Supporting information

Large-scale prospective serum metabolomic profiling reveals candidate predictive biomarkers for suspected preeclampsia patients

Yan Cao [1,2†], Lanlan Meng [1,2†], Yifei Wang[3], Shenglong Zhao[4], Yuanyuan Zheng[4], Rui Ran[5], Jie Du[5], Hongqiang Wu[5], Jiaqi Han[1,2], Zhengwen Xu[1,2], Yifan Lu[1,2], Lin Liu[1,2], Lu Chen[1,2], Jing Wang[1,2], Youran Li[1,2], Yanhong Zhai[1,2*], Zhi Sun[6*], Zheng Cao[1,2*]

[1]Department of Laboratory Medicine, Beijing Obstetrics and Gynecology Hospital, Capital Medical University. Beijing Maternal and Child Health Care Hospital, Beijing, China

[2]Center of Clinical Mass Spectrometry, Beijing Obstetrics and Gynecology Hospital, Capital Medical University. Beijing Maternal and Child Health Care Hospital, Beijing, China

[3]Department of Obstetrics and Gynecology, Shanghai Tenth People's Hospital, Tongji University

[4]Department of Obstetrics, Beijing Obstetrics and Gynecology Hospital, Capital Medical University. Beijing Maternal and Child Health Care Hospital, Beijing, China

[5]Biotree Metabolomics Technology Research Center, Shanghai, China

[6]Department of Pharmacy, The First Affiliated Hospital of Zhengzhou University

†Yan Cao and Lanlan Meng contributed equally to this work and share first authorship.

* Corresponding author

Zheng Cao

Beijing Obstetrics and Gynecology Hospital, Capital Medical University

251 Yaojiayuan Road, Beijing, China 100026.

Tel: 86-10-52276406

Email: zhengcao2011@ccmu.edu.cn

Zhi Sun

The First Affiliated Hospital of Zhengzhou University

No.1 Jianshe East Road, Zhengzhou Henan Province, 450052, P. R. China,

Email: sunzhi2013@163.com

Yanhong Zhai

Beijing Obstetrics and Gynecology Hospital, Capital Medical University

251 Yaojiayuan Road, Beijing, China 100026.

Tel: 86-10-52276406

Email: zhaiyanhong2006@ccmu.edu.cn

Figure legends

Supplemental Fig 1 Total ion chromatogram of plasma samples in the positive (A) and negative ion mode (B).

Supplemental Fig 2 Score scattering plot generated from principal component analysis of all plasma samples in the four groups.

Supplemental Fig 3 Statistical validation of the corresponding PLS-DA model by permutation analysis (permutation No.: 100).

Supplemental Fig 4 Score scattering plots generated from principal component analysis of all QC samples in the positive (A) and negative ion mode (B).

Supplemental Fig 5 Score scattering plots (A), loading plots (B) and permutation (C) of the corresponding OPLS-DA model (HC *vs.* PE).

Supplemental Fig 6 The minimum penalty coefficient model constructed using the LASSO regression model.

Supplemental Fig 7 Score scattering plots of the PCA model (suspected-PE^+ *vs.* suspected-PE^-).

Supplemental Fig 8 Permutation of the OPLS-DA model (suspected-PE^+ *vs.* suspected-PE^-).

Table captions

Table S1 The metabolites with the highest potential prognostic significance were identified by LASSO regression analysis.

Table S2 Compared with healthy controls, pathways involved in differential down-regulated metabolites in PE patients ($p < 0.05$, FDR < 0.25).

Table S3 Compared with healthy controls, pathways involved in differential up-regulated metabolites in PE patients ($p < 0.05$, FDR < 0.25).

