1. LC-MS method descriptions

a. C8-pos: Reversed-phase C8 chromatography/positive ion mode MS detection to measure lipids (Broad Institute). Analyses of polar and non-polar plasma lipids were conducted using an LC-MS system comprised of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.) coupled to an Exactive Plus orbitrap mass spectrometer (Thermo Fisher Scientific). Plasma samples (10 µL) were extracted for lipid analyses using 190 µL of isopropanol containing 1,2-didodecanoyl-sn-glycero-3-phosphocholine (Avanti Polar Lipids) as an internal standard. After centrifugation, supernatants were injected directly onto a 100 x 2.1 mm, 1.7 µm ACQUITY BEH C8 column (Waters). The column was eluted isocratically with 80% mobile phase A (95:5:0.1 vol/vol/vol 10mM ammonium acetate/methanol/formic acid) for 1 minute followed by a linear gradient to 80% mobile-phase B (99.9:0.1 vol/vol methanol/formic acid) over 2 minutes, a linear gradient to 100% mobile phase B over 7 minutes, then 3 minutes at 100% mobile-phase B. MS analyses were carried out using electrospray ionization in the positive ion mode using full scan analysis over 200–1100 m/z at 70,000 resolution and 3 Hz data acquisition rate. Other MS settings were: sheath gas 50, in source CID 5 eV, sweep gas 5, spray voltage 3 kV, capillary temperature 300°C, S-lens RF 60, heater temperature 300°C, microscans 1, automatic gain control target 1e6, and maximum ion time 100 ms. Raw data were processed using TraceFinder software (Thermo Fisher Scientific) for targeted peak integration and manual review of a subset of identified lipids and using Progenesis QI (Nonlinear Dynamics) for peak detection and integration of both lipids of known identify and unknowns. Lipid identities were determined based on comparison to reference plasma extracts and are denoted by total number of carbons in the lipid acyl chain(s) and total number of double bonds in the lipid acyl chain(s).

b. C18-neg: Reversed-phase C18 chromatography/negative ion mode MS detection to measure free fatty acids, bile acids, and metabolites of intermediate polarity (Broad Institute). Analyses of free fatty acids and bile acids were conducted using an LC-MS system comprised of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.) coupled to a Q Exactive hybrid quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific). Plasma samples (30 µL) were extracted using 90 uL of methanol containing 15R-15-methyl-PGA_2, 15R-15-methyl-PGF_{2alpha}, 15S-15-methyl-PGD_2, 15S-15-methyl-PGE_1, and 15S-15-methyl-PGE_2 (Cayman Chemical Co.) internal standards and centrifuged (10 min, 9,000 x g, 4°C). The samples were injected onto a 150 x 2 mm ACQUITY BEH C18 column (Waters). The column was eluted isocratically at a flow rate of 400 µL/min with 60% mobile phase A (0.1% formic acid in water) for 4 minutes followed by a linear gradient to 100% mobile phase B (acetonitrile with 0.1% acetic acid) over 8 minutes. MS analyses were carried out in the negative ion mode using electrospray ionization, full scan MS acquisition over 200-550 m/z, and a resolution setting of 70,000. Other MS settings were: sheath gas 45, sweep gas 5, spray voltage -3.5 kV, capillary temperature 320°C, S-lens RF 60, heater temperature 300°C, microscans 1, automatic gain control target 1e6, and maximum ion time 250 ms. Raw data were processed using TraceFinder software (Thermo Fisher Scientific) for targeted peak integration and manual review of a subset of identified metabolites and using Progenesis QI (Nonlinear Dynamics) for peak detection and integration of both metabolites of known identify and unknowns. Metabolite identities were confirmed using authentic reference standards.

c. HILIC-pos: Hydrophilic interaction liquid chromatography/positive ion mode MS detection to measure polar metabolites (Broad Institute). HILIC analyses of water soluble metabolites in the positive ionization mode were conducted using an LC-MS system comprised of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.) coupled to a Q Exactive hybrid quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific). Plasma samples (10 µL) were prepared via protein precipitation with the addition of nine volumes of 74.9:24.9:0.2 v/v/v

acetonitrile/methanol/formic acid containing stable isotope-labeled internal standards (valine-d8, Sigma-Aldrich; St. Louis, MO; and phenylalanine-d8, Cambridge Isotope Laboratories; Andover, MA). The samples were centrifuged (10 min, 9,000 x g, 4°C), and the supernatants were injected directly onto a 150 x 2 mm, 3 µm Atlantis HILIC column (Waters). The column was eluted isocratically at a flow rate of 250 µL/min with 5% mobile phase A (10 mM ammonium formate and 0.1% formic acid in water) for 0.5 minute followed by a linear gradient to 40% mobile phase B (acetonitrile with 0.1% formic acid) over 10 minutes. 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, 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. Raw data were processed using TraceFinder software (Thermo Fisher Scientific) for targeted peak integration and manual review of a subset of identified metabolites and using Progenesis QI (Nonlinear Dynamics) for peak detection and integration of both metabolites of known identify and unknowns. Metabolite identities were confirmed using authentic reference standards.

