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	规格	货号	品牌
甲醇（Methanol）	4L/瓶	A-456-4	Fisher
乙腈（Acetonitrile）	4L/瓶	955-4	Fisher
甲酸（Formic acid）	50ml/瓶	A17-50	Fisher

表 2. 实验仪器列表

名称	型号	品牌
天平	BSA124S-CW	Sartorius
纯水仪	明澈 D24 UV	Merck Millipore
组织破碎仪	JXFSTPRP-24	净信
色谱柱	T3 column (100mm*2.1mm, 1.8μm)	Waters
超高效液相	Vanquish Flex UHPLC	Thermo
高分辨率质谱	Q-Exactive	Thermo
高分辨率质谱	TripleTOF 6600	Sciex

2.实验流程（关于实验流程，此部分为常规方法，可能会根据您具体样本情况进行调整。如有调整会在 src/method 里附上 word 文档）

2.1 代谢物提取：

液体样本：

采集的样品在冰上解冻，用 50%甲醇缓冲液提取代谢物。简单地说，用 120 μ L 预冷的 50%甲醇提取 20 μ L 样品，旋涡 1 min，室温培养 10 min;提取液在-20℃下保存过夜，4000 g 离心 20 分钟后，将上清转移到新的 96 孔板中。在 LC-MS 分析之前，将样品保存在-80°C。另外，每种提取液各取 10 μ L，制备混合 QC 样品。

固体样本：

采集的样品在冰上解冻，用 50%甲醇缓冲液提取代谢物。简单地说，用 1ml 预冷的 50%甲醇提取 100mg 样品，旋涡 1min，室温培养 10min;提取液在-20℃下保存过夜。4000 g 离心 20 分钟后，将上清液转移到新的 96 孔板中。在 LC-MS 分析之前，将样品保存在-80°C。另外，每种提取液各取 10 μ L，制备混合 QC 样品。

2.2 液相参数描述

所有样品均由 LC-MS 系统按照机器指令采集。首先，所有色谱分离均使用 UltiMate 3000 UPLC 系统(Thermo Fisher Scientific, Bremen, Germany)进行。采用 ACQUITY UPLC T3 色谱柱(100mm*2.1mm, 1.8μm, Waters, Milford, USA)进行反相分离。柱箱保持在 40℃，后加 5mM 乙酸铵和 5mM 乙酸)和溶剂 B(乙腈)。低流速为 0.3 ml/min，流动相为溶剂 A。梯度洗脱条件设定为:0~0.8

min, 2% B;0.8~2.8 min, 2% ~ 70% B;2.8~5.6 min, 70% ~ 90% B;5.6~6.4 min, 90% ~ 100% B;6.4~8.0 min, 100% B;8.0~8.1 min, 100% ~ 2% B;8.1~10 min, 2%B。

2.3 质谱参数描述（具体使用 Q-Exactive 或者 TripleTOF6600，请查看 raw 文件格式。如果是 raw 格式，使用 Q-Exactive；如果是 wiff 及 wiff.scan 格式，使用 TripleTOF6600）

TripleTOF6600：

使用高分辨率串联质谱仪 Q-Exactive (Thermo Scientific)用于检测从柱中洗脱的代谢物。Q-Exactive 在正离子和负离子模式下工作。在 70000 的分辨率下收集代谢离子前体光谱(70-1050 m/z)，达到 AGC 为 3e6。最大注射时间设置为 100 ms。在 DDA 模式下设置了获取数据的 top 3 配置。在 17500 分辨率下采集片段光谱，AGC 为 1e5，最大注入时间为 80 ms.为了评估整个采集过程中 LC-MS 的稳定性，每 10 个样品后采集一个质量控制样品(混合样品)。

Q-Exactive：

使用高分辨率串联质谱仪 TripleTOF 6600 (SCIEX, Framingham, MA, USA)检测从柱中洗脱的代谢物。Q-TOF 在正离子和负离子模式下工作。幕气设置为 30 PSI，离子源气 1 设置为 60 PSI，离子源气 2 设置为 60 PSI，界面加热器温度为 500℃。对于正离子模式，离子喷雾浮动电压分别设置为 5000 V。对于负离子模式，离子喷雾浮动电压分别设置为-4500V。质谱数据采用 IDA 模式获取。TOF 质量范围为 60 ~ 1200 Da。在 150 毫秒内获得调查扫描，如果超过每秒 100 次计数(计数/秒)的阈值并具有 1+电荷状态，则收集多达 12 次产物离子扫描。动态排斥设置为 4 s。在采集过程中，每 20 个样品对质量精度进行校准。

此外，为了评估整个采集过程中 LC-MS 的稳定性，每 10 个样本后采集一个质量控制样本(混合样品)。

3.信息分析流程

3.1 信息分析描述

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。

3.2 XCMS 主要参数

峰提取主要通过开源软件 XCMS 实现。包括峰对齐、峰提取、归一化、去卷积和化合物鉴定等步骤。峰提取和鉴定的主要参数设置如下表所示。

Item	Parameter
method	centWave
minfrac	0.5
snthr	6
ppm	30
peakwidth	5,25
bw2	5
mzwid	0.015
mzdiff	0.01
profStep.OBIWarp	0.1