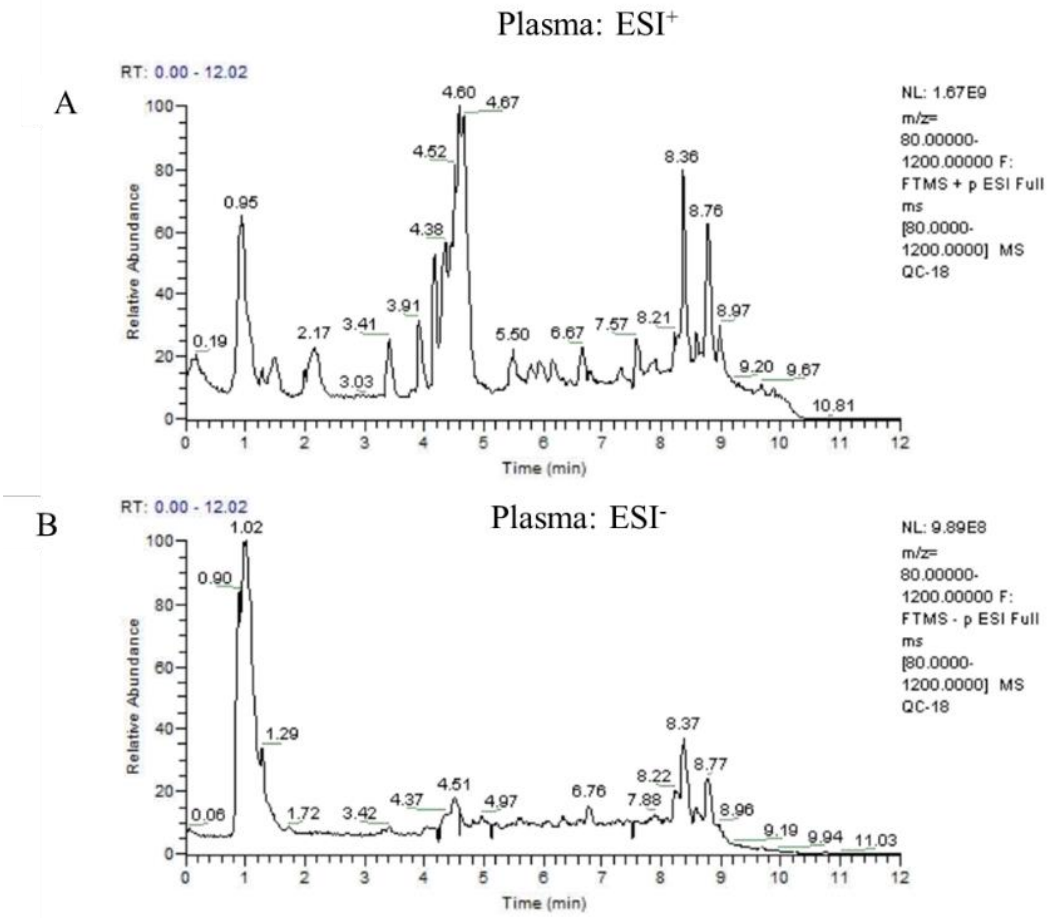
Table S4 Differential metabolites between suspected-PE^+ group and suspected-PE^- group.

Supplementary Methods

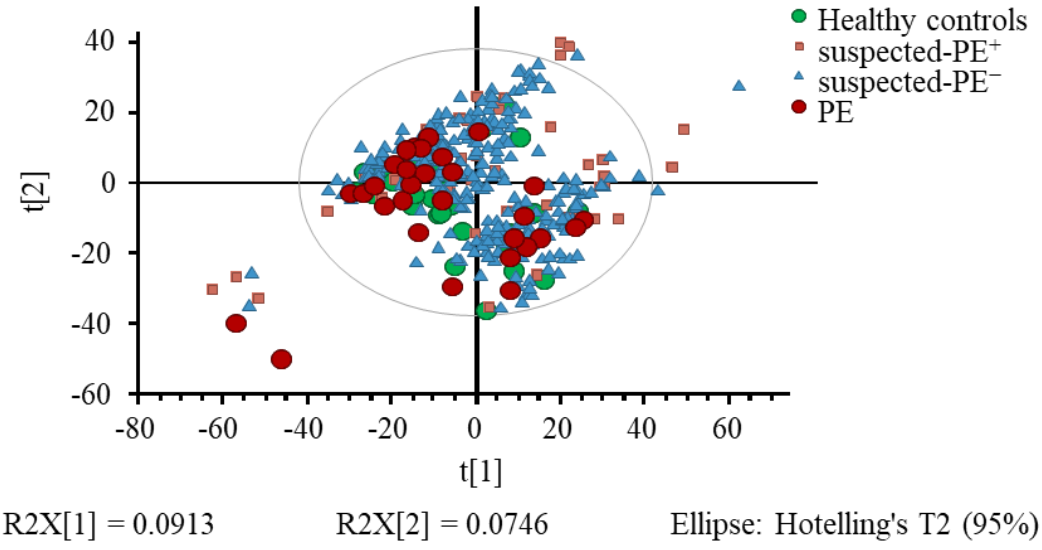
Untargeted metabolomics profiling analysis for serum