d. Amide-neg: Targeted negative ion mode analysis of central metabolites (BIDMC). Central metabolites including sugars, sugar phosphates, organic acids, purine, and pyrimidines, were extracted from 30 µL of plasma using acetonitrile and methanol and separated using a 100 x 2.1 mm XBridge Amide column (Waters). A high sensitivity Agilent 6490 QQQ MS (Agilent) was used to profile metabolites in the negative ion mode via multiple reaction monitoring (MRM) scanning. MRM parameters for approximately 200 metabolites were previousluy optimized by infusing authentic reference standards. Raw data are processed using MassHunter Quantitative Analysis Software (Agilent).

2. Nontargeted data processing

Raw LC-MS data were acquired to the data acquisition computer interfaced to each LC-MS system and then stored on a robust and redundant file storage system (Isilon Systems) accessed via the Broad's internal network. For data processing, the platform is equipped with >10 powerful workstations configured with multi-core XEON processors, >32 GB of RAM and 2 TB of fast storage (RAID 0 arrays of four drives or NVME ssd). Targeted data processing of known metabolites was achieved using TraceFinder software (Thermo Fisher Scientific). Identities of >600 plasma metabolites have been confirmed using authentic reference standards (**MSI Level 1 ID**) and mixtures of reference standards and reference samples were included in each analysis queue to confirm IDs in every dataset. High resolution, nontargeted data were processed using Progenesis QI software (Nonlinear Dynamics) to detect peaks, perform chromatographic retention time alignment, and integrate peak areas. Metabolites of confirmed identity were then annotated in the dataset and unknowns are "tagged" using their measured mass to charge ratio (m/z) and retention time (RT). A significant challenge for large-scale, nontargeted metabolomics studies is accurate "alignment" on unknowns among batches of samples acquired over time. This challenge arises because of the large number of nontargeted features detected in every dataset and the occurrence of minor deviations in measured m/z and RT for each peak as a function of differences in instrument calibration and LC column performance over time. The Broad lab has developed an innovative feature alignment algorithm that overcomes this challenge. The software tool is deployed as a web app and uses a unique approach to detect landmark features and non-parametric retention time scaling to accurately match unknowns between datasets. This workflow is illustrated in **Figure A1**. The output from this workflow is a table of concatenated data from each of the methods, expressed as individual

samples in columns and metabolite abundances in rows. The format is easily shared and is amenable to analyses using standard software (e.g. R, Excel, Matlab, etc.).

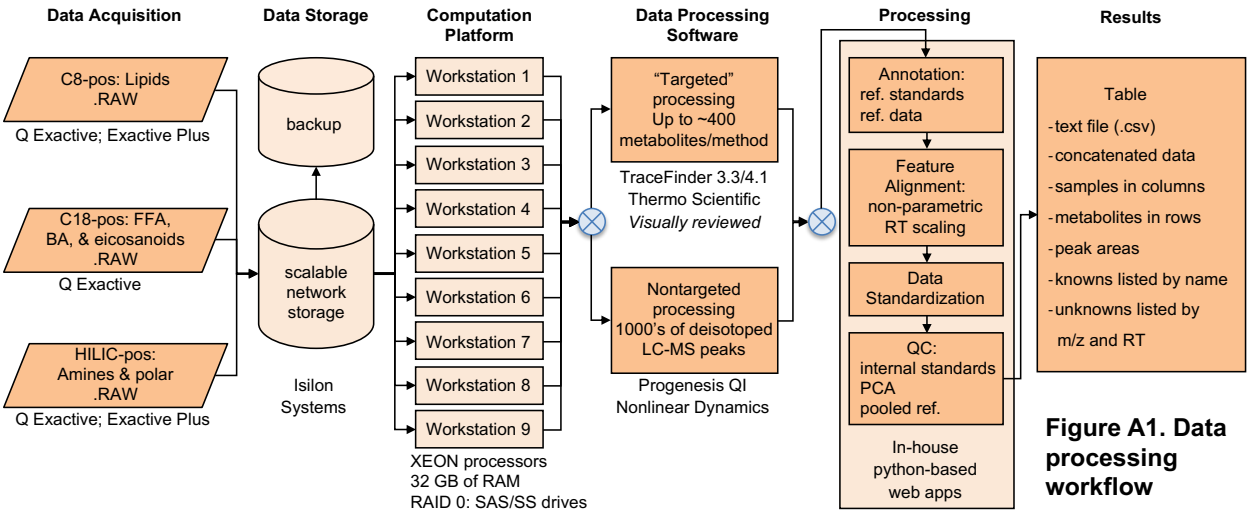

Figure A1. Data processing workflow