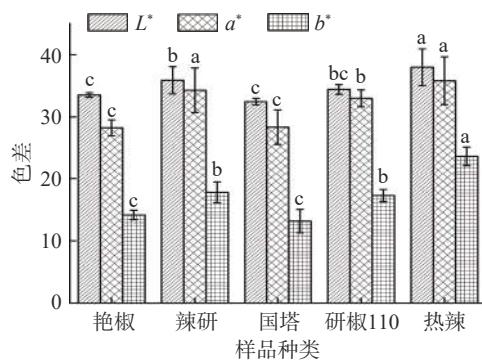


图 2 不同样品的色差测定

Fig.2 Color difference determination of different samples
注: 同一指标不同字母表示不同辣椒样品差异显著($P<0.05$)。

2.2 不同品种辣椒碱、二氢辣椒碱含量分析

辣椒碱是一种存在于辣椒中的生物碱, 赋予辣椒特有的辛辣味道。辣味是辣椒碱与疼痛感觉神经系统中的瞬时受体电位—香草酸受体 1(TRPV1), 即辣椒素受体相结合, 出现灼烧和疼痛等感觉^[14], 然后发展为辛辣感。辣椒碱类物质具有抗氧化^[15]、抗肿瘤^[16]、降脂^[17]等活性。此外, 人们对辣味的接受程度因人而异, 因此辣椒碱类物质的含量是影响辣椒品质的一项重要指标。从图 3 可以看出, 样品 RL 的辣椒碱与二氢辣椒碱含量最高, 分别为 533.897 ± 62.187 、 $264.526\pm28.532 \mu\text{g/g}$, 说明 RL 品种辣椒的辣度最辣, 与朱志妍等^[18]研究相比结果偏高, 可能是辣椒的品种不同导致。对于艳椒(YJ)、辣研(LY)、国塔(GT)、研椒 110(YJA)和热辣(RL)来说, 利用辣椒碱、二氢辣椒碱含量仅仅可以区分出热辣辣椒产品。

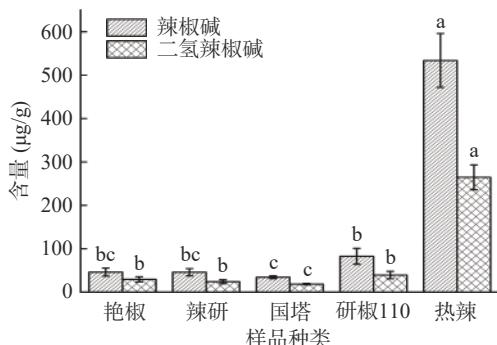


图 3 不同样品中辣椒碱、二氢辣椒碱含量

Fig.3 Contents of capsaicin and dihydrocapsaicin in different samples
注: 同一指标不同字母表示不同辣椒样品差异显著($P<0.05$)。

2.3 不同品种辣椒 V_C 含量分析

辣椒应用价值丰富, 含有丰富的 V_C , 其含量在

蔬菜中占首位, 是番茄的 7 倍~15 倍^[19]。辣椒中含有的 V_C , 作为天然抗氧化剂, 可消除有害氧自由基的毒性, 可用于抗氧化^[20]与抗癌^[21]。五种辣椒中 V_C 含量见图 4, 从图中可以看出, 五种辣椒中 V_C 含量范围在 $85.4\sim146.9 \text{ mg}/100 \text{ g}$ 之间, 与王蜜蜜等^[22]的研究结果相当。样品 YJ 中 V_C 的含量最高, 为 $146.9\pm2.9 \text{ mg}/100 \text{ g}$, LY 中 V_C 含量最低, 为 $85.4\pm6.3 \text{ mg}/100 \text{ g}$ 。利用 V_C 含量, 仅可以区分出艳椒(YJ)、研椒 110(YJA)和热辣(RL), 而辣研(LY)和国塔(GT)不容易区分开来。

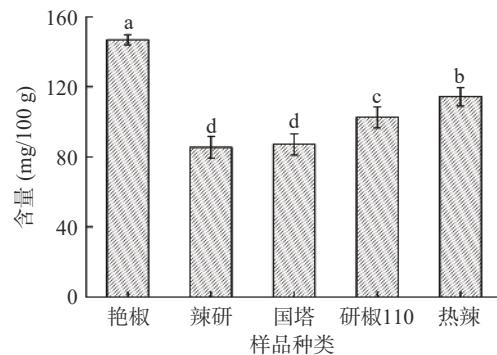
图 4 不同样品中 V_C 含量测定

Fig.4 Determination of V_C content in different samples
注: 同一指标不同字母表示不同辣椒样品差异显著($P<0.05$)。

