

Supplemental Methods

To further characterize the potentially novel deoxyceramide observed in plasma, the photochemical Paternò-Büchi reaction with acetone was used to create double bond cleavage products similar to the approach described by Murphy et. al³⁸. Briefly, plasma fractions or reference compounds were dried and resuspended in acetone solution and exposed to UV light using a mercury lamp. This reaction produces ozone cleavage products at each double bond location which, by tracking the formation of aldehyde products, can be used to map the original location of the double bond.

Experimental conditions for the UV treatments were determined using the Cer(m18:0/24:1) reference compound, which eluted later in the chromatogram but produced identical MS/MS product ion spectra compared to the deoxyceramide detected in plasma. The reference compound was suspended in acetone/water solution (75/25, v/v), placed approximately 5 mm from a mercury lamp, exposed to UV emissions over several time intervals ranging from 2 to 20 minutes. LC-MS was used to monitor consumption of the starting material and formation of the expected cleavage aldehyde product at carbon 4 of Cer(m18:1/24:0) with a calculated m/z 454.4255. Essentially complete conversion of the starting material occurred within 15 minutes of UV exposure.

C8 chromatography coupled with fraction collecting was then applied toward isolating the deoxyceramide from human plasma. The fractions were dried under nitrogen using a TurboVap LV (Biotage) and resuspended in acetone/water solution (75/25, v/v). Prior to UV exposure, LC-MS was used to monitor the relative abundance of the deoxyceramide in each fraction (Supplemental Figure 3A). The solutions were then exposed to UV light for 15 minutes as described above and LC-MS was used to monitor for double bond cleavage products. Similar to the Cer(m18:1/24:0) reference material, a major product of m/z 454.4253 was observed following UV exposure, indicating the presence of the 4,5 double bond in the parent compound similar to Cer(m18:1/24:0). The signal intensity of the product also correlated with the signal intensity of the parent compound in each fraction, indicating that the cleavage product was derived from the parent compound (Supplemental Figure 3B). Of note, the cleavage product for the plasma deoxyceramide peak appeared as a multiplet at an earlier retention time compared to the linear alkyl chain-containing cleavage product of Cer(m18:1/24:0), similar to the RT difference observed between the parent compounds. The cleavage products contain the 24:0 alkyl chain and therefore it is possible that branching of the C24:0 fatty acid in the plasma deoxyceramide contributes to the retention time difference.

