HILIC LC-MS positive ion mode non-targeted metabolomics

Hydrophilic interaction liquid chromatography (HILIC) analyses of polar metabolites in the positive ionization mode were conducted at the Broad Institute of MIT and Harvard.

Sample preparation

Plasma samples: Samples were stored at -80°C until extraction and were thawed and maintained on wet ice throughout processing steps. Plasma samples (10 µL) were extracted using 90 µL of 74.9/24.9/0.2 v/v/v acetonitrile/methanol/formic acid containing valine-d8 (0.2 ng/µL, Sigma-Aldrich) and phenylalanine-d8 (0.2 ng/µL, Cambridge Isotope Laboratories) internal standards. Samples were centrifuged (10 minutes, 9,000 x g, 4°C) and supernatants were transferred to autosampler vials with deactivated glass inserts (Waters).

Tissue samples: Samples were stored at -80°C until extraction and were thawed and maintained on wet ice throughout processing steps. Powdered tissue samples (10 mg) were homogenized at 4°C in 300 µL of 10/67.4/22.4/0.018 v/v/v/v water/acetonitrile/methanol/formic acid containing valine-d8 (0.2 ng/µL, Sigma-Aldrich) and phenylalanine-d8 (0.2 ng/µL, Cambridge Isotope Laboratories) internal standards using a TissueLyser II (QIAGEN) bead mill set to two 2 min intervals at 20 Hz. Samples were centrifuged (10 minutes, 9,000 x g, 4°C) and supernatants were transferred to autosampler vials with deactivated glass inserts (Waters).

LC-MS analysis

Data were acquired using an LC-MS system comprised of a Nexera X2 UHPLC (Shimadzu) coupled to a Q-Exactive hybrid quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific). Metabolite extracts were injected onto a 150 x 2 mm, 3 µm Atlantis HILIC column (Waters). The column was eluted using the gradient program in Table 2. MS analyses were carried out using electrospray ionization in the positive ion mode using full scan analysis over 70-800 *m/z* at 70,000 resolution and 3 Hz data acquisition rate. Other MS settings were: sheath gas 40, auxiliary gas 10, sweep gas 2, spray voltage 3.5 kV, capillary temperature 350°C, S-lens RF 40, heater temperature 300°C, microscans 1, automatic gain control target 1e6, and maximum ion time 250 ms.

Table 2. HILIC-pos gradient program

Time (minutes)	%A	%B	Flow (mL/min)
0	5	95	0.25
0.5	5	95	0.25
10.5	60	40	0.25
15	60	40	0.25
17	5	95	0.25
17.5	5	95	0.25
18	5	95	0.4
30.5	5	95	0.4
31.5	5	95	0.25

32	5	95	0.25

Data processing and quality control

Feature detection and alignment: Raw data from each tissue type were processed independently using TraceFinder software (version 3.3; Thermo Fisher Scientific) for targeted peak integration and manual review of a subset of identified metabolites and Progenesis QI (Nonlinear Dynamics, Waters) for peak detection and integration of both metabolites of known identity, confirmed using reference standards, and unknowns. Data for each sample type were acquired in a single batch and unknowns were aligned using Progenesis QI.

Data cleaning and degeneracy removal: Unknown features detected and integrated using Progenesis QI were de-isotoped by the software and tacked by measured mass-to-charge ratios (m/z) and retention times (RT). In order to filter redundant ion features, Spearman correlation coefficients were calculated among all LC-MS features that eluted within a sliding retention time window of 0.025 minutes. Cliques of highly correlated, co-eluting features (Spearman correlation coefficients >0.8) were then identified. When possible, mass differences among features within each clique were used to determine ion adduct types. If an [M+H]+ ion could be identified, that feature was retained; otherwise the most abundant feature was retained. All other redundant features were flagged and removed.

Drift correction and quality control: Data quality was assured by i) initially confirming LC-MS system performance by analyzing a mixture of >140 well-characterized synthetic reference compounds as well as repeated analyses of extracts from human pooled plasma (BioIVT); ii) daily evaluation of internal standard signals to ensure that each sample injected properly and to monitor MS sensitivity; and iii) analysis of four pairs of pooled extract samples per sample type that were inserted in the analysis queue at regular intervals. One sample from each pair was used to correct for instrument drift using “nearest neighbor” scaling while the second reference sample served as a passive QC for determination of the analytical coefficient of variation of every identified metabolite and unknown.

Compound identification: Except for lipids, metabolite identities were confirmed using authentic reference standards to match measured MS and RT data in the method. Lipids were identified using representative standards and identified by RT patterns and high resolution and accurate mass MS data.