Supplementary Information

Title: The ratio of circulatory levels of sphingolipids to steroids predicts asthma exacerbations

Authors:
Yulu Chen[1⊥], Pei Zhang[2,3⊥], Mengna Huang[1], Priyadarishini Kachroo[1], Antonio Checa[2], Qingwen Chen[1], Kevin Mendez[1], Meryl Stav[1], Nicole Prince[1], Sofina Begum[1], Andrea Aparicio[1], Tao Guo[1], Rinku Sharma[1], Su H. Chu[1], Rachel S. Kelly[1], Julian Hecker[1], Ayobami Akenroye[4], Amber Dahlin[1], Scott T. Weiss[1], Michael McGeachie[1], Craig E. Wheelock[2,3,5*], Jessica A. Lasky-Su[1*]

[1] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[2] Unit of Integrative Metabolomics, Institute of Environmental Medicine, Karolinska Institute, 171 77, Stockholm, Sweden
[3] Gunma University Initiative for Advanced Research (GIAR), Gunma University, Maebashi, Japan
[4] Division of Allergy and Clinical Immunology, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[5] Department of Respiratory Medicine and Allergy, Karolinska University Hospital, 171 76, Stockholm, Sweden

[⊥] These authors contributed equally to this manuscript
* These authors contributed equally to this manuscript

Keywords: asthma exacerbations, targeted metabolomics, metabolite ratios, electronic medical records, biobank

Supplemental Methods – untargeted metabolomics and lipidomics methods applied in the MGBB-KAS cohort. These methods have not been previously published and are therefore provided in detail here.

1. HILIC-Q/TOF-MS (positive ionization mode)

1.1 Preparation of internal standard stock solution

The following chemicals were used as technical internal standards: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES; Dojindo Chemicals, #GB10-10), piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES; Dojindo Chemicals, #GB15-10), N-Cyclohexyl-2-aminoethanesulfonic acid (CHES; Dojindo Chemicals, #GB07-10), Fluorocytosine (COMBI-BLOCKS, #OR-1262), Pyrantel (Wako, #167-18391), and L-Histidine-15N3 (Taiyo Nippon Sanso, #B03-0005). Stock solutions were prepared as follows: HEPES at 1 mM in water, PIPES at 1 mM in water, CHES at 1 mM in water, 5-Fluorocytosine at 1 mM in methanol (MeOH), Pyrantel pamoate at 0.5 mM in MeOH, and L-Histidine-$^{15}N_3$ at 1 mM in water.

1.2 Sample preparation

Serum samples were thawed overnight at 4 °C, vortexed, and 300 µL was transferred to each well of a 2 mL 96-deepwell plate placed on a metal block on ice. An additional aliquot of each sample was pooled to create the study quality control (QC) sample, which was then transferred to the dedicated wells of the deepwell plate. To avoid multiple freeze−thaw cycles, the deepwell plate content was directly aliquoted in aliquots of 20 µL into 0.2 mL 96-well plates using a Bravo automated liquid-handling platform (Agilent Technologies, Inc.) equipped with a cooling unit set at 4 °C. The plates were sealed with a pierceable seal (4titude; Wotton, UK) for 3 s at 170 °C by using a plate sealer (BioRad PX-1; CA, USA). Aliquot plates were stored at −80 °C until use.

One aliquot plate (20 µL) was thawed on ice, and protein precipitation was performed by adding 120 µL of ice-cold methanol containing a mixture of internal standards (HEPES PIPES, CHES, Fluorocytosine, Pyrantel, and L-Histidine-$^{15}N_3$ at final concentrations of 0.025 µM, 0.5 µM, 0.1 µM, 1 µM, 1 µM, and 2 µM, respectively). Samples were vortexed for 1 min, mixed for 10 min (Eppendorf ThermoMixer® C), centrifuged (Kubota 7000, Japan) at 4390 *g* for 30 min at 4 °C. An aliquot of 40 µL of the supernatant was transferred to a new plate and extracts were evaporated to dryness. The dried samples were reconstituted in 80 µL acetonitrile:water (80:20, v/v). After centrifuge at 4390 *g* for 15 min at 4 °C, 50 µL of the supernatant was transferred to a new plate for LC-MS measurement. QC samples were extracted and injected every 10 samples across the whole cohort for monitoring the instrument performance.

1.3 Instrumental conditions

The instrumental system used was a 1290 Infinity II ultrahigh performance liquid chromatography (UHPLC) system coupled to a 6550 iFunnel quadrupole-time-of-flight (Q-TOF)

mass spectrometer equipped with a dual AJS electrospray ionization source (Agilent Technologies).

Metabolites were separated using a SeQuant ZIC-HILIC column (Merck, Darmstadt, Germany; 100 × 2.1 mm, 3.5 μm) coupled with a guard column (2.1 × 2 mm, 3.5 μm). The mobile phase consisted of solvent A (0.1% formic acid in water) and solvent B (acetonitrile with 0.1% formic acid). The elution gradient was as follows: an isocratic step at 95% B for 1.5 min, a gradual increase from 95% to 40% B over 10.5 min, held at 40% B for 2 min, then decreased to 25% B at 14.2 min, maintained at 25% B for 2.8 min, followed by a return to initial conditions over 1 min, and finally re-equilibration at initial conditions for 7 min. The flow rate was 0.3 mL/min, the injection volume was 1 μL, and the column oven was maintained at 25 °C.