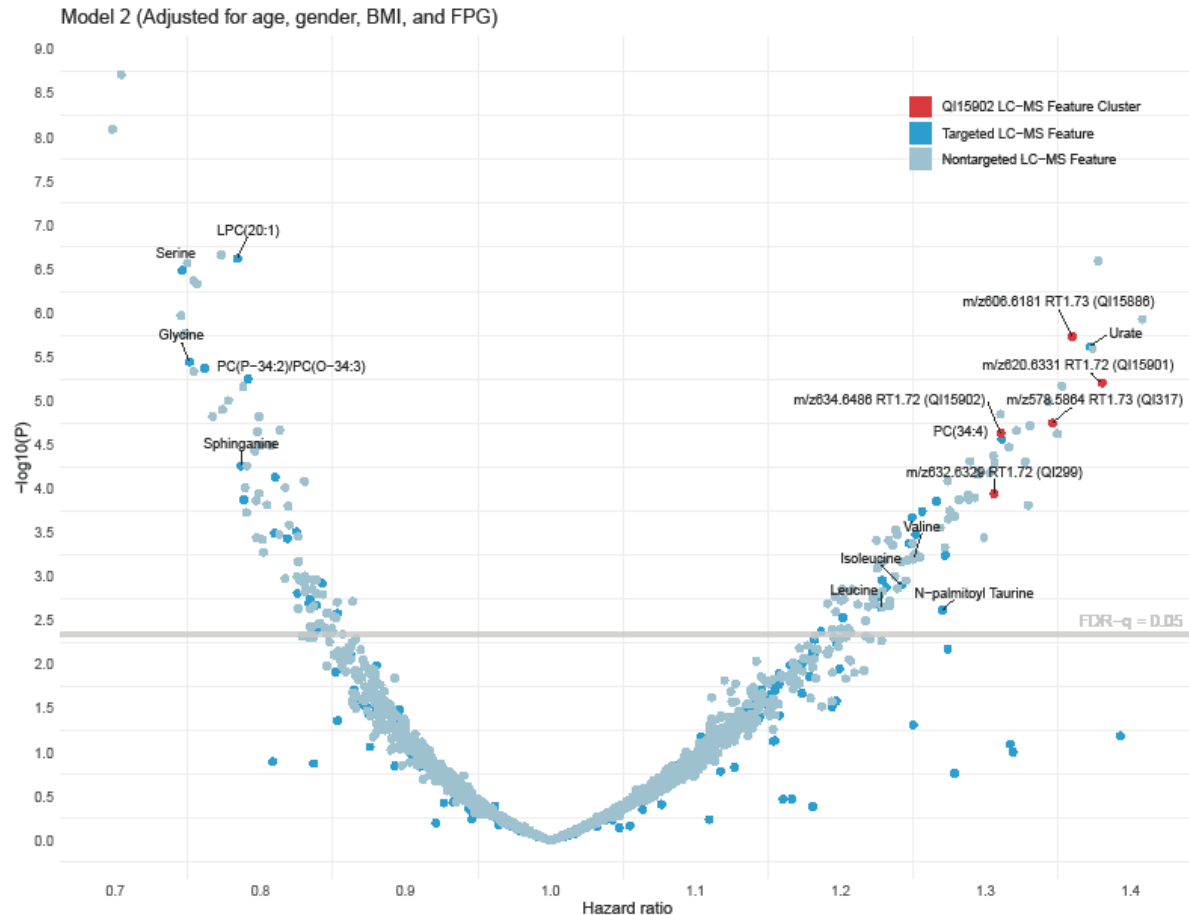
Inspection of the LC-MS data for theoretical cleavage products formed by other potential double bond locations in either the sphingoid base or fatty acyl moieties of the parent structure did not reveal any signals (not shown). Last, the MS/MS spectrum of the plasma ceramide cleavage product had an m/z 104.0706 ion consistent with the region of the molecule containing the aldehyde derivative (Supplemental Figure 3D). A complete list of the relevant MS/MS spectra generated from these studies is also listed on the next page which have been uploaded to a publicly available data repository.

Relevant MS/MS Spectra Available on Publicly Accessible MS Spectra Database

1. MS/MS of reference compound N-(tetracosanoyl)-1-deoxysphing-4-enine (Cer(m18:1/24:0), PubChem CID 73242199)
2. MS/MS of reference compound N-(15Z-tetracosenoyl)-1-deoxysphinganine (Cer(m18:0/24:1), PubChem CID 73242204)
3. MS/MS of human plasma unknown QI15902 (m/z 634.6486, RT 1.72 min)
4. MS/MS of Paterno-Büchi reaction double bond cleavage product of reference compound Cer(m18:1/24:0)
5. MS/MS of Paterno-Büchi reaction double bond cleavage product of reference compound Cer(m18:0/24:1)
6. MS/MS of Paterno-Büchi reaction double bond cleavage product of unknown QI15902 fractionated from human plasma

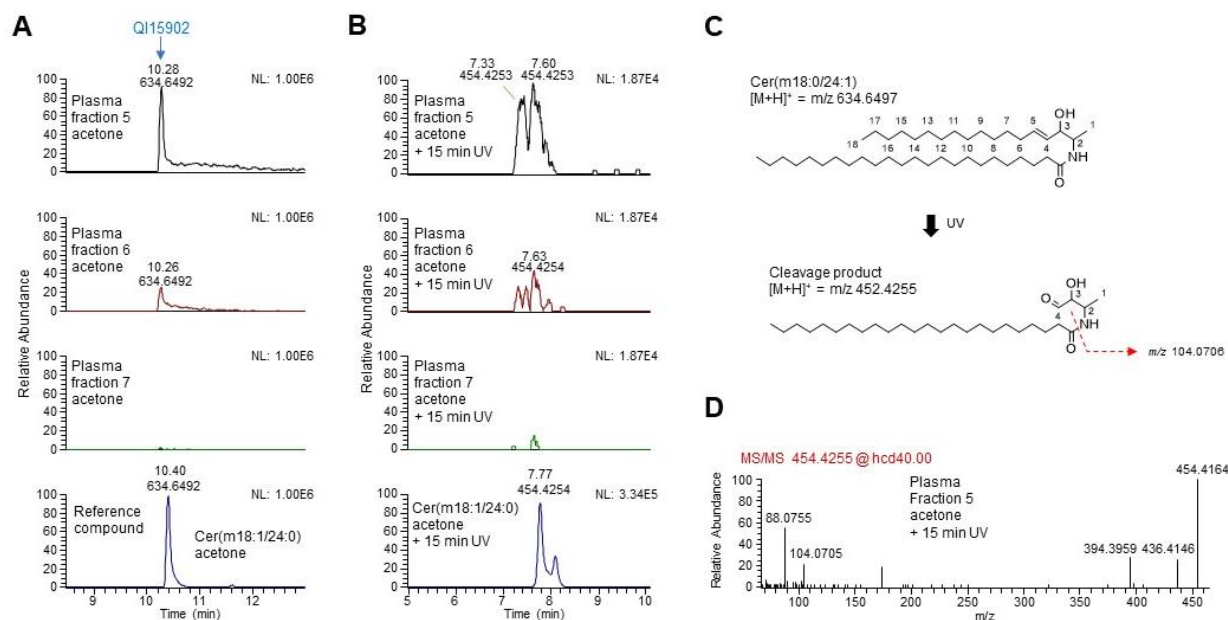
These MS/MS spectra have been uploaded to the MassIVE database on the University of California, San Diego Center for Computational Mass Spectrometry (CCMS) website ([doi:10.25345/C5V97ZW46](https://doi.org/10.25345/C5V97ZW46)) under the MassIVE dataset MSV000090113.