2.4 辣椒代谢产物分析

2.4.1 PCA 分析 考虑到利用单一理化指标无法对不同品种的辣椒进行识别, 进一步利用 GC-MS 非靶向代谢组学进行差异代谢物筛选。基于 GC-MS 分析, 在五种辣椒样品中共得到 95 种代谢物, 为清晰得出它们之间的代谢物差异, 首先利用主成分分析法对五种辣椒代谢物进行降维处理, 得到二维图进而判别辣椒之间的差异。在无监督 PCA 得分图中, 每个实心圆代表一个样本, 对代谢物进行主成分分析可以总体反映样本组间和组内差异, 结果如图 5 所示。由图可知, 得分图的所有样本基本都在 95% 置信椭圆内, 但无法得到有效分离, 且 PC1 与 PC2 只解释了总方差的 34%, 因此还需进一步采用有监督的识别模式进行分析。

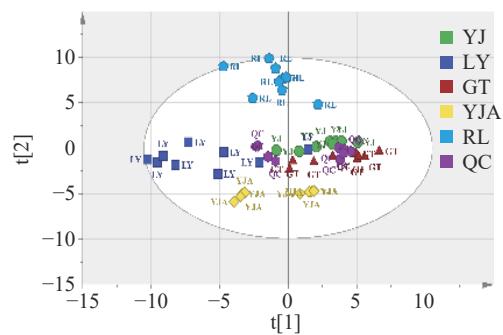
 $R^2X[1]=0.18 \quad R^2X[2]=0.16 \quad \text{Ellipse: Hotelling's } T^2(95\%)$

图 5 PCA 得分图

Fig.5 Score chart of PCA

2.4.2 OPLS-DA 分析 为了更好地实现高水平的组间分离,采用有监督的 OPLS-DA 进行分析^[23]。在 OPLS-DA 模型中, R^2Y 和 Q^2 的值常用于判断其稳定性和观察原始数据的可靠性,两者的值越接近 1, 模型越可靠。如图 6 所示, $R^2Y=0.838$, $Q^2=0.716$, 表明模型能很好地解释样本间的差异。为防止模型过拟合,采用 200 次响应的置换检验(Response Permutation Testing, RPT 检验)对模型进行验证,其中 R^2 、 Q^2 为回归直线与 y 轴的截距值, R^2 表示模型能够解释的方差总和, Q^2 表示模型的预测能力; 使用 RPT 检验时,一般要求 $Q^2<0$ 。由图 7 可知, $R^2=0.204$, $Q^2=-0.407$, 回归线呈向上的趋势, $Q^2<0$, 没有出现过拟合的现象,说明 OPLS-DA 模型可靠,可用于后续的分析。

