Reverse phase and ion pairing LC-MS non-targeted metabolomics

Reverse-phase and ion pairing profiling of polar metabolites was conducted at the University of Michigan.

Sample preparation

Plasma samples: Samples were stored at -80°C until extraction and were thawed and maintained on wet ice throughout processing steps. Plasma (50 µl aliquot) was extracted by adding 200 µL of extraction solvent (1:1:1 v:v methanol:acetonitrile:acetone containing internal standards listed in Table 3; chemical standards were purchased from Sigma Aldrich or Cambridge Isotope Labs). Samples were vortexed 10 seconds, incubated in ice for 10 minutes, then centrifuged for 10 minutes at 15,000 rcf at 4°C to pellet precipitated proteins. 150 µL supernatant was transferred to a glass autosampler vial with a low-volume insert and brought to dryness using a nitrogen blower at ambient temperature. Dried samples were reconstituted in 37.5 µL water:methanol (8:2 v:v) for LC-MS analysis. A QC sample was generated by pooling residual supernatant from multiple samples, then drying and reconstituting as above.

Table 3. Internal standard concentrations in extraction solvent

Internal standard compound	**Concentration in extraction solvent for plasma (µM, except as noted)**	**Concentration in extraction solvent for tissues (µM, except as noted)**
$^{13}C_3$ lactic acid	50	10
$^{13}C_5$ alpha-ketoglutaric acid	0.5	0.5
$^{13}C_6$ citric acid	5	0.5
$^{13}C_4$ succinic acid	0.5	0.5
$^{13}C_4$ malic acid	0.5	0.5
$^{13}C_6$ fructose 6-phosphate	-	1
$^{13}C_6$ fructose 1,6-bisphosphate	-	0.5
$^{13}C_{10}{}^{15}N_5$ adenosine monophosphate	-	0.5
$^{13}C_{10}{}^{15}N_5$ adenosine triphosphate	-	1
U-^{13}C amino acid mix (Sigma 426199)	10 µg/ml	5 µg/ml
$^{13}C_5$ glutamine	25	5
$^{13}C_6$ cystine	5	1
$^{15}N_2$ asparagine	5	1
$^{15}N_2$ tryptophan	5	1
$^{13}C_6$ glucose	250	2
D_9 L-carnitine	0.38	0.38
D_4 thymine	1	1
^{15}N anthranilic acid	1	1
Gibberelic acid	1	1
Epibrassinolide	1	1

D_3 acetylcarnitine (Cambridge NSK-B)	0.095	0.095
D_3 propionylcarnitine (Cambridge NSK-B)	0.019	0.019
D_3 butyrylcarnitine (Cambridge NSK-B)	0.019	0.019
D_9 isovalerylcarnitine (Cambridge NSK-B)	0.019	0.019
D_3 octanoylcarnitine (Cambridge NSK-B)	0.019	0.019
D_9 myristoylcarnitine (Cambridge NSK-B)	0.019	0.019
D_3 palmitoylcarnitine (Cambridge NSK-B)	0.038	0.038

Tissue samples: Frozen tissue samples were rapidly weighed into pre-tared, pre-chilled Eppendorf tubes and tissue mass was recorded to the nearest 0.1 mg. Extraction solvent was 1:1:1:1 methanol:acetonitrile:acetone:water containing internal standards. To extract samples, chilled extraction solvent was added to a tissue sample at the ratio of 1 ml solvent to 50 mg wet tissue mass. Immediately following solvent addition, the sample was homogenized using a Branson 450 probe sonicator set to output level 4, 40% duty cycle, for 30 seconds. Tubes were subsequently mixed several times by inversion and then incubated on ice for 10 minutes. Samples were centrifuged at 15,000 rcf for 10 minutes. 300 µL of supernatant was transferred to two autosampler vials with flat-bottom inserts, dried using the nitrogen blower, and stored at -80C until the day of analysis. Samples were reconstituted in 60 µL of 8:2 water:methanol and submitted for LC-MS analysis. A QC sample was generated by pooling residual supernatant from multiple samples, then drying and reconstituting as above.

LC-MS analysis

Non-targeted reverse phase LC-MS: Samples were analyzed on an Agilent 1290 Infinity II / 6545 qTOF MS system with a JetStream electrospray ionization (ESI) source (Agilent Technologies, Santa Clara, California) using a Waters Acquity HSS T3 column, 1.8 µm 2.1 x 100 mm equipped with a matched Vanguard precolumn (Waters Corporation). Mobile phase A was 100% water with 0.1% formic acid and mobile phase B was 100% methanol with 0.025% formic acid. The gradient was as follows: Linear ramp from 0% to 100% B from 0-10 minutes, hold 100% B until 17 minutes, linear return to 0% B from 17 to 17.1 minutes, hold 0% B until 20 minutes. The flow rate was 0.45 ml/min, the column temperature was 55°C, and the injection volume was 5 µL. All solvents and mobile phase additives were LC-MS grade and purchased from Sigma-Aldrich. Each sample was analyzed twice, once in positive and once in negative ion mode MS, scan rate 2 spectra/sec, mass range 50-1200 *m/z*. Source parameters were: drying gas temperature 350°C, drying gas flow rate 10 L/min, nebulizer pressure 30 psig, sheath gas temperature 350°C and flow 11 L/minute, capillary voltage 3500 V, internal reference mass correction enabled. A QC sample run was performed at minimum every tenth injection.