The Agilent 6550 Q-TOF-MS system was calibrated and tuned following the manufacturer's recommended protocols. High-purity nitrogen (>99.999%) was used as both the sheath gas and drying gas, with flow rates of 8 L/min and 15 L/min, respectively. The drying and sheath gas temperatures were set at 250 °C, and the nebulizer pressure was maintained at 35 psig. Positive ionization mode was employed with a voltage of +3000 V. The fragmentor voltage was set to 180 V. Data acquisition covered a mass range of 40–1200 m/z in the positive all ion fragmentation mode, including 3 sequential experiments at alternating collision energies: one full scan at 0 eV, followed by one MS/MS scan at 10 eV, and then one MS/MS scan at 30 eV.

2. RPLC-Q/TOF-MS (lipidomics in positive and negative ionization modes)

2.1 Preparation of internal standard stock solution

The following chemicals were used as internal standards: 1-nonadecanoyl-2-hydroxy-sn-glycero-3-phosphocholine (LysoPC 19:0; Avanti, #855776P-25MG-A-022), N-lauroyl-D-erythro-sphingosylphosphorylcholine (SM d18:1/12:0; Avanti, #860583P-5MG-A-017), N-heptadecanoyl-D-erythro-sphingosine (Cer d18:1/17:0; Avanti, #860517P-5MG-A-013), 1,2-diheptadecanoyl-sn-glycero-3-phosphoethanolamine (PE 17:0/17:0; Avanti, #830756P-25MG-B-020), 1,2,3-tripentadecanoyl-sn-glycerol (TG 15:0/15:0/15:0; Sigma, #MKBV2973V-100MG), 1-pentadecanoyl-2-oleyol(d7)-sn-glycerol (DG 15:0-18:1-d7; Avanti, #791647C-1MG), cholesteryl-d7 palmitate (16:0 cholesteryl-d7 ester; Avanti, #700149P-10MG-C-010) and 16,16,16-trideuteriohexadecanoic acid (Palmitic acid-16,16,16-d3; SRL, #BX-224).

Stock solutions were prepared as follows: LysoPC 19:0 at 1 mM in chloroform, SM d18:1/12:0 at 1 mM in methanol, Cer d18:1/17:0 at 1 mM in methanol, PE 17:0/17:0 at 1 mM in butanol-methanol (1:1, v/v), DG 15:0-18:1-d7 at 0.85 mM in chloroform, TG 15:0/15:0/15:0 at 1 mM in chloroform-methanol (1:1, v/v), 16:0 cholesteryl-d7 ester at 1 mM in chloroform-methanol (1:1, v/v), and Palmitic acid-16,16,16-d3 at 1 mM in chloroform. An internal standard mixture was created by pooling and diluting the stock solutions with methanol, resulting in the following final concentrations: LysoPC 19:0 at 1.7 μM, SM d18:1/12:0 at 0.85 μM, Cer d18:1/17:0 at 0.85 μM, PE 17:0/17:0 at 3.4 μM, DG 15:0-18:1-d7 at 11.33 μM, TG 15:0/15:0/15:0 at 0.57 μM, 16:0 cholesteryl-d7 ester at 17 μM, and Palmitic acid-16,16,16-d3 at 28.33 μM.

2.2 Sample preparation

60 µL of the internal standard mixture was added to 20 µL of serum. The mixture was vortexed for 2 min at 4°C and 2000 rpm using an Eppendorf ThermoMixer® C. Then, 200 µL of methyl tert-butyl ether (MTBE) was added, vortexed for 5 min at 4°C and 2000 rpm. Next, 60 µL of water was added, and the sample was vortexed for 2 min at 4°C and 2000 rpm. The samples were then centrifuged at 14,000 *g* for 10 min at 4°C. A 100 µL aliquot of the supernatant was transferred to a new Eppendorf tube and dried under nitrogen gas flow at room temperature for 10 min. For LC-MS analysis, the dried pellet was reconstituted in 100 µL of isopropanol (IPA) and mixed for 5 min at 4°C and 2000 rpm. The samples were centrifuged again at 14,000 *g* for 10 min at 4°C, and 60 µL of the supernatant was transferred to an LC vial with insert. QC samples were extracted and injected every 10 samples across the whole cohort for monitoring the instrument performance.

2.3 Instrumental conditions

The analytical system comprised a 1290 Infinity II ultrahigh-performance liquid chromatography coupled to a 6550 iFunnel quadrupole-time-of-flight (Q-TOF) mass spectrometer equipped with a dual AJS electrospray ionization source (Agilent Technologies).

Lipids were separated on a Zorbax Eclipse Plus C18 column (100 × 2.1 mm, 1.8 µm) with an attached guard column. The mobile phase consisted of solvent A (acetonitrile-water 60:40, v/v, 10 mM ammonium formate) and solvent B (isopropanol:acetonitrile 90:10, v/v, 10 mM ammonium formate). The elution gradient was as follows: from 20% to 60% B over 3 min, then from 60% to 85% B over 7 min, maintained at 85% B for 5 min, rapidly increased to 97% B in 0.1 min, held at 97% B for 2 min, decreased back to 20% B at 17.1 min, and held for an additional 2 min. The flow rate was 0.4 mL/min, with an injection volume of 1 µL. The column oven was maintained at 50 °C.

