Methods S2: Pharmacometabonomics analysis for cohort 1 and 2

The profiling of the plasma EMCs was completed at the University of California, Davis Genome Center’s West Coast Metabolomics Center. Global pharmacometabonomics analysis of their primary metabolism assay was carried out using a gas chromatography time-of-flight mass spectrometry (GC-TOF MS). The primary metabolism assay, which covers carbohydrates and sugar phosphates, amino acids, hydroxyl acids, free fatty acids, purines, pyrimidines, aromatics, exposome-derived chemicals, was used. It provided the normalized relative intensities are reported for up to 200 identified EMCs in addition to 200-300 compounds of unknown chemical structure.

The data were acquired using the following chromatographic parameters. Additional chromatographic details can be found in Fiehn et al.1

Column: The column was a Restek corporation Rtx-5Sil MS (30 m length x 0.25 mm internal diameter with 0.25 μm film made of 95% dimethyl/5%diphenylpolysiloxane) with helium mobile phase with a flow-rate of 1 mL/min. The injection volume was 0.5 μL with the injection being 25 splitless time into a multi-baffled glass liner Injection temperature: 50°C ramped to 250°C by 12°C/s.

For the oven temperature program, 50°C for 1 min, then ramped at 20°C min-1 to 330°C, held constant for 5 min. The analytical GC column was protected by a 10 m long empty guard column which was cut by 20 cm intervals whenever the reference mixture quality control (QC) samples indicate problems caused by column contaminations. The WCMC validated that at this sequence of column cuts, no detrimental effects are detected with respect to peak shapes, absolute or relative metabolite retention times or reproducibility of quantifications. This chromatography method yields excellent retention and separation of primary metabolite classes (amino acids, hydroxyl acids, carbohydrates, sugar acids, sterols, aromatics, nucleosides, amines and miscellaneous compounds) with narrow peak widths of 2–3 s and very good within-series retention time reproducibility of better than 0.2 s absolute deviation of retention times. The WCMC uses automatic liner exchanges after each set of 10 injections which they could show to reduce sample carryover for highly lipophilic compounds such as free fatty acids. The mass spectrometry parameters that were used were: a Leco Pegasus IV mass spectrometer is used with unit mass resolution at 17 spectra s-1 from 80-500 Da at -70 eV ionization energy and 1800 V detector voltage with a 230°C transfer line and a 250°C ion source.

For the data processing, the raw data files are preprocessed directly after data acquisition and stored as ChromaTOF-specific *.peg files, as generic *.txt result files and additionally as generic ANDI MS *.cdf files. ChromaTOF vs. 2.32 was used for data preprocessing without smoothing, 3 s peak width, baseline subtraction just above the noise level, and automatic mass spectral deconvolution and peak detection at signal/noise levels of 5:1 throughout the chromatogram. Apex masses were reported for use in the BinBase algorithm.2 Result *.txt files are exported to a data server with absolute spectra intensities and further processed by a filtering algorithm implemented in the metabolomics BinBase database.

References for Methods S2:
1. Fiehn, O. Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling. Curr Protoc Mol Biol. 2016;114:30 34 31-30 34 32.
2. Lai, Z., Tsugawa, H., Wohlgemuth, G., et al. Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics. Nature methods. 2018;15:53-56.