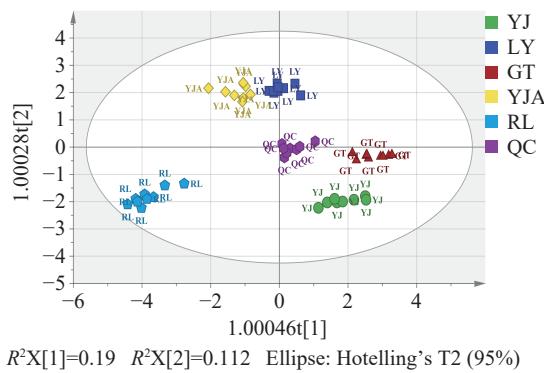


图 6 OPLS-DA 得分图

Fig.6 Score chart of OPLS-DA

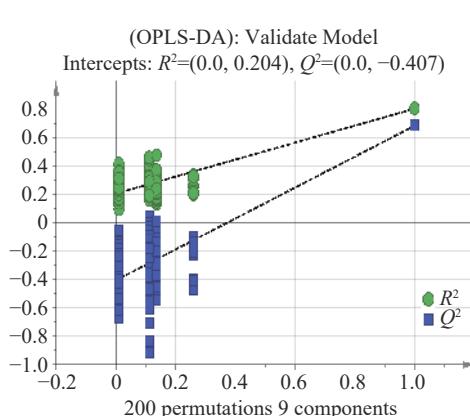


图 7 OPLS-DA 模型置换检验

Fig.7 OPLS-DA model permutation test

2.4.3 差异代谢物筛选 采用多维分析和单维分析相结合的方法,可以进一步筛选组间差异代谢物。根据 OPLS-DA 模型分析结果,筛选主成分的 VIP 值不低于 1、 t 检验 $P<0.05$ 的差异代谢物进行汇总(表 1),共筛选出 16 种差异代谢物,含有糖类、醇类、有机酸、氨基酸及其衍生物等物质,每种辣椒的各类代谢物如图 8 所示,在 YJ、LY、GT 和 YJA 中有机酸的相对含量最高,而 RL 品种中糖类的相对含量较高,而氨基酸及其衍生物在所有品种的辣椒中相对含量最少,说明可能有机酸、糖是区分 5 种辣椒的

表 1 不同品种辣椒代谢差异物筛选
Table 1 Screening of metabolic differences in different varieties of chili peppers

代谢物	英文对照	P值	VIP
乳酸	Lactic acid	0.001	1.676
蔗糖	Sucrose	0.001	2.090
苹果酸	Malic acid	0.001	1.533
肌醇	Myo-Inositol	0.001	1.689
山梨醇	D-Sorbitol	0.001	1.784
天门冬氨酸	Asparagine	0.001	1.326
奎宁酸	Quinic acid	0.003	1.580
D-果糖	D-Fructose	0.006	1.035
柠檬酸	Citric acid	0.007	3.143
葡萄糖	Glucose	0.008	1.517
2-甲氧基雌二醇	2-Methoxyestradiol	0.019	1.574
β -D-吡喃葡萄糖醛酸	β -D-Glucopyranuronic acid	0.008	1.026
苏氨酸-2,5-外二脲	threo-2,5-Hexodiulose	0.011	1.144
D-塔格糖	D-Tagatose	0.027	2.606
D-甘露糖	d-Mannose	0.033	2.764
乌头酸	Aconitic acid	0.0422	2.297

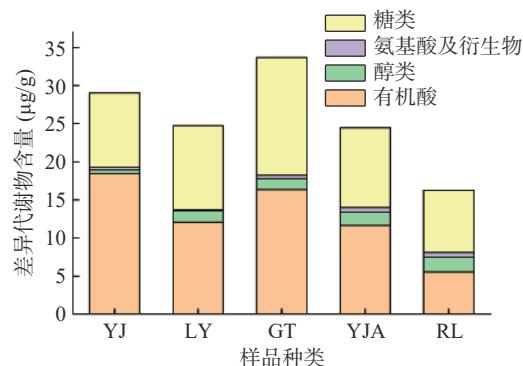


图 8 不同品种辣椒差异代谢物含量图

Fig.8 Difference in metabolite content among different varieties of chili peppers

关键差异代谢物。

为了更直观地观察差异代谢物在五种辣椒间的含量差异,对筛选出的差异化合物进行聚类,绘制所有差异代谢物的层次聚类热图,如图 9 所示。在五种辣椒中,差异代谢产物有明显不同,以代谢产物类别差异对辣椒进行分析,差异代谢物中有机酸有 6 种,包含乳酸、柠檬酸、奎宁酸、乌头酸、 β -D-吡喃葡萄糖醛酸和苹果酸。乳酸在品种 GT 中含量明显高于其他品种,其可能由糖酵解生成,也可能由苹果酸脱羧生成^[24]; 品种 GT 跟 LY 中富含柠檬酸,它是柠檬酸循环与乙醛酸和二羧酸代谢的中间产物,是辣椒能量代谢和新陈代谢过程中的重要组成部分; 奎宁酸、乌头酸和苹果酸在辣椒中分布广泛,其中奎宁酸影响辣椒的生长过程与代谢过程,通常涉及细胞壁合成、酸碱平衡和抗氧化活性等,而乌头酸是一种食用增味剂,可用于中药,这两种有机酸在 YJ 中含量最高,苹果酸则为水果的重要成分,酸爽可口,在 YJ、