The Agilent 6550 Q-TOF-MS system was calibrated and tuned according to the manufacturer's recommended protocols. Nitrogen (>99.999% purity) was used as both the sheath gas and drying gas, with flow rates of 12 L/min and 15 L/min, respectively. Gas temperatures were set at 250 °C, and the nebulizer pressure was maintained at 40 psig. The instrument voltage was +3000 V in positive mode or -3000 V in negative mode, with the fragmentor voltage fixed at 360 V. Data-dependent acquisition (DDA) mode was used, covering a mass range of 40 to 1200 m/z. Collision energies were set at 20 V for positive mode and 25 V for negative mode.

3. Data quality check and preprocessing

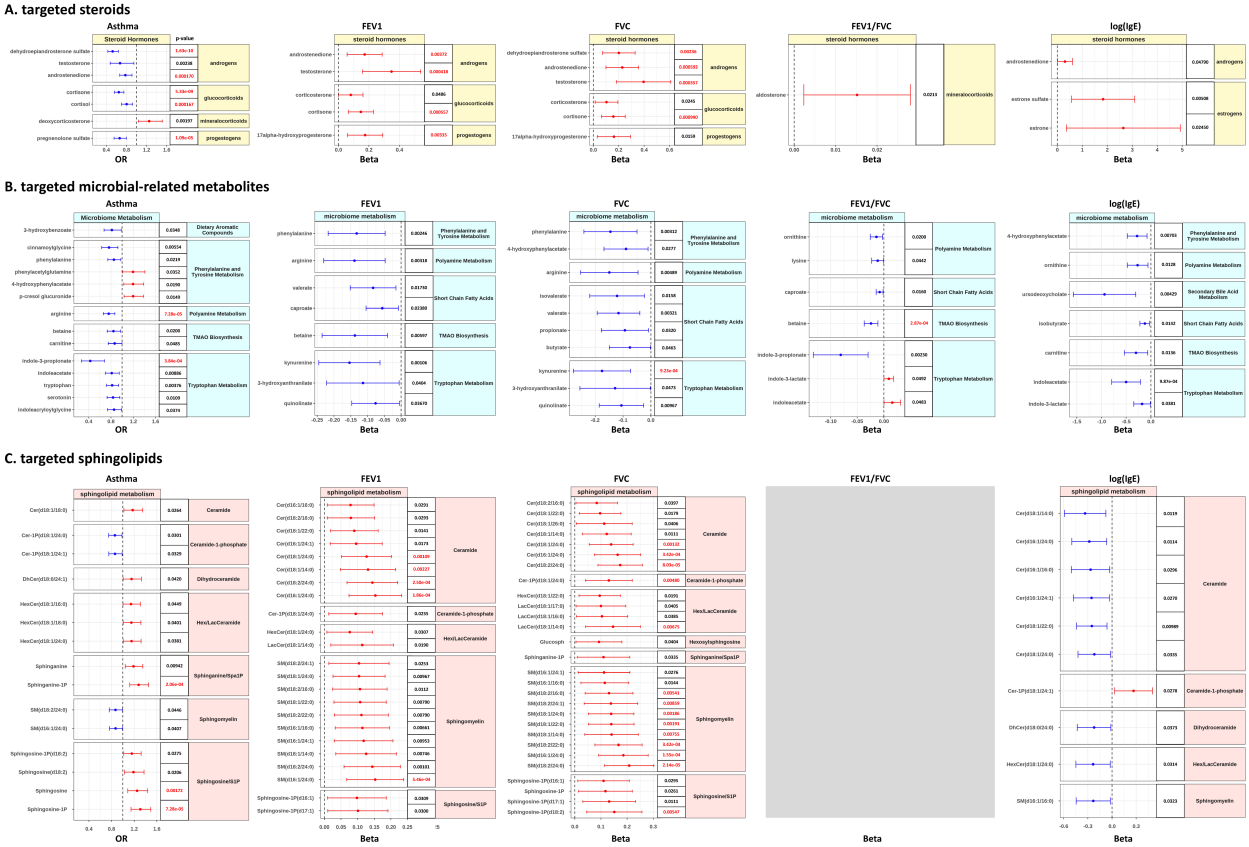
After completing each batch, data quality was evaluated based on the performance of internal standards in both QC samples and study samples. Coefficients of variation were calculated, with acceptance criteria set at less than 15% for QC samples and under 25% for all samples. Retention time (RT) correction algorithms were applied to raw data files to correct for any RT drift across batches. Data deconvolution was performed using the CorrDec method in MS-DIAL (Version 4.20), and peak intensity drift was corrected using QC samples within the preprocessed

dataset. To ensure data integrity, features missing in 75% or more of samples were excluded. Missing values were imputed as half the minimum value for each feature across all samples. PCA plots were examined to evaluate data variability, and the interquartile range (IQR) and skewness of each feature were calculated before and after transformations. All features were subsequently log-10 transformed and Pareto-scaled, with IQR and skewness recalculated post-transformation to assess normalization effectiveness.