3.3 metaX 主要参数

对提取的峰进行后处理进行代谢物的一、二级鉴定。

一级鉴定 metaX 参数

Item	Parameter
adduct ion	pos: [M+H]+,[M+Na]+,[M+K]+,[M+NH4]+ neg: [M-H]-,[M+NH4-2H]-,[M+2Cl]2-,[2M-3H]3-
ms1 mass tolerance	10 ppm
database	HMDB、KEGG、Lipidmaps

二级鉴定参数

Item	Parameter
ms1 mass tolerance	0.01 Da
ms2 mass tolerance	0.05 Da
identification score cut off	75%
database	in-house,Massbank,HMDB,Lipidblast

实验和分析方法（英文版）

1 实验流程

1.1 代谢物提取描述

液体样本：

The collected samples were thawed on ice, and metabolite were extracted with 80% methanol Buffer. Briefly, 100 μl of sample was extracted with 400 μl of precooled methanol. The extraction mixture was then stored in 30 min at -20° C. After centrifugation at 20,000 g for 15 min, the supernatants were transferred into new tube to and vacuum dried. The samples were redissolved with 100 μL 80% methanol and stored at -80° C prior to the LC-MS analysis. In addition, pooled QC samples were also prepared by combining 10 μL of each extraction mixture.

固体样本：

The collected samples were thawed on ice, and metabolite were extracted with 80% methanol Buffer. Briefly, 50 mg of sample was extracted with 0.5 ml of precooled 80% methanol. The extraction mixture was then stored in 30 min at -20° C. After centrifugation at 20,000 g for 15 min, the supernatants were transferred into new tube to and vacuum dried. The samples were redissolved with 100 μL 80% methanol and stored at -80° C prior to the LC-MS analysis. In addition, pooled QC samples were also prepared by combining 10 μL of each extraction mixture.

1.2 液相参数描述

All samples were acquired by the LC-MS system followed machine orders. Firstly, all chromatographic separations were performed using an UltiMate 3000 UPLC System (Thermo Fisher Scientific, Bremen, Germany). An ACQUITY UPLC T3 column (100mm*2.1mm, 1.8μm, Waters, Milford, USA) was used for the reversed phase separation. The column oven was maintained at 40° C. The fter, 5mM ammonium acetate and 5mM acetic acid) and solvent B (Acetonitrile). low rate was 0.3 ml/min and the mobile phase consisted of solvent A . Gradient elution conditions were set as follows: 0～0.8 min, 2% B; 0.8～2.8 min, 2% to 70% B; 2.8～5.6 min, 70% to 90% B; 5.6~6.4 min, 90% to 100% B; 6.4~8.0 min, 100% B; 8.0～8.1 min, 100% to 2% B; 8.1～10 min, 2%B.

1.3 质谱参数描述（具体使用 Q-Exactive 或者 TripleTOF6600，请查看 raw 文件格式。如果是 raw 格式，使用 Q-Exactive；如果是 wiff 及 wiff.scan 格式，使用 TripleTOF6600）

TripleTOF6600：

A high-resolution tandem mass spectrometer TripleTOF 6600 (SCIEX, Framingham, MA, USA) was used to detect metabolites eluted form the column. The Q-TOF was operated in both positive and

negative ion modes. The curtain gas was set 30 PSI, Ion source gas1 was set 60 PSI, Ion source gas2 was set 60 PSI, and an interface heater temperature was 500 ℃.For positive ion mode, the Ionspray voltage floating were set at 5000 V, respectively. For negative ion mode, the Ionspray voltage floating were set at -4500V, respectively. The mass spectrometry data were acquired in IDA mode. The TOF mass range was from 60 to 1200 Da. The survey scans were acquired in 150 ms and as many as 12 product ion scans were collected if exceeding a threshold of 100 counts per second (counts/s) and with a 1+ charge-state. Dynamic exclusion was set for 4 s.During the acquisition, the mass accuracy was calibrated every 20 samples. Furthermore, in order to evaluate the stability of the LC-MS during the whole acquisition, a quality control sample (Pool of all samples) was acquired after every 10 samples.

Q-Exactive：

A high-resolution tandem mass spectrometer Q-Exactive (Thermo Scientific) was used to detect metabolites eluted form the column. The Q-Exactive was operated in both positive and negative ion modes. Precursor spectra (70 – 1050 m/z) were collected at 70,000 resolution to hit an AGC target of 3e6. The maximum inject time was set to 100 ms. A top 3 configuration to acquire data was set in DDA mode. Fragment spectra were collected at 17,500 resolution to hit an AGC target of 1e5 with a maximum inject time of 80 ms.In order to evaluate the stability of the LC-MS during the whole acquisition, a quality control sample (Pool of all samples) was acquired after every 10 samples.

2 信息分析流程

2.1 信息分析描述

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.

杭州联川生物技术股份有限公司
杭州经济技术开发区下沙 6 号大街 260 号中自科技园 16 幢 4 层
邮编：310018