Untargeted metabolomics analysis was performed on an UHPLC system (Vanquish, Thermo Fisher Scientific) tandem Q Exactive mass spectrometer (Thermo Fisher Scientific) equipped with an electrospray ionization (ESI) interface. Chromatographic separation was performed on a ACQUITY UPLC BEH C18 column (100 mm*2.1 mm，1.7 μm，Waters，USA), with a column temperature of 40 °C, mobile phase A (water containing 0.1% formic acid) and B (ACN), an elution gradient (0-1.0 min, 5%B; 1.0-9.0 min, 5~100%B; 9.0-12.0 min, 100%B). The auto-sampler temperature was 4 °C, and the injection volume was 5 μL.

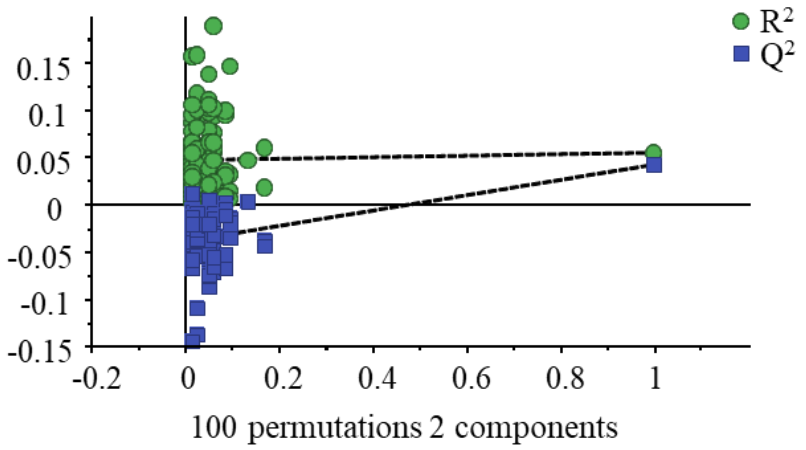
A Q-Exactive Orbitrap mass spectrometer was used to acquire MS/MS spectra on data-dependent acquisition mode. In this mode, the acquisition software (Xcalibur 4.0.27, Thermo Fisher Scientific) continually assessed the full-scan MS data. Ionization conditions for MS were positive ion spray (ESI^+) mode detection, spray voltage of 3500 V, sheath gas and auxiliary gas of 40 and 10 arb, capillary temperature of 300 °C, full MS resolution of 70,000, and MS/MS resolution of 17500. When the negative ion spray (ESI^-) mode was detected, the spray voltage was adjusted to 2800 V, and the other parameters remained the same, as those for the ESI^+ mode.



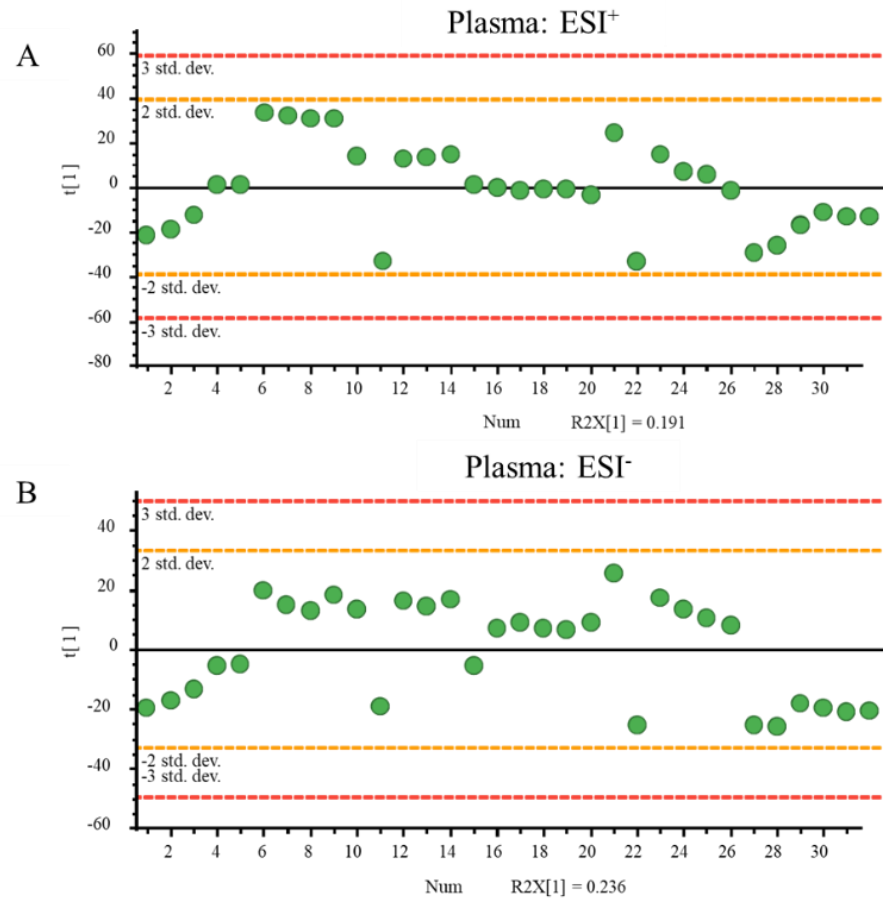
Supplemental Fig 1 Total ion chromatogram of serum samples in the positive (A) and negative ion mode (B).