Non-targeted ion pairing LC-MS: Samples were analyzed on an identically-configured LC-MS system using an Agilent Zorbax Extend C18 1.8 µm RRHD column, 2.1 x 150 mm ID, equipped with a matched guard column. Mobile phase A was 97% water, 3% methanol. Mobile phase B was 100% methanol. Both mobile phases contained 15 mM tributylamine and 10 mM acetic acid. Mobile phase C was 100% acetonitrile. Elution was carried out using a linear gradient followed by a multi-step column wash including automated (valve-controlled) backflushing (see Table 4). Column temperature was 35°C and the injection volume was 5 µL. MS acquisition was

performed in negative ion mode, scan rate 2 spectra/sec, mass range 50-1200 *m/z*. Source parameters were: drying gas temperature 250°C, drying gas flow rate 13 L/min, nebulizer pressure 35 psig, sheath gas temp 325°C and flow 12 L/min, capillary voltage 3500V, internal reference mass correction enabled. A QC sample run was performed at minimum every tenth injection.

Table 4. IPC-MS gradient program

Time (minutes)	%A	%B	%C	Flow (ml/min)	Flow direction
0	100	0	0	0.25	Normal
2	100	0	0	0.25	Normal
12	1	99	0	0.25	Normal
18	1	99	0	0.25	Normal
18.05	5	0	95	0.25	Backflush
21	5	0	95	0.25	Backflush
21.50	5	0	95	0.80	Backflush
23	5	0	95	0.80	Backflush
23.2	5	0	95	0.60	Backflush
24.00	100	0	0	0.40	Backflush
27.99	100	0	0	0.40	Backflush
28.0	100	0	0	0.40	Normal
29.9	100	0	0	0.40	Normal
30	100	0	0	0.25	Normal

Iterative Data Dependent MS/MS data acquisition (iDDA): To aid in compound identification, iterative MS/MS data was acquired for both reverse phase and ion pairing methods using the pooled sample material. Eight repeated LC-MS/MS runs of the QC sample were performed at three different collision energies (10, 20, and 40) with iterative acquisition enabled. The software excluded precursor ions from MS/MS acquisition within 0.5 minute of their MS/MS acquisition time in prior runs, resulting in deeper MS/MS coverage of lower-abundance precursor ions[37,38].

Data analysis

Feature detection and alignment: Data analysis was performed using a hybrid targeted/untargeted approach. Targeted compound detection and relative quantitation was performed by automatic integration followed by manual review and correction using Profinder v8.0 (Agilent Technologies, Santa Clara, CA.) Non-targeted feature detection was performed using custom scripts that automate operation of the “find by molecular feature” workflow of the Agilent Masshunter Qualitative Analysis (v7) software package. Feature alignment and recursive feature detection were performed using Agilent Mass Profiler Pro (v8.0) and Masshunter Qualitative Analysis (“find by formula” workflow), yielding an aligned table including *m/z*, RT, and peak areas for all features.

Data Cleaning and Degeneracy Removal: A combined feature set was generated by merging untargeted features and named metabolites into a single feature list. Features missing from over 50% of all samples in a batch or over 30% of QC samples were removed prior to subsequent

normalization steps. Next, the combined feature set underwent data reduction using Binner[39]. Briefly, Binner first performs RT-based binning, followed by clustering of features by Pearson's correlation coefficient, and then assigns annotations for isotopes, adducts or in-source fragments by searching for known mass differences between highly correlated features.

Normalization and Quality Control: Data were normalized using a Systematic Error Removal Using Random Forest (SERRF) approach[40], which helps correct for drift in peak intensity over the batch using data from the QC sample runs. When necessary to correct for residual drift, peak area normalization to closest-matching internal standard was also applied to selected compounds. Both SERRF correction and internal standard normalization were implemented in R. Parameters were set to minimize batch effects and other observable drift, as visualized using principal component analysis score plots of the full dataset. Normalization performance was also validated by examining relative standard deviation values for additional QC samples not included in the drift correction calculations. Quality control reports containing these data were generated for all datasets and uploaded to the MoTrPAC data repository along with raw and processed data.

Compound identification: Metabolites from the targeted analysis workflow were identified with high confidence (MSI level 1)[41] by matching retention time (+/- 0.1 minute), mass (+/- 10 ppm) and isotope profile (peak height and spacing) to authentic standards. MS/MS data corresponding to unidentified features of interest from the untargeted analysis were searched against a spectral library (NIST 2020 MS/MS spectral database or other public spectral databases) to generate putative identifications (MSI level 2) or compound-class level annotations (MSI level 3) as described previously[38].