非靶代谢组学实验和分析方法

1.试剂和仪器信息

表 1. 实验试剂列表

名称	CAS	纯度	品牌
甲醇（Methanol）	67-56-1	LC-MS 级	CNW Technologies
乙腈（Acetonitrile）	75-05-8	LC-MS 级	CNW Technologies
乙酸铵（Ammonium acetate）	631-61-8	LC-MS 级	SIGMA-ALDRICH
氨水(Ammonium hydroxide)	1336-21-6	LC-MS 级	Fisher Chemical
超纯水(ddH2O)	-	-	屈臣氏

表 2. 实验仪器列表

仪器	型号	品牌
超高效液相	Vanquish	Thermo Fisher Scientific
高分辨质谱	Orbitrap Exploris 120	Thermo Fisher Scientific
离心机	Heraeus Fresco17	Thermo Fisher Scientific
天平	BSA124S-CW	Sartorius
超声仪	PS-60AL	深圳市雷德邦电子有限公司

2. 实验流程

2.1 代谢物提取

移取 100 μL 样品至 EP 管中，加入 400 μL 提取液（甲醇：乙腈=1：1（V/V），含同位素标记内标混合物），涡旋混匀 30 s；

超声 10 min（冰水浴）；

-40 ℃静置 1 h；

将样品 4 ℃，12000 rpm（离心力 13800(×g)，半径 8.6cm）离心 15 min；

取上清于进样瓶中上机检测；

- 所有样品另取等量上清混合成 QC 样品上机检测。

2.2 上机检测

本项目使用 Vanquish (Thermo Fisher Scientific)超高效液相色谱仪，通过 Waters ACQUITYUPLC BEH Amide (2.1 mm × 50 mm, 1.7 μm)液相色谱柱对目标化合物进行色谱分离 [7]。液相色谱 A 相为水相，含 25 mmol/L 乙酸铵和 25 mmol/L 氨水，B 相为乙腈。样品盘温度：4 ℃，进样体积：2 μL。

Orbitrap Exploris 120 质谱仪能够在控制软件（Xcalibur, 版本：4.4，Thermo）控制下进行一级、二级质谱数据采集。详细参数如下： sheath gas flow rate as 50 Arb, Aux gas flow rate as 15 Arb, capillary temperature 320 ℃, full MS resolution as 60000, MS/MS resolution as15000, collision energy: SNCE 20/30/40, spray voltage as 3.8 kV (positive) or -3.4 kV (negative), respectively.

2.3 数据处理

原始数据经 ProteoWizard 软件转成 mzXML 格式后，使用自主编写的 R 程序包（内核为 XCMS）进行峰识 别、峰提取、峰对齐和积分等处理 [8]，然后与自建二级质谱数据库匹配进行物质注释，算法打分的 Cutoff 值设为 0.3。

3.信息分析流程

采用 XCMS 软件对采集到的质谱数据进行峰拾取、峰分组、保留时间校正、二次峰分组、同位素和加合物标注等预处理。将 LC - MS 原始数据文件转换成 mzXML 格式，然后用 R 软件实现的 XCMS、CAMERA 和 metaX 工具箱进行处理。结合保留时间(RT)和 m/z 数据对各离子进行鉴定。记录每个峰的强度，并生成一个三维矩阵，其中包含任意指定的峰指数(保留时间-m/z 对)、样品名称(观察值)和离子强度信息(变量)。

利用在线 KEGG, HMDB 数据库对代谢物进行标注，将样品的精确分子质量数据(m/z)与数据库中的数据进行匹配。如果观测值与数据库值之间的质量差小于 10 ppm，则将对代谢物进行注释，并通过同位素分布测量进一步鉴定和验证代谢物的分子式。我们还使用了 in-house 数据库来验证代谢物鉴定。

数据的统计分析主要由 R 软件（version 4.0）完成，蛋白的原始强度值会经过中位数（medium）归一化，聚类热图由 R 包 pheatmap 绘制，PCA 分析和显著差异蛋白分析由 R 包 metaX 完成，PLSDA 分析由 R 包 ropls 进行，并计算各变量的 VIP 值，相关性分析由 R 包 cor 的 Pearson 相关系数进行，由 T 检验所得的 P Value<0.05、差异倍数>1.2、PLSDA 分析计算的 VIP 三个条件同时满足筛选出最终的显著差异代谢物。基于超几何检验进行 KEGG Pathway 的差异富集分析，统计检验的 Pvalue <0.05 的功能条目为差异蛋白显著富集的功能条目。使用软件 GSEA (v4.1.0)和 MSigDB 进行基因集富集分析，以确定一组基因是否在具体的 KEGG pathway 的差异情况，满足此条件|NES|>1, NOM p value<0.05, FDR q-val<0.25 被认为是在两组有显著差异。网络图根据代谢物所在的通路进行绘制。