4. Metabolite/lipid annotation

For the HILIC-Q/TOF-MS data, an in-house MS2 spectral library containing experimental MS2 spectra and RTs for 391 compounds derived from standards was used to annotate detected metabolites based on three criteria: (i) accurate mass (AM) match (tolerance: 0.01 Da), (ii) RT match (tolerance: 1 min), and (iii) MS2 spectrum similarity (greater than 80%). For features not included in the in-house library, annotation was performed by comparison with reference MS2 spectra available in MS-DIAL (Version 4.20). Lipid annotations were assigned based on retention time, m/z, and MS/MS using MS-DIAL (Version 4.20).

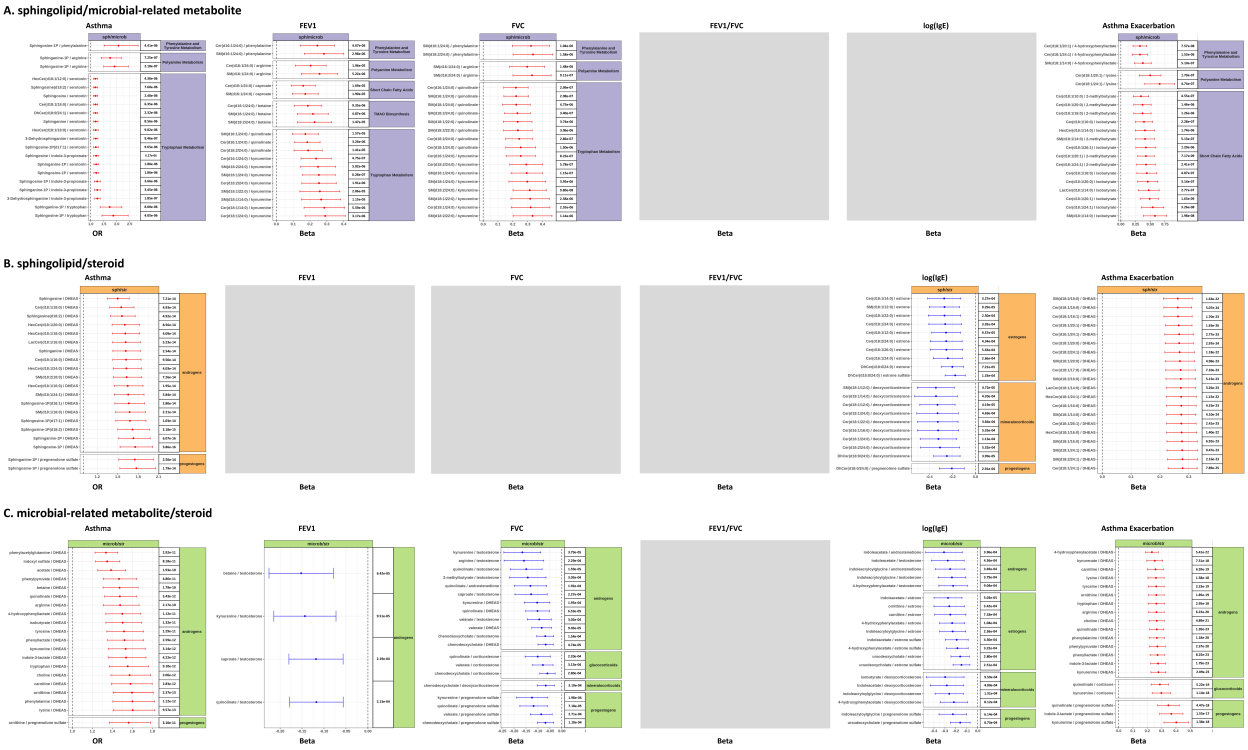
Supplemental Figure 1.1 Forest plot of significant targeted metabolites association results in MGBB-KAS. The significance levels of the metabolite ratio regression analyses were FDR ≤ 0.05. The black metabolites represent those with p-value ≤ 0.05 in the regression model while the red metabolites indicate those with FDR ≤ 0.05 in the regression model. **A)** steroid hormones panel, **B)** microbial-derived metabolites panel, and **C)** sphingolipid panel.



Gray blank areas indicate that no metabolites met the significance threshold for the corresponding asthma phenotype. For asthma (binary outcome), points denote adjusted odds ratios (ORs) from logistic regression; the centre is the point estimate and error bars are 95% confidence intervals. The x-axis is on a log(OR) scale. Two-sided Wald tests of β=0 are reported (test statistic z = β/SE(β)) with exact P values. For other asthma-related continuous traits, points denote adjusted regression coefficients (β) from linear regression; the centre is the point estimate and error bars are 95% confidence intervals based on model-based standard errors. Two-sided t-tests of β=0 are reported (test statistic t(df)), with exact P values.

Abbreviations: OR odds ratio; FEV_1, forced expiratory volume in 1 second; FVC, forced vital capacity; FEV_1/FVC, the ratio of FEV_1 to FVC.