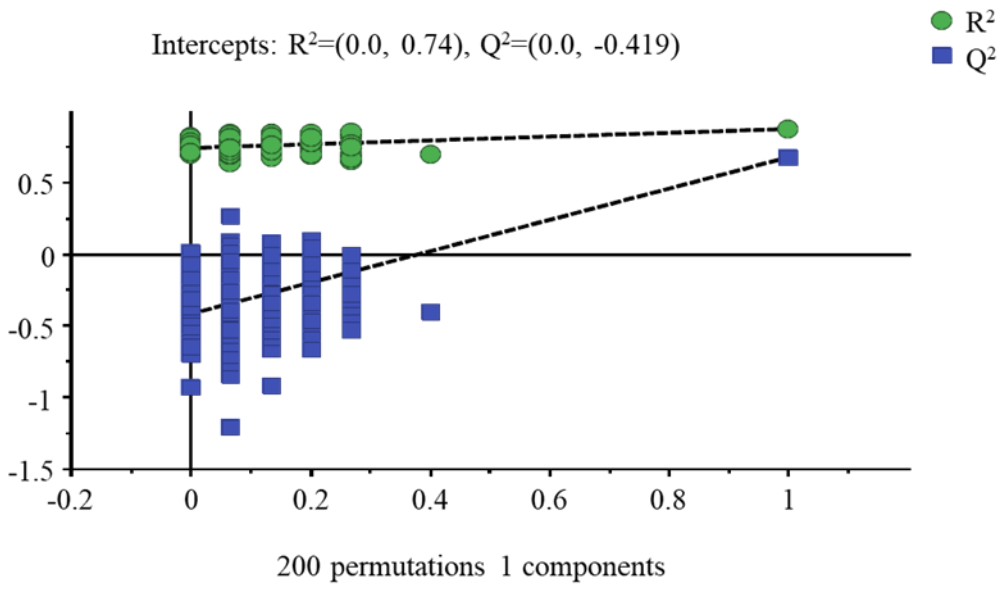
Supplemental Fig 2 Score scattering plot generated from principal component analysis of all serum samples in the four groups.



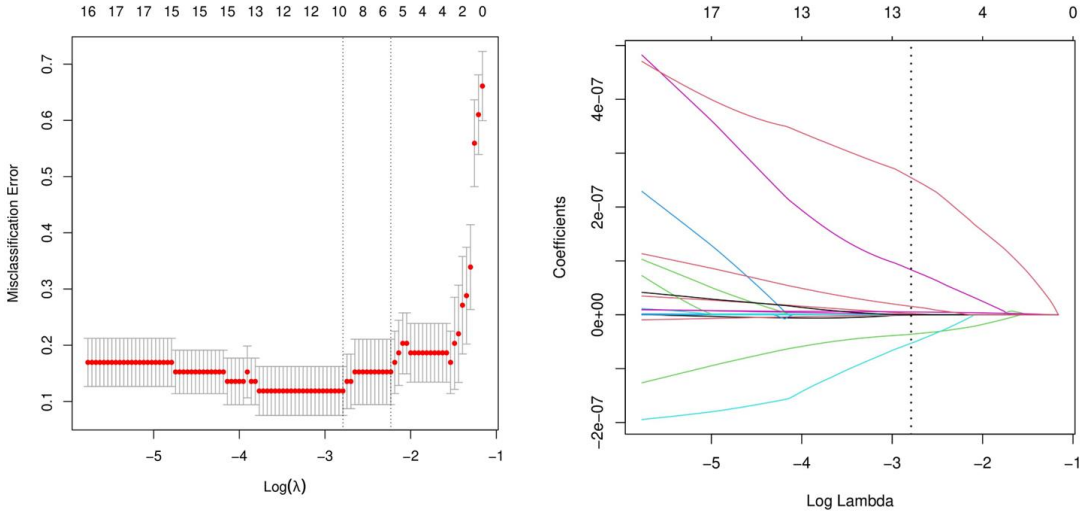
Supplemental Fig 3 Statistical validation of the corresponding PLS-DA model by permutation analysis (permutation No.: 100).