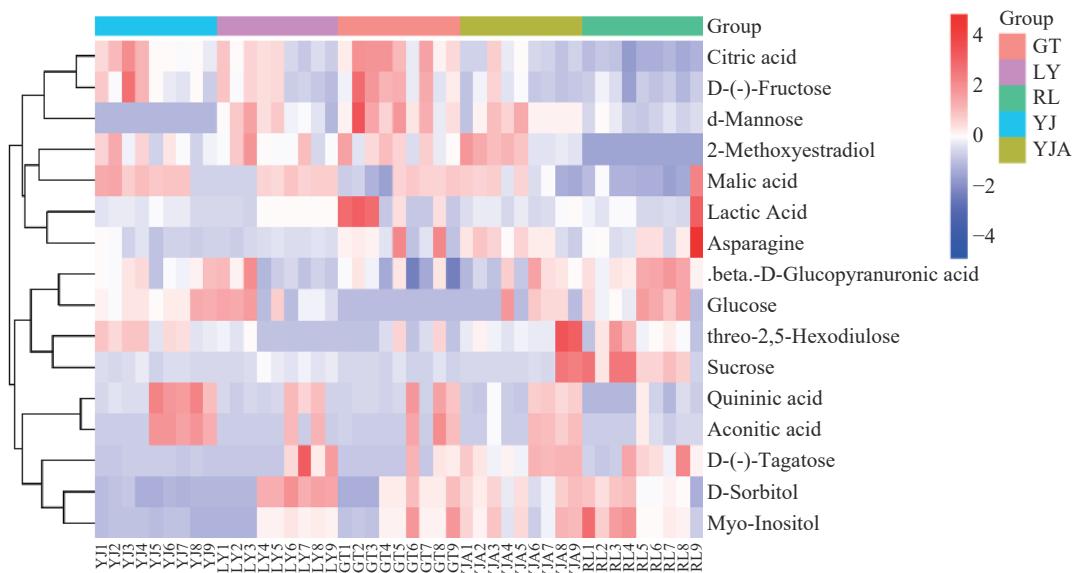


图 9 不同样品组间差异代谢物热图

Fig.9 Heat map of differential metabolites between different sample groups

LY 和 GT 当中比较丰富。

醇类作为广泛分布在植物中的生物标记物之一, 差异代谢物中的醇类有三种为 D-山梨醇、肌醇和 2-甲氧基雌二醇, 肌醇是化学上非常稳定的极性小分子, 是一种碳环糖多元醇, 具有抗炎、抗氧化和降脂活性^[25], 其在 RL 中含量最高; D-山梨醇改善辣椒果实的着色, 影响辣椒的糖代谢^[26], 五种辣椒中, LY 的 D-山梨醇含量最高; 2-甲氧基雌二醇有抗癌抗肿瘤的生物活性, 并不是天然存在的, 可能是在衍生化的过程中产生。糖类代谢物包含葡萄糖、蔗糖、D-果糖、D-塔格糖和 D-甘露糖, 其中, D-果糖相对含量高于其它代谢物, 参与果糖与甘露糖代谢, 影响辣椒生长和发育, 同时也可以影响辣椒的口感和风味特性, 在 GT 中, D-果糖与 D-甘露糖含量明显高于其它品种, 而葡萄糖与蔗糖在 RL 中含量最高, D-塔格糖在 YJA 中最丰富。

氨基酸及其衍生物有两种, 为天门冬氨酸与苏氨酸-2,5-外二脲, 天冬氨酸被称为味氨基酸, 呈鲜味, 能影响果实风味^[27], 可以优化根系形态, 促进植物对养分的吸收^[28], 在品种 YJA 跟 GT 中含量较高; 苏氨酸是人体八种必需氨基酸之一, 对于生物体的生长和发育具有重要作用^[29], 在品种 YJA 广泛存在。氨基酸在辣椒的代谢通路中也广泛存在, Lozada 等^[30]揭示了 Jalapeño 和 Serrano 两种辣椒差异代谢组分包含有机酸、核苷酸及其衍生物、氨基酸及其衍生物、生物碱等, 其中参与了萜类主干生物合成和光合作用。Tilen 等^[31]对不同辣椒品种果实代谢差异的 HPLC/MS 分析, 分析了包括糖类、有机酸、辣椒素、酚类物质等 16 种代谢产物。在三个品种中, 果糖是最常见糖, 其次为葡萄糖和蔗糖, 这与本文研究结果相似, 有机酸中的柠檬酸、苹果酸、奎宁酸和草酸含量都有所差异, 抗坏血酸仅在一个品种中有差异, 不

同品种间辣椒素总含量无显著差异。Nemesio 等^[32]对墨西哥的 10 个新品种辣椒核磁共振代谢组学分析, 发现柠檬酸、甲酸、苹果酸、果糖、蔗糖和半乳糖为差异代谢物, 对不同品种的辣椒分类有重要作用。