Supplemental Figure 1. Incident diabetes association of the LC-MS peaks associated with nontargeted compound QI15902 in JHS



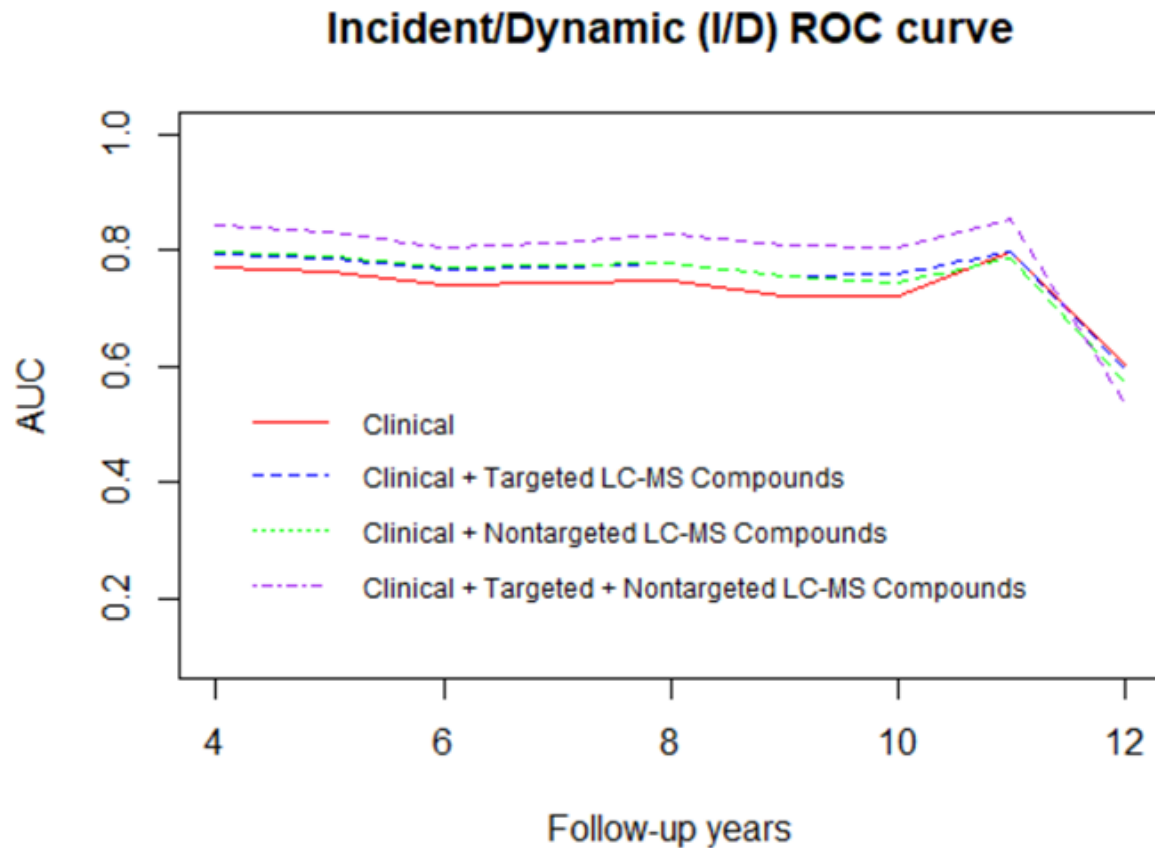
Volcano plot showing *all* targeted and nontargeted LC-MS peak (not restricting to only primary features) associations with incident diabetes in JHS model 2 (Cox proportional hazard model adjusted for age, gender, BMI, and PFG). Diabetes hazard ratios over a mean 10.2 years follow up for every 1 SD increase in transformed and normalized LC-MS peak is plotted on the x-axis. The horizontal grey line denotes an FDR $q < 0.05$. Dark blue dots are targeted LC-MS peaks, light blue dots are nontargeted LC-MS peaks. Red dots represent the LC-MS peaks that are member of the compound QI15902 cluster. Select targeted LC-MS compounds are named as reference. LPC: lipophosphatidylcholine. PC: phosphatidylcholine.