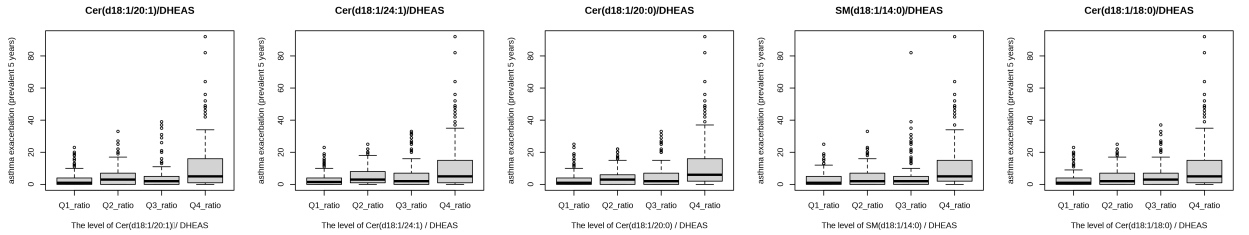
Supplemental Figure 1.2 Forest plot of top 20 significant metabolite ratios association results in MGBB-KAS. The significance levels of the metabolite ratio regression analyses were FDR ≤ 0.05. **A**) ratio of sphingolipids/microbial-derived metabolites, **B**) ratio of sphingolipids/steroids, and **C**) ratio of microbial-derived metabolites/steroids.



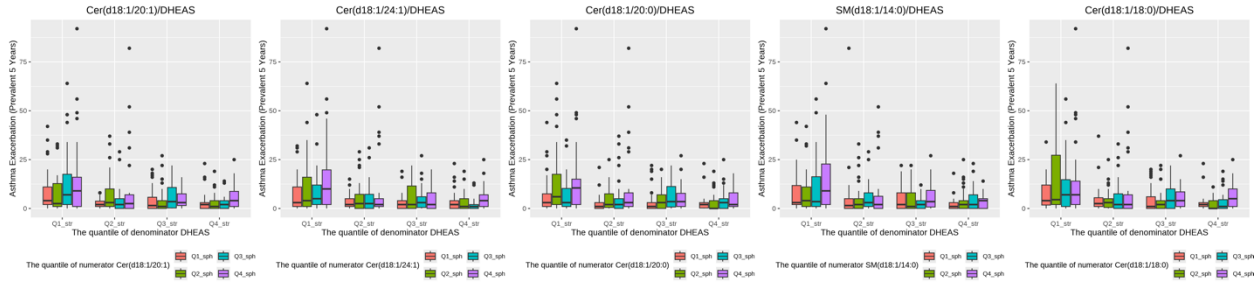
Gray blank areas indicate that no metabolites met the significance threshold for the corresponding asthma phenotype. For asthma (binary outcome), points denote adjusted odds ratios (ORs) from logistic regression; the centre is the point estimate and error bars are 95% confidence intervals. The x-axis is on a log(OR) scale. Two-sided Wald tests of β=0 are reported (test statistic z = β/SE(β)) with exact P values. For other asthma-related continuous traits, points denote adjusted regression coefficients (β) from linear regression; the centre is the point estimate and error bars are 95% confidence intervals based on model-based standard errors. Two-sided t-tests of β=0 are reported (test statistic t(df)), with exact P values.
Abbreviations: OR odds ratio; FEV_1, forced expiratory volume in 1 second; FVC, forced vital capacity; FEV_1/FVC, the ratio of FEV_1 to FVC.

Supplemental Figure 2. The associations with asthma exacerbations are driven by the metabolite ratios, not the denominator or numerator: **A)** Boxplots of asthma exacerbation by top 5 metabolite ratios; **B)** Boxplots of asthma exacerbation by combined quartiles of numerators and denominators from the top 5 metabolite ratios.

A)



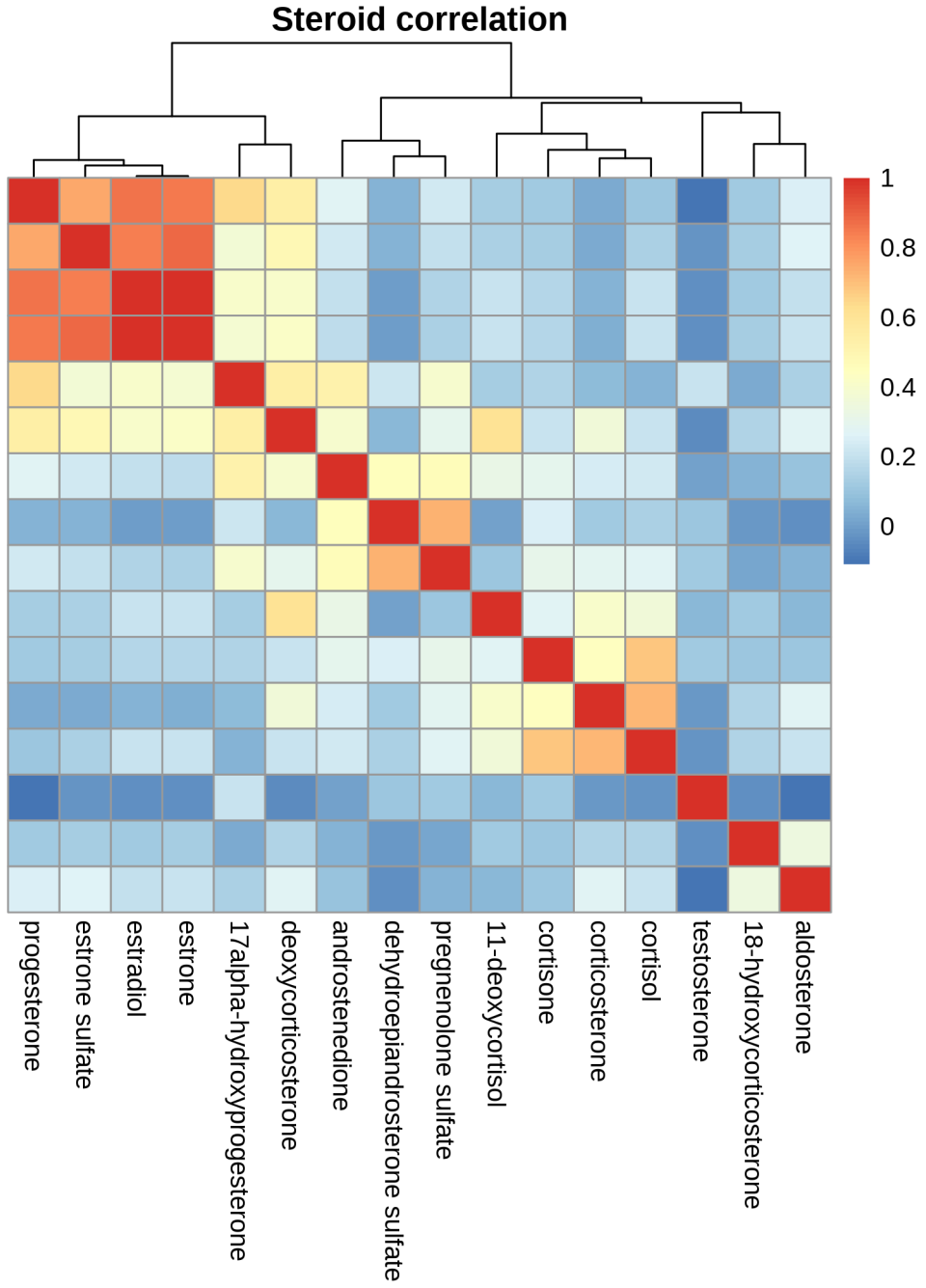
B)