2.4.4 代谢通路分析

为了进一步了解不同品种辣椒代谢途径差异, 运用 Metaboanalyst 6.0 对五种辣椒进行富集分析。将筛选出的 16 个差异性代谢物投影至 KEGG 数据库中的 Pathway 数据库并与其进行匹配, 由 $P < 0.05$ 筛选出 6 条主要的代谢通路, 主要为糖醇代谢与有机酸代谢, 分别为半乳糖代谢、果糖和甘露糖代谢、柠檬酸循环、淀粉和蔗糖代谢、乙醛酸和二羧酸代谢、丙酮酸代谢(图 10)。

半乳糖代谢、果糖和甘露糖代谢、淀粉和蔗糖代谢主要作用于辣椒的糖代谢, 分解和生成相关糖类物质; 柠檬酸循环、乙醛酸和二羧酸代谢、丙酮酸代谢涉及乌头酸、柠檬酸、苹果酸和乳酸, 是碳水化合物、脂肪和氨基酸的最终常见氧化途径, 它是身体能量供应最重要的代谢途径, 也是连接几乎所有个体代谢途径的最重要的中枢途径, 其中一些二羧酸与丙酮酸是一种重要的代谢中间产物, 参与到柠檬酸循环中。

2.4.5 机器学习识别

为了对辣椒的品种识别进行验证, 采用了三种机器学习方法提高验证结果的可靠性, 以 16 种差异代谢物作为数据集减少运算的复杂度, 将数据集随机划分为训练集(70%)和测试集(30%)进行机器学习方法测试, 采用 XGBoost、随机森林(RF)和 BP 神经网络三种方法建立分类模型, 为了减小过拟合现象, 使用十折交叉验证, 比较不同品种辣椒的分类准确率。结果由表 2 可知, 在预测准确率上, RF 模型与 XGBoost 模型识别准确率较高, 分别为 100% 和 92.9%, 能够得到预期结果, 说明上述数据可以完成分类识别。

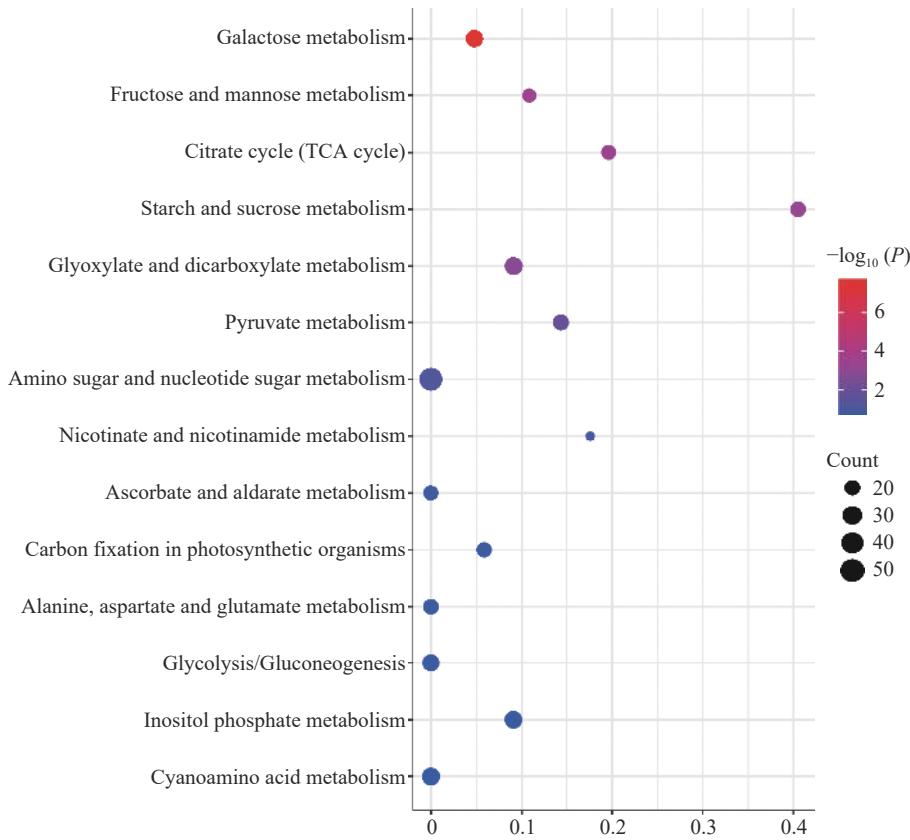


图 10 不同品种辣椒差异代谢物代谢通路图

Fig.10 Metabolic pathways of different metabolites in different varieties of chili peppers