Supplemental Figure 2. Extracted ion chromatograms (EIC) of plasma lipid fractions and UV cleavage products.



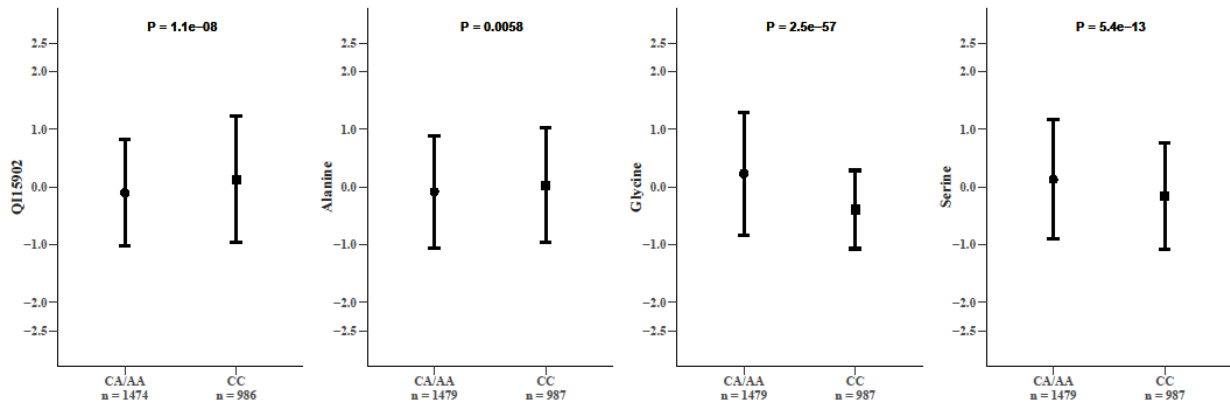
(A) EIC at m/z 634.6497 of deoxyceramide in three consecutive chromatographic fractions of human plasma (top panels) and of neat Cer(m18:1/24:0) reference compound suspended in acetone. (B) EIC at m/z 454.4255 of expected products from cleavage of the double bond at carbon 4 of the deoxyceramide following UV treatment in acetone. Top three panels are products from human plasma fractions and the bottom panel is the product from the neat Cer(m18:1/24:0) reference material. (C) Cleavage product expected for the double bond position at carbon 4 of Cer(m18:1/24:0). (D) MS/MS product ion spectra for plasma UV cleavage product detected in fraction 5.

Supplemental Figure 3. Incident receiver operating characteristic curve (ROC) for the different incident diabetes prediction models in JHS



Clinical risk factors included age, gender, BMI, systolic blood pressure, HDL cholesterol, triglycerides, FPG, and family history of diabetes and were used in all models. Elastic net regularization was used to select targeted, nontargeted, and targeted plus nontargeted LC-MS primary features to be included as predictors in the respective models.

Supplemental Figure 4. The CPS1 rs1047891 SNP association with select metabolite levels in JHS.



Homozygote and heterozygote carriers of the C>A missense substitution within the *CPS1* gene demonstrate lower levels of Q115902 and alanine levels and higher levels of glycine and serine. The mean and SD of natural log-transformed and standardized relative metabolite concentrations are depicted with *P* values generated from an age- and sex-adjusted regression model