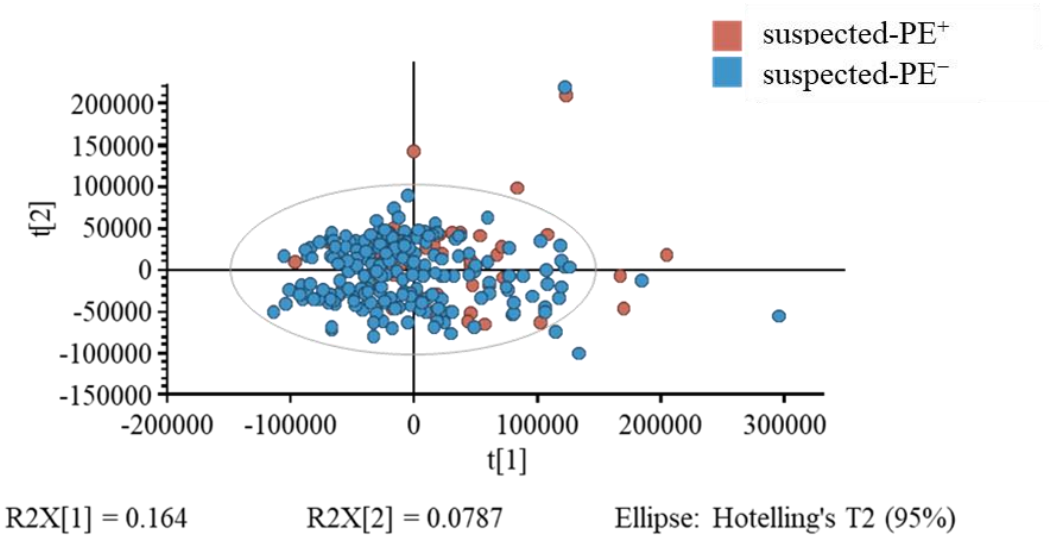
Supplemental Fig 4 Score scattering plots generated from principal component analysis of all QC samples in the positive (A) and negative ion mode (B).



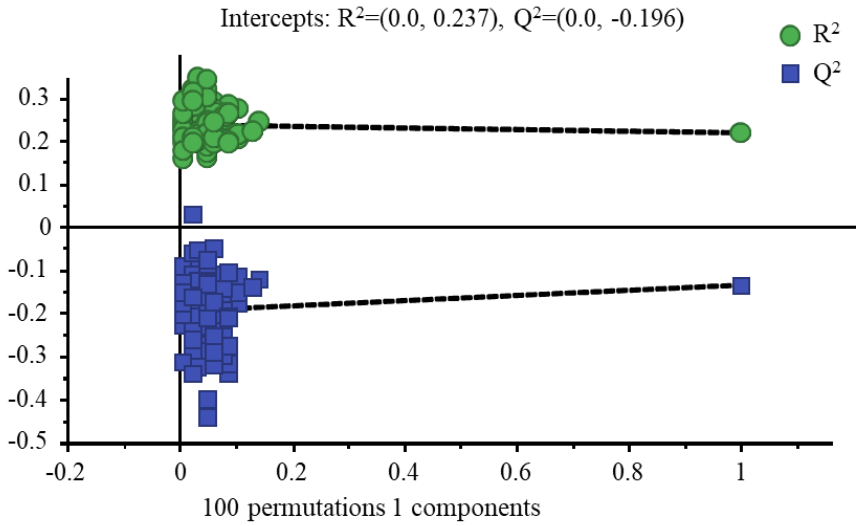
Supplemental Fig 5 Score scattering plots (A), loading plots (B) and permutation (C) of the corresponding OPLS-DA model (HC *vs.* PE).