表 2 三种机器学习方法对辣椒的识别准确率(%)

Table 2 Recognition accuracy of two machine learning methods for chili peppers (%)

模型	RF	XGBoost	BP神经网络
训练集	100	100	100
测试集	100	92.9	78.6

3 结论
本研究对五个品种的望都辣椒的色泽、 V_C 、辣椒碱类的含量及代谢产物进行了检测和统计分析,五种辣椒中 RL 品种辣椒的 L^* 、 a^* 和 b^* 值为最高,而其余品种色泽度区分不明显;辣椒碱类物质中,品种 RL 含量最高,辣度最辣,其余品种差异并不显著;而利用 V_C 含量,艳椒(YJ)的含量最高,而辣研(LY)和国塔(GT)含量最低,不容易区分。采用 GC-MS 非靶向代谢组学方法,利用 PCA 分析与 OPLS-DA 分析,筛选出 16 种差异代谢物,包含有机酸、醇类、糖类和氨基酸及其衍生物,其中在艳椒(YJ)、辣研(LY)、国塔(GT)和研椒 110(YJA)中有机酸的相对含量最高,而热辣(RL)品种中糖类的相对含量较高,不同品种辣椒所富集的差异代谢物也不同,奎宁酸和乌头酸等有机酸在艳椒(YJ)中含量较高,乌头酸为中药成分,故艳椒(YJ)可以应用于药品开发;D-山梨醇在辣研(LY)中含量最高,国塔(GT)当中柠檬酸、D-果糖、D-甘露糖和乳酸富集程度最高,可应用于功能性食品开发,研椒 110(YJA)中 D-塔格糖与氨基

酸含量最高,故可以应用于香料制作,而热辣(RL)中葡萄糖跟肌醇含量在所有品种中占优势,该结果可以作为区分不同品种辣椒的判断指标。最后利用机器学习对辣椒的差异代谢物进行识别验证,RF 模型、XGBoost 和 BP 神经网络模型识别准确率分别为 100%、92.9% 和 78.6%,RF 模型、XGBoost 能够达到预期结果,而 BP 神经网络模型准确率较低,还需要进一步训练优化。本研究在为进一步进行品种鉴别提供了参考,对开发品质优良的新品种发挥积极作用。

© The Author(s) 2024. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

参考文献

- [1] GUO G J, PAN B G, YI X X, et al. Genetic diversity between local landraces and current breeding lines of pepper in China[J]. *Scientific Reports*, 2023, 13(1): 4058.
- [2] 邹学校, 杨莎, 朱凡, 等. 中国高口感品质鲜食辣椒产业发展与未来趋势[J]. 园艺学报, 2024, 51(1): 27–38. [ZOU X X, YANG S, ZHU F, et al. The development and future trends of China's high taste and quality fresh chili industry[J]. *Journal of Horticulture*, 2024, 51(1): 27–38.]
- [3] 张子峰. 我国辣椒产业发展现状、主要挑战与应对之策[J]. 北方园艺, 2023(14): 153–158. [ZHANG Z F. The current development status, main challenges, and countermeasures of China's chili industry[J]. *Northern Horticulture*, 2023(14): 153–158.]

