

Online Only Supplemental Methods

Table of Contents

A. Untargeted Metabolomics Measurements	2
B. Metabolite Annotation	3

A. Untargeted Metabolomics Measurements

Untargeted metabolomics was measured in plasma samples from the two-hour time point of the oral glucose tolerance test via liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) as described previously (1). This analysis was run in batches of 70 study samples and ten pooled QA/QC samples. The analysis for the SOLAR and CHS cohorts was performed consecutively within seven days. To maximize the detection of endogenous metabolites, we expanded established methods (2) to allow analysis using a dual column and dual polarity approach that included analyses with both reverse phase (RP) and hydrophilic interaction (HILIC) chromatography. Untargeted analysis was accomplished using a Vanquish Duo liquid chromatography (LC) system (Thermo Fisher Scientific, Rockford, IL, USA) equipped with dual pumps and columns with independent flow paths interfaced to a Q-Exactive HF-X Orbitrap MS system (Thermo Fisher Scientific, Rockford, IL, USA). The dual LC system was configured to enable parallel analytical separation and flushing using columns with the same stationary phase and mobile phases optimized for positive or negative ionization. To facilitate measurement by all four analytical configurations, all study samples were initially analyzed using RP analysis, after which the system was switched to HILIC.

Before analysis, samples were thawed at 4°C and plasma (40 µL for RP and 30 µL for HILIC) was extracted with ice-cold acetonitrile (80 µL for RP and 90 µL for HILIC). Treated samples were vortexed for 10 seconds, equilibrated on ice for 30 minutes, and then centrifuged for 15 minutes at 18,000×g and 4 °C. The supernatant (40 µL for RP and 30 µL for HILIC) was added to 250 µL LC vials containing water (80 µL for RP) or 1:1 (v/v) water/ acetonitrile (90 µL for HILIC analysis) and placed in a refrigerated autosampler. Samples were analyzed with mobile phases optimized for positive or negative ionization. For both positive and negative modes, RP analyte separation was accomplished by C₁₈ (TARGA C₁₈ 5µm 50x2.1mm, Higgins Analytical, Inc, Mountain View, CA, USA). Mobile phases for RP included water and acetonitrile containing 0.1% formic acid for positive mode and 10mM ammonium acetate and 95/5 (v/v) acetonitrile/water (A) for negative mode. For HILIC, positive ESI analysis was completed using a SeQuant ZIC-HILIC column (3.5µm, 200A 4.6x50mm; MilliporeSigma, Burlington, MA, USA), and mobile phase including 0.1% formic acid in water and acetonitrile. For negative mode, a Waters XBridge Amide column (3.5µm 3.0x50mm, Waters Corporation, Milford, MA, USA) was used with mobile phases consisting of 10mM ammonium acetate in water adjusted to pH 9.60 with ammonium hydroxide and 95/5 (v/v) acetonitrile. The total runtime for each analysis was 7.5 minutes. Mass spectral data was collected over the scan range 85-1275 at 120,000 (FWHM) resolution. Spray voltages were maintained at 3.5 and 4.0 kV for positive and negative modes, respectively. Sheath and auxiliary gas temperatures were 300°C and 250°C, respectively, while sheath and auxiliary gas flow rates were set to 45 and 25 (arbitrary units). The RF funnel level was set at 35 to minimize analyte fragmentation at the source. In addition to full scan data collection, a subset of samples was selected for data-dependent MSMS, which collected MSMS spectra for the top 20 most abundant peaks at MS2 resolution of 15,000 using normalized collision energies of 20, 40, and 60.

Following analysis of all samples from the SOLAR and CHS cohorts, mass spectral peaks for metabolites were extracted with apLCMS and xMSanalyzer (2). Extraction was performed across all

study samples concurrently and was performed separately for each of the four LC-HRMS modes. After extracting LC-MS features, inter- and intra-batch variation was corrected using a random forest signal correction algorithm based on quality control samples run in tandem with study samples (3). As part of the random forest signal correction algorithm, features not detected in more than 25% of samples were removed from further analysis. Features were also removed from further analysis if the coefficient of variability in all quality control samples post-correction was more than 30%. After LC-MS data processing, the total number of features included in data analysis was 23,166, including 3,711 features from the C18 negative mode, 5,069 features from the C18 positive mode, 7,442 features from the HILIC negative mode, and 6,944 features from the HILIC positive mode.

B. Metabolite Annotation

Metabolite annotation was performed as follows. First, tentative metabolite annotation for the 595 features included in the feature selection was performed using version 2 of the MS peaks to paths module from MetaboAnalyst v5.0 (4; 5). Following feature selection, the identities of the three selected metabolites were confirmed by comparison to a database of authentic standards analyzed on the same instrument using identical instrumental methods based upon mass error and retention difference of 10ppm and 20 seconds. The standard database was developed using both solvent and plasma-spiked samples, and the m/z , retention time, and MSMS were confirmed by visual inspection. Comparison to the three annotated metabolites of interest in this study supported the initial identification of caprylic acid and taurocholic acid based on a level 1 annotation defined by Schymanski et al. (2008) (6). However, the third metabolite ($m/z = 214.0261$; retention time 35 seconds; Mode: HILIC negative), which MetaboAnalyst initially annotated as hippuric acid, was found to be inaccurately identified.

To address this issue, we re-annotated this feature using the following steps. We first identified two features that were significantly correlated ($r^2 > 0.8$) with the original feature across both cohorts and that each had retention times within 1 second of the original feature. These three features had m/z 's of 133.0660 and 213.0228, respectively. Through analysis of MS1 spectra, we identified the peak at 213.0228 as the M-H feature, which was annotated to Allylphenol sulfate, and 133.0660 as a potential in-source fragment. Accordingly, the peak at 214.0261 was designated as the M(C13)-H peak, with a theoretically predicted mass intensity of 10%. This process allowed us to confirm the chemical formula of the feature as $C_9H_{10}O_4S$. The sole metabolite in the HMDB database that matched this chemical formula was Allylphenol sulfate, and the presence of the in-source fragment was consistent with the database MSMS spectra, allowing for level two annotation as defined by Schymanski et al. (2008) (6).

References

1. Goodrich JA, Walker DI, He J, Lin X, Baumert BO, Hu X, Alderete TL, Chen Z, Valvi D, Fuentes ZC, Rock S, Wang H, Berhane K, Gilliland FD, Goran MI, Jones DP, Conti DV, Chatzi L. Metabolic Signatures of Youth Exposure to Mixtures of Per- and Polyfluoroalkyl Substances: A Multi-Cohort Study. *Environ Health Perspect* 2023;131:27005
2. Liu KH, Nellis M, Uppal K, Ma C, Tran V, Liang Y, Walker DI, Jones DP. Reference Standardization for Quantification and Harmonization of Large-Scale Metabolomics. *Analytical Chemistry* 2020;92:8836-8844
3. Luan H, Ji F, Chen Y, Cai Z. statTarget: A streamlined tool for signal drift correction and interpretations of quantitative mass spectrometry-based omics data. *Analytica Chimica Acta* 2018;1036:66-72
4. Pang Z, Zhou G, Ewald J, Chang L, Hacariz O, Basu N, Xia J. Using MetaboAnalyst 5.0 for LC-HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. *Nat Protoc* 2022;17:1735-1761
5. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, Gauthier C, Jacques PE, Li S, Xia J. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res* 2021;49:W388-W396
6. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol* 2014;48:2097-2098