Supplemental Fig 6 The minimum penalty coefficient model constructed using the LASSO regression model.



Supplemental Fig 7 Score scattering plots of the PCA model (suspected-PE^{+} *vs.* suspected-PE^{-}).



Supplemental Fig 8 Permutation of the OPLS-DA model (suspected-PE^{+} *vs.* suspected-PE^{-}).

Table S1 The metabolites with the highest potential prognostic significance were identified by LASSO regression analysis.

	Coefficient
(Intercept)	-2.4328758
Dehydroepiandrosterone sulfate	-8.272E-10
Cortisol	-3.609E-08
Valproic acid glucuronide	-5.293E-08
L-Proline	5.1717E-10
3-Hydroxybutyric acid	2.4022E-09
L-Histidine	8.3445E-08
2-Hydroxybutyric acid	4.7476E-11
L-Glutamic acid	2.5446E-07
Citric acid	5.1712E-09
3-Phosphonooxypyruvate	1.5572E-08

Table S2 Compared with healthy controls, pathways involved in differential down-regulated metabolites in PE patients ($p < 0.05$, FDR < 0.25).

Pathway	*p*	-log(*p*)	Holm *p*	FDR	Impact	Match metabolites
Steroid hormone biosynthesis	0.0028859	2.5397	0.23088	0.23088	0.02707	Dehydroepiandrosterone sulfate, Cortisol, Estrone 3-sulfate

Table S3 Compared with healthy controls, pathways involved in differential up-regulated metabolites in PE patients ($p < 0.05$, FDR < 0.25).

Pathway	*p*	-log(*p*)	Holm *p*	FDR	Impact	Match metabolites
Phenylalanine, tyrosine and tryptophan biosynthesis	0.001203	2.9197	0.096237	0.096237	1	L-Phenylalanine, L-Tyrosine
Nitrogen metabolism	0.0029542	2.5296	0.23338	0.11817	0	L-Glutamate, L-Glutamine
Phenylalanine metabolism	0.0070274	2.1532	0.54111	0.14055	0.3109	L-Phenylalanine, L-Tyrosine
Glyoxylate and dicarboxylate	0.010241	1.9896	0.77833	0.14878	0.03175	Citrate, Glutamate,

metabolism						L-Glutamine
Glycine, serine and threonine metabolism	0.011158	1.9524	0.83688	0.14878	0.02475	Choline, L-Threonine, 3-Phosphonooxypyruvate
Biosynthesis of unsaturated fatty acids	0.014191	1.848	1	0.16218	0	Linoleate, Arachidonate
Arginine biosynthesis	0.016691	1.7775	1	0.16691	0.11675	L-Glutamate, L-Glutamine
Butanoate metabolism	0.019089	1.7192	1	0.16968	0	(R)-3-Hydroxybutanoate, L-Glutamate
Histidine metabolism	0.021623	1.6651	1	0.17299	0.22131	L-Glutamate, L-Histidine

Table S4 Differential metabolites between suspected-PE^+ group and suspected-PE^- group.

Metabolite	VIP value	*P*-value	Log (FC)
2-Methyl-3-hydroxy-5-formylpyridine-4-carboxylate	2.071	0.001	-0.388
LysoPC(20:2(11Z,14Z))	2.215	$9.572e^{-5}$	-0.326
L-Proline	2.086	$1.323e^{-5}$	-0.292
Dehydro-ascorbide	1.546	0.009	-0.270
Leucinic acid	2.431	0.000	0.260
gamma-Glutamylleucine	2.305	0.000	0.268
D-Glucono-1,5-lactone 6-phosphate	2.533	0.000	0.272
gamma-Glutamylalanine	1.948	0.009	0.275
2-Hydroxyvaleric acid	2.023	0.016	0.297

LysoPC(16:1(9Z)/0:0)	1.671	0.003	0.306
PC(DiMe(13,5)/MonoMe(13,5))	1.691	0.002	0.308
ADP-D-glycero-beta-D-manno-heptose	1.780	0.011	0.310
Phenylalanyl-Tryptophan	1.966	0.016	0.340
Allantoic acid	1.689	0.000	0.379
LysoPC(16:0)	1.738	0.031	0.386
TG(14:0/22:0/o-18:0)	1.975	0.020	0.402
TG (18:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z)/24:1(15Z))	2.264	0.000	0.461
CDP-ribitol	1.598	0.020	0.483
2-O-Galloyl-1,4-galactarolactone	1.680	0.038	0.491