实验和分析方法（英文版）

实验方法：

1.Metabolites Extraction：

1.1.Metabolites Extraction of Liquid：

100 μL of sample was taken, mixed with 400 μL of extraction solution (MeOH:ACN, 1:1 (v/v)), the extraction solution contain deuterated internal standards, the mixed solution were vortexed for 30 s, sonicated for 10 min in 4 ℃ water bath, and incubatedfor 1 h at -40 ℃ to precipitate proteins. Then the samples were centrifuged at 12000 rpm (RCF=13800(×g),R= 8.6cm) for 15 min at 4 ℃. The supernatant was transferred to afresh glass vial for analysis. The quality control (QC) sample was prepared by mixing an equal aliquot of the supernatant of samples.

1.2 Metabolites Extraction of Solid：

The animal tissue samples (25 mg±1 mg) were taken, mixed with beads and 500 μL of extraction

solution (MeOH:ACN:H2O, 2:2:1 (v/v)). The extraction solution containdeuterated internal standards. The mixed solution were vortexed for 30 s.

The soil samples (100 mg±1 mg) were taken, mixed with beads and 500 μL ofextraction solution (MeOH:ACN:H2O, 2:2:1 (v/v)). The extraction solution containdeuterated internal standards. The mixed solution were vortexed for 30 s.

The plant samples (20 mg±1 mg) were taken and lyophilized, mixed with beads and1000 μL of extraction solution (MeOH:ACN:H2O, 2:2:1 (v/v)). The extraction solution contain deuterated internal standards. The mixed solution were vortexed for 30 s.

Then the mixed samples were homogenized (35 Hz,4 min) and sonicated for 5 min in 4 ℃ water bath, the step repeat for three times.

The samples were incubated for 1 h at -40 ℃ to precipitate proteins. Then thesamples ware centrifuged at 12000 rpm (RCF=13800(×g),R= 8.6cm) for 15 min at 4 ℃. The supernatant was transferred to a fresh glass vial for analysis. The quality control (QC) sample was prepared by mixing an equal aliquot of the supernatant of samples.

1.3 Metabolites Extraction of Cell or Bacteria:

The cell pellets (about 10^7 cells) or bacteria pellets(about 10^7 bacteria) weretaken, mixed with 1000 μL of extraction solution (MeOH:ACN:H2O, 2:2:1 (v/v)), theextraction solution contain deuterated internal standards, the mixed solution were vortexed for 30 s and incubated in liquid nitrogen for 1 min. The samples were thenallowed to thaw at room temperature and vortexed for 30 s. This freeze - thaw cycle was repeated three times. Then the samples were sonicated for 10 min in 4 ℃ water bath, and incubated for 1 h at -40 ℃ to precipitate proteins. The samples were centrifuged at12000 rpm (RCF=13800(×g),R= 8.6cm) for 15 min at 4 ℃. The supernatant was transferredto a fresh glass vial for analysis. The quality control (QC) sample was prepared bymixing an equal aliquot of the supernatant of samples.

2. LC-MS/MS Analysis：

LC-MS/MS analyses were performed using an UHPLC system (Vanquish, Thermo FisherScientific) with a Waters ACQUITY UPLC BEH Amide (2.1 mm × 50 mm, 1.7 μm) coupled to Orbitrap Exploris 120 mass spectrometer (Orbitrap MS, Thermo). The mobile phase consisted of 25 mmol/L ammonium acetate and 25 mmol/L ammonia hydroxide in water（pH = 9.75）(A) and acetonitrile (B). The auto-sampler temperature was 4 ℃, and the injection volume was 2 μL. The Orbitrap Exploris 120 mass spectrometer was used for its ability to acquire MS/MS spectra on information-dependent acquisition (IDA) mode in the controlof the acquisition software (Xcalibur, Thermo). In this mode, the acquisition software continuously evaluates the full scan MS spectrum. The ESI source conditions were set as following: sheath gas flow rate as 50 Arb, Aux gas flow rate as 15 Arb, capillarytemperature 320 ℃, full MS resolution as 60000, MS/MS resolution as 15000, collision energy: SNCE 20/30/40, spray voltage as 3.8 kV (positive) or -3.4 kV (negative), respectively.