- [4] TANJA C, JASMINA R, STELA J. Quality of pepper seed by-products: A review [J]. *Foods*, 2022, 11(5): 748.
- [5] NACCARATO A, FURIA E, SINDONA G, et al. Multivariate class modeling techniques applied to multielement analysis for the verification of the geographical origin of chili pepper [J]. *Food Chemistry*, 2016, 206: 217–222.
- [6] IVAN I M, POPOVICI V, CHITESCU C L, et al. Phytochemical profile, antioxidant and cytotoxic potential of *Capsicum annuum*(L.) dry hydro-ethanolic extract [J]. *Pharmaceutics*, 2024, 16(2): 245.
- [7] 沈央红, 方金玉, 朱军莉, 等. 代谢组学在食品质量安全领域的应用进展 [J]. *食品科学*, 2023, 44(5): 282–289. [SHEN Y H, FANG J Y, ZHU J L, et al. The application progress of metabolomics in the field of food quality and safety [J]. *Food Science*, 2023, 44(5): 282–289.]
- [8] 杨创创, 何建文, 张正海, 等. 13 份干辣椒挥发性风味品质评价 [J]. *中国蔬菜*, 2023(10): 37–49. [YANG C C, HE J W, ZHANG Z H, et al. Evaluation of volatile flavor quality of 13 dried chili peppers [J]. *Chinese Vegetables*, 2023(10): 37–49.]
- [9] ELVIA B M, ELIDETH F R, NURY P H, et al. ¹H NMR-based metabolomic fingerprinting to determine metabolite levels in serrano peppers (*Capsicum annuum* L.) grown in two different regions [J]. *Food Research International*, 2017, 102: 163–170.
- [10] 杨芳, 邓凤琳, 袁海彬, 等. 基于气相色谱-离子迁移谱结合多元统计方法研究辣椒产地对辣椒油理化性质和风味成分的影响 [J]. *核农学报*, 2023, 37(7): 1393–1402. [YANG F, DENG F L, YUAN H B, et al. Based on gas chromatography ion migration spectroscopy combined with multivariate statistical methods, the influence of chili pepper production areas on the physicochemical properties and flavor components of chili oil [J]. *Journal of Nuclear Agriculture*, 2023, 37(7): 1393–1402.]
- [11] 刘录, 陈海云, 刘菲菲, 等. HPLC 法测定云南省不同辣椒品种中辣椒碱及二氢辣椒碱的含量 [J]. *食品工业科技*, 2013, 34(5): 299–303. [LIU L, CHEN H Y, LIU F F, et al. HPLC method for determining the content of capsaicin and dihydrocapsaicin in different pepper varieties in Yunnan Province [J]. *Food Industry Technology*, 2013, 34(5): 299–303.]
- [12] 蒋依辉, 凡超, 刘伟, 等. 运用 RP-HPLC 同时测定火龙果中 7 种有机酸和 Vc 含量的方法研究 [J]. *江西农业大学学报*, 2015, 37(5): 798–803. [JIANG N H, FAN C, LIU W, et al. Study on the simultaneous determination of seven organic acids and Vc content in dragon fruit using RP-HPLC [J]. *Journal of Jiangxi Agricultural University*, 2015, 37(5): 798–803.]
- [13] 李平平, 张祥, 刘雨婷, 等. 辣椒 63 份种质果皮颜色与呈色物质的关系 [J]. *园艺学报*, 2022, 49(7): 1589–1601. [LI P P, ZHANG X, LIU Y T, et al. The relationship between peel color and coloring substances of 63 chili germplasms [J]. *Journal of Horticulture*, 2022, 49(7): 1589–1601.]
- [14] 程杰, 罗洲, 肖书剑, 等. 辣椒、花椒及其制品辣度和麻度检测研究现状 [J]. *中国调味品*, 2023, 48(6): 209–215. [CHENG J, LUO Z, XIAO S J, et al. Current research status on the detection of spiciness and numbness in chili peppers, Sichuan peppercorns and their products [J]. *Chinese Seasoning*, 2023, 48(6): 209–215.]
- [15] KEITHELLAKPAM S. Antioxidant potential and factors influencing the content of antioxidant compounds of pepper: A review with current knowledge [J]. *Comprehensive Reviews in Food Science and Food Safety*, 2023, 22(4): 3011–3052.
- [16] WANG W L, HUANG Y L, HU W X, et al. Synthesis, structures and anti-tumor properties of capsaicin analogues [J]. *Chemical Journal of Chinese Universities-Chinese*, 2010, 31(12): 2400–2406.
- [17] LEE H, SONG J, CHUNG S, et al. Gochujang elicits anti-obesity effects by increasing capsaicin-independent brown adipogenesis and thermogenesis in high-fat diet induced obese mice [J]. *Journal of Functional Foods*, 2023, 111: 239.
- [18] 朱志妍, 田浩, 潘俊, 等. 云南省主要辣椒品种活性成分及其加工适应性 [J]. *食品安全质量检测学报*, 2022, 13(21): 7091–7100. [ZHU Z Y, TIAN H, PAN J, et al. The active ingredients and processing adaptability of the main chili pepper varieties in Yunnan Province [J]. *Journal of Food Safety and Quality Testing*, 2022, 13(21): 7091–7100.]
- [19] XU Z N, ZHONG D J, WANG J Y, et al. Colour, physico-chemical, microbiological and bioactive component quality analysis of peppers in Guizhou [J]. *International Journal of Food Science & Technology*, 2022, 58(2): 665–675.
- [20] HUNGERFORD A J, BAKOS H W,AITKEN R J. Addition of vitamin C mitigates the loss of antioxidant capacity, vitality and DNA integrity in cryopreserved human semen samples [J]. *Antioxidants*, 2024, 13(2): 247.
- [21] CHEN D, WEI X Y, YANG K, et al. Piperlongumine combined with vitamin C as a new adjuvant therapy against gastric cancer regulates the ROS-STAT3 pathway [J]. *The Journal of International Medical Research*, 2022, 50(4): 3000605221093308.
- [22] 王蜜蜜, 胡能兵, 庞丹丹, 等. 辣椒主要营养品质比较 [J]. *安徽科技学院学报*, 2019, 33(6): 38–42. [WANG M M, HU N B, PANG D D, et al. Comparison of the main nutritional qualities of chili peppers [J]. *Journal of Anhui University of Science and Technology*, 2019, 33(6): 38–42.]
- [23] 李萌, 卢奎丞, 黄云艳, 等. 基于代谢组学技术分析不同纬度山羊乳的差异性 [J]. *分析化学*, 2021, 49(11): 1864–1875. [LI M, LU X C, HUANG Y Y, et al. Analysis of differences in goat milk at different latitudes based on metabolomics techniques [J]. *Analytical Chemistry*, 2021, 49(11): 1864–1875.]
- [24] CHEN D, LIU S Q. Transformation of chemical constituents of lychee wine by simultaneous alcoholic and malolactic fermentations [J]. *Food Chemistry*, 2016, 196.
- [25] ROSTAMI S, AREFHOSSEINI S, TUTUNCHI H, et al. Does myo-inositol supplementation influence oxidative stress biomarkers in patients with non-alcoholic fatty liver disease? [J]. *Food Science & Nutrition*, 2024, 12(2): 1279–1289.
- [26] 周平, 郭瑞, 颜少宾, 等. 外源山梨醇影响桃叶片和果实糖代谢的分子机制研究 [J]. *园艺学报*, 2023, 50(5): 959–971.
- [27] ZHOU P, GUO R, YAN S B, et al. A molecular mechanism study on the effect of exogenous sorbitol on sugar metabolism in peach leaves and fruits [J]. *Journal of Horticulture*, 2023, 50(5): 959–971.]
- [28] 彭真汾, 王威, 叶清华, 等. 高效液相色谱-串联质谱法定量分析橄榄果实氨基酸组分 [J]. *食品科学*, 2018, 39(24): 231–238. [PENG Z F, WANG W, YE Q H, et al. Quantitative analysis of amino acid components in olive fruit using high-performance liquid chromatography tandem mass spectrometry [J]. *Food Science*, 2018, 39(24): 231–238.]
- [29] 刘媛, 袁亮, 张水勤, 等. 不同分子量聚天冬氨酸对小麦根系