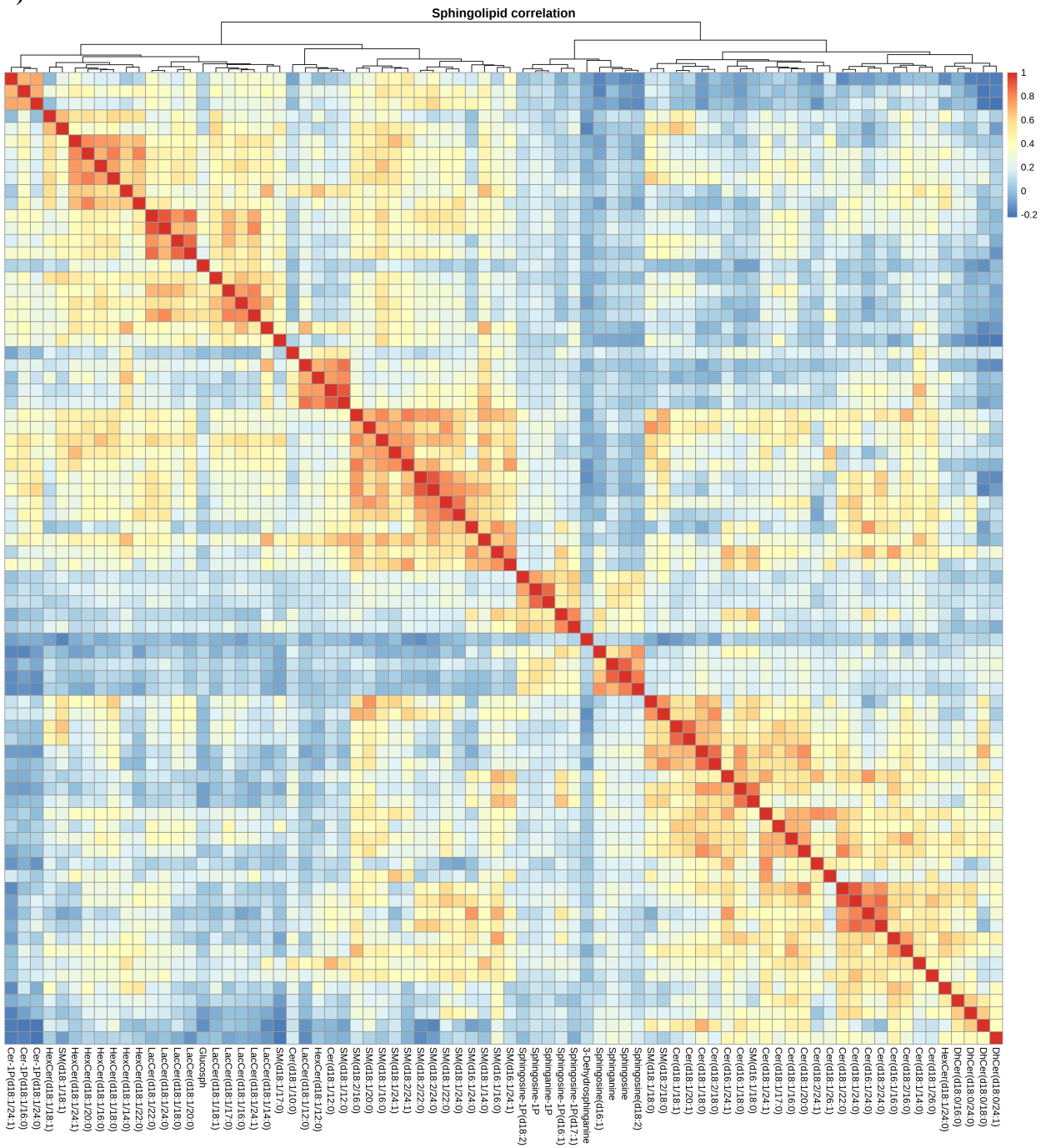
A) For each ratio, participants were grouped into quartiles of the ratio (Q1–Q4; Q1 = lowest, Q4 = highest). Overlaid points show individual observations. Boxes display the median (centre line) and IQR (box); whiskers extend to 1.5×IQR and points beyond whiskers denote outliers. The y-axis is the 5-year asthma exacerbation burden derived from EMR.
B) Two-dimensional stratification by denominator and numerator quantiles. Along the x-axis, groups correspond to the quartiles of the denominator steroid (DHEAS; Q1_str–Q4_str), and box colours encode the quartiles of the numerator sphingolipid (Q1_sph–Q4_sph). Boxplot elements reporting are as in (A).
Abbreviations: DHEAS, dehydroepiandrosterone sulfate; Cer, ceramide; SM, sphingomyelin; IQR, interquartile range.