4.1.2. Data preprocessing and annotation：

The raw data were converted to the mzXML format using ProteoWizard and processed with an in-house program, which was developed using R and based on The raw data were converted to the mzXML format using ProteoWizard and processed with an in-house program, which was developed using R and based on XCMS, for peak detection, extraction, alignment, and integration. Then an in-house MS2 database was applied in metabolite annotation. The cutoff for annotation was set at 0.3.

2 信息分析

The acquired MS data pretreatments including peak picking, peak grouping, retention time correction, second peak grouping, and annotation of isotopes and adducts was performed using XCMS software. LC−MS raw data files were converted into mzXML format and then processed by the XCMS, CAMERA and metaX toolbox implemented with the R software. Each ion was identified by combining retention time (RT) and m/z data. Intensities of each peaks were recorded and a three dimensional matrix containing arbitrarily assigned peak indices (retention time-m/z pairs), sample names (observations) and ion intensity information (variables) was generated.

The online KEGG, HMDB database was used to annotate the metabolites by matching the exact molecular mass data (m/z) of samples with those from database. If a mass difference between observed and the database value was less than 10 ppm, the metabolite would be annotated and the molecular formula of metabolites would further be identified and validated by the isotopic distribution measurements. We also used a in-house fragment spectrum library of metabolites to validate the metabolite identidification.

Statistical analysis was performed in R (version 4.0.0). The raw protein intensity will be normalized by method "medium", Hierarchical clustering was performed using pheatmap package. Principal component analysis (PCA) was performed using metaX package. The PLSDA analysis is performed by the R package ropls and the VIP values of each variable are calculated.Correlation analysis was performed by Pearson correlation coefficient of cor package .The three conditions of P Value<0.05, difference multiple >1.2 obtained by T test and VIP calculated by PLSDA analysis simultaneously met the screening of the final metabolites with significant differences.. Hypergeometric-based enrichment analysis with KEGG Pathway was performed to annotate protein sequences. individually.The software GSEA (v4.1.0) and MSigDB were used for gene set enrichment analysis to determine whether a set of genes in a specific KEGG pathway in different situations. Meeting this condition |NES|>1, NOM p-val<0.05, FDR q-val<0.25 were considered to be significantly different between the two groups. .The network map is drawn according to the pathway where the metabolite is located.

非靶向代谢组应用文章参考

1. XCMS： Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem. 2006 Feb 1;78(3):779-87. doi: 10.1021/ac051437y. PMID: 16448051.
2. CAMERA： Kuhl C, Tautenhahn R, Böttcher C, Larson TR, Neumann S. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Anal Chem. 2012 Jan 3;84(1):283-9. doi: 10.1021/ac202450g. Epub 2011 Dec 12. PMID: 22111785; PMCID: PMC3658281.
3. metaX： Wen B, Mei Z, Zeng C, Liu S. metaX: a flexible and comprehensive software for

processing metabolomics data. BMC Bioinformatics. 2017 Mar 21;18(1):183. doi: 10.1186/s12859-017-1579-y. PMID: 28327092; PMCID: PMC5361702.
4. PQN：Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. Anal Chem. 2006 Jul 1;78(13):4281-90. doi: 10.1021/ac051632c. PMID: 16808434.
5. Han Q, Li H, Jia M, Wang L, Zhao Y, Zhang M, Zhang Q, Meng Z, Shao J, Yang Y, Zhu L. Age-related changes in metabolites in young donor livers and old recipient sera after liver transplantation from young to old rats. Aging Cell. 2021 Jul;20(7):e13425. doi: 10.1111/acel.13425. Epub 2021 Jun 22. PMID: 34157207; PMCID: PMC8282239.
6. Chen Y, Yang L, Lu Y, Liu N, Ma W, Fan H, Hu Q, Han X, Gan W, Li D. Up-regulation of NMRK2 mediated by TFE3 fusions is the key for energy metabolism adaption of Xp11.2 translocation renal cell carcinoma. Cancer Lett. 2022 Jul 10;538:215689. doi: 10.1016/j.canlet.2022.215689. Epub 2022 Apr 18. PMID: 35447281.

7. Ke M, Kang L, Wang L, Yang S, Wang Y, Liu H, Gu C, Huang H, Yang Y. CAR-T therapy alters synthesis of platelet-activating factor in multiple myeloma patients. J Hematol Oncol. 2021 Jun 9;14(1):90. doi: 10.1186/s13045-021-01101-6. PMID: 34108020; PMCID: PMC8191024.
8. Li R, Huang X, Liang X, Su M, Lai KP, Chen J. Integrated omics analysis reveals the alteration of gut microbe-metabolites in obese adults. Brief Bioinform. 2021 May 20;22(3):bbaa165. doi: 10.1093/bib/bbaa165. PMID: 32770198.