- 生长和养分吸收的影响[J]. *中国农业科学*, 2022, 55(13): 2526–2537. [LIU Y, YUAN L, ZHANG S Q, et al. The effects of polyaspartic acid with different molecular weights on the growth and nutrient uptake of wheat roots[J]. *Chinese Agricultural Science*, 2022, 55(13): 2526–2537].]
- [29] DANIEL P F, DIRLEISE P, EDUARDO K B, et al. L-threonine requirement of silver catfish (*Rhamdia quelen*): Growth performance, body composition, plasma and liver metabolites[J]. *Journal of Applied Animal Research*, 2023, 51(1): 358–365.
- [30] LOZADA D N, PULICHERLA S R, HOLGUIN F O. Wide-ly targeted metabolomics reveals metabolite diversity in *Jalapeno* and *Serrano chile* peppers (*Capsicum annuum* L.)[J]. *Metabolites*, 2023, 13(2): 288.
- [31] TILEN Z, ALJAŽ M, ROBERT V, et al. Metabolic variation among fruits of different chili cultivars (*Capsicum* spp.) using HPLC/MS[J]. *Plants*, 2021, 11(1): 101.
- [32] NEMESIO V R, MOISÉS R M, REINALDO M A, et al. ¹H NMR-based metabolomics profiling of ten new races from *Capsicum annuum* cv. *serrano* produced in Mexico[J]. *Food Research International*, 2019, 119: 785–792.