Supplemental Figure 3. Correlation heatmap of targeted assays: **A)** steroids, **B)** sphingolipids and **C)** ratios of sphingolipids to steroids.

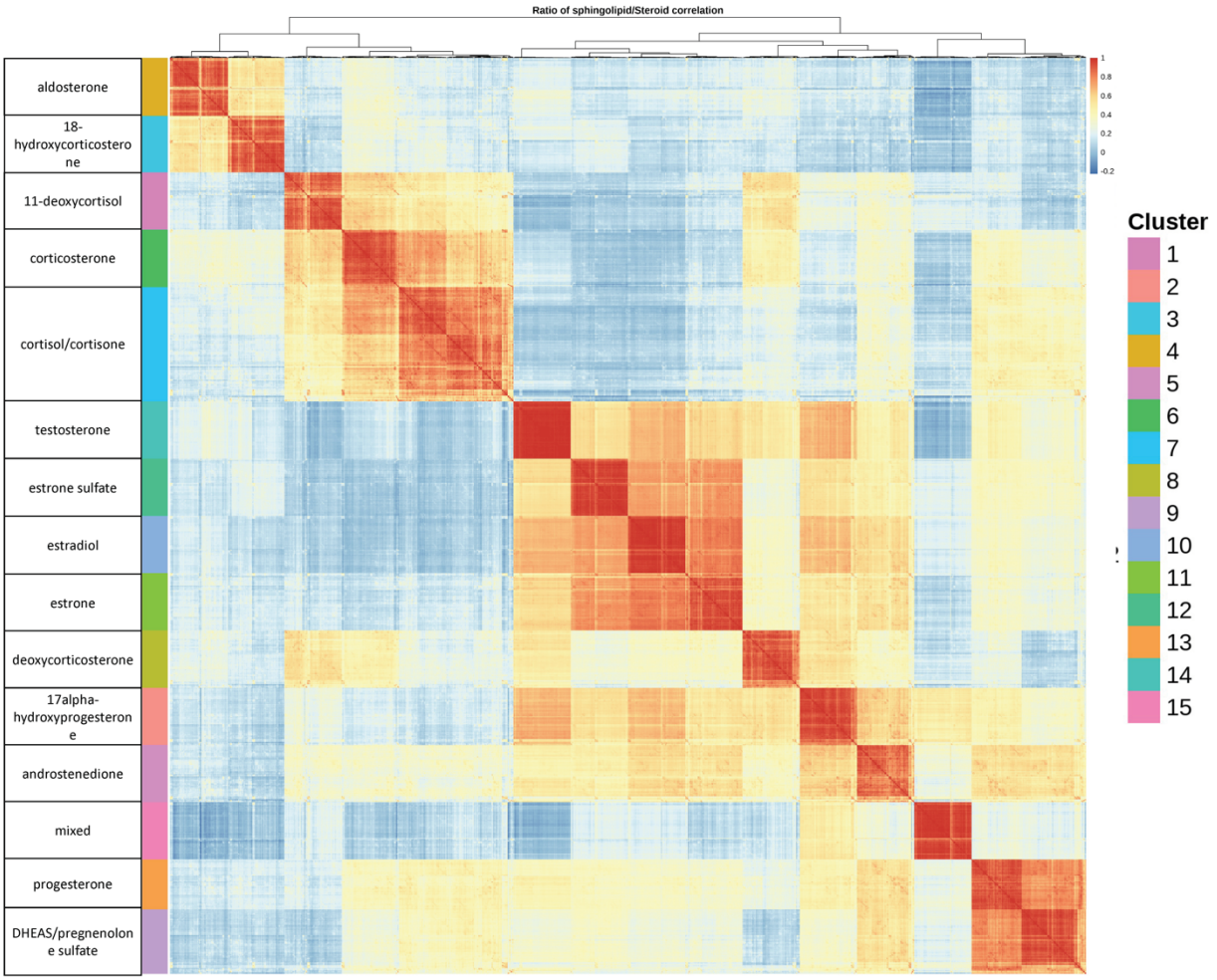
A)



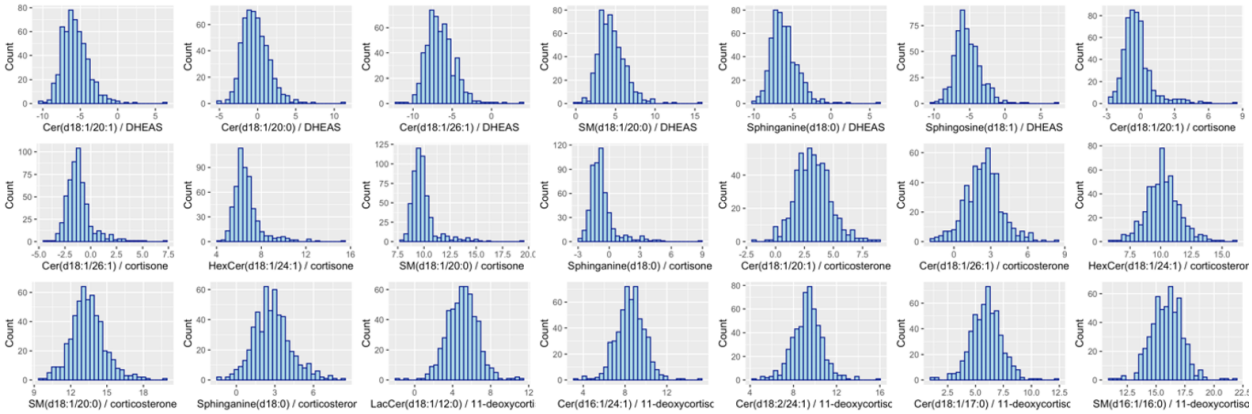
B)


Sphingolipid correlation
1
0.8
0.6
0.4
0.2
0
-0.2
Cer-1P(d18:1/24:1)
Cer-1P(d18:1/16:0)
Cer-1P(d18:1/24:0)
HexCer(d18:1/18:1)
SM(d18:1/18:1)
HexCer(d18:1/24:1)
HexCer(d18:1/20:0)
HexCer(d18:1/16:0)
HexCer(d18:1/18:0)
HexCer(d18:1/14:0)
HexCer(d18:1/22:0)
LacCer(d18:1/22:0)
LacCer(d18:1/24:0)
LacCer(d18:1/18:0)
LacCer(d18:1/20:0)
Glucosph
LacCer(d18:1/18:1)
LacCer(d18:1/17:0)
LacCer(d18:1/16:0)
LacCer(d18:1/14:0)
LacCer(d18:1/24:1)
SM(d18:1/17:0)
Cer(d18:1/10:0)
LacCer(d18:1/12:0)
HexCer(d18:1/12:0)
Cer(d18:1/12:0)
SM(d18:2/16:0)
SM(d18:1/20:0)
SM(d18:1/16:0)
SM(d18:2/24:1)
SM(d18:1/24:1)
SM(d18:2/22:0)
SM(d18:2/24:0)
SM(d18:1/22:0)
SM(d18:1/24:0)
SM(d16:1/24:0)
SM(d18:1/14:0)
SM(d16:1/16:0)
SM(d16:1/24:1)
Sphingosine-1P(d18:2)
Sphinganine-1P
Sphinganine-1P(d16:1)
Sphingosine-1P(d17:1)
3-Dehydrosphinganine
Sphingosine(d16:1)
Sphinganine
Sphingosine
Sphingosine(d18:2)
SM(d18:1/18:0)
SM(d18:2/18:0)
Cer(d18:1/18:1)
Cer(d18:1/20:1)
Cer(d18:1/18:0)
Cer(d18:2/18:0)
Cer(d16:1/24:1)
Cer(d16:1/18:0)
SM(d16:1/18:0)
Cer(d18:1/17:0)
Cer(d18:1/16:0)
Cer(d18:1/20:0)
Cer(d18:2/24:1)
Cer(d18:1/26:1)
Cer(d18:1/22:0)
Cer(d18:1/24:0)
Cer(d16:1/24:0)
Cer(d18:2/24:0)
Cer(d16:1/16:0)
Cer(d18:1/14:0)
Cer(d18:1/26:0)
Cer(d18:1/24:0)
HexCer(d18:1/24:0)
DhCer(d18:0/16:0)
DhCer(d18:0/24:0)
DhCer(d18:0/18:0)
DhCer(d18:0/24:1)

C)


Ratio of sphingolipid/Steroid correlation
aldosterone
18-hydroxycorticosterone
11-deoxycortisol
corticosterone
cortisol/cortisone
testosterone
estrone sulfate
estradiol
estrone
deoxycorticosterone
17alpha-hydroxyprogesterone
androstenedione
mixed
progesterone
DHEAS/pregnenolone sulfate
1
0.8
0.6
0.4
0.2
0
-0.2
Cluster
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

Supplemental Figure 4. The distribution of selected 21 sphingolipid:steroid ratios from the prediction model.



Abbreviations: DHEAS, dehydroepiandrosterone sulfate; Cer, ceramide; SM, sphingomyelin

Supplemental Figure 5. Cox model results and cumulative incidence plots for the time until the first asthma exacerbation were revealed by all 16 replicated ratios in the predictive models in MGBB-LLF (formerly ODOLLFA).