

SUPPLEMENTARY MATERIALS

Umbilical Cord Blood and Urine Collection

For each infant, 10 mL umbilical cord blood was collected using a 10 mL heparin anticoagulant vacuum tube. Samples were frozen thereafter and sent to the Qingdao CDC to reserve at $-70\text{ }^{\circ}\text{C}$.

For morning urine samples, parents of participants brought home five 15-mL polypropylene centrifuge tubes to collect 15 mL of their child's morning urine each time for five days in a row. Samples were frozen immediately after collection and sent back to their hospital at birth, and then were reserved out of the sun at $-20\text{ }^{\circ}\text{C}$. These samples were finally sent to the Qingdao CDC and kept at $-70\text{ }^{\circ}\text{C}$ for later analysis.

Urinary Hydroxyl PAHs Testing

We planned to collect urine samples in infants at the age of one year. But most infants at the age of one year were not able to urinate spontaneously. Thus, we collected the morning first-void urine samples when the children were able to urinate spontaneously between 2 and 3 years old. The caregivers took back the bottles and collected the urine samples in a consecutive 5 days. After mixing the urine samples collected continuously, we accurately removed 10 mL of the mixed sample to a blowing bottle, and then added 30 μe hydrolysate of lglucoaldosidase and aryl sulfatase. Later we placed the fully blended mixture to a $37\text{ }^{\circ}\text{C}$ thermostatic incubator to stay overnight. Afterwards, we collected these hydrolyzed samples to enrich and purify in a C18 SPE column (6 mL/500 mg). The C18 columns were first activated by a 5 mL of acetonitrile and a 10 mL of water. Samples passed columns at a speed of 1 mL/min. After samples passed, these C18 columns were cleaned with a 3 mL of water and drained under negative pressure. Then, they were eluted with an 8 mL of formic acid-acetonitrile (0.3%, v/v). Eluents were then dried and condensed to less than 1 mL by nitrogen under $35\text{ }^{\circ}\text{C}$, and diluted to 1 mL under mobile phase with the initial ratio. They were ready for LC-MS/MS testing after being filtered by a 0.22 μm of membrane.

An Acquity UPLC BEH-C18 (100 mmBEH-C18 BEH-C18C18der mobUSA) was equipped. The flow rate was 0.5 mL/min. Mobile phases A and B were water and acetonitrile, respectively. The gradient of mobile phase A was programmed as follows: 60% A at 0.0–0.2 min, then the percentage of A was decreased linearly to 5% between 0.2–3.0 min; 5% A between 3.0–4.0 min, then the percentage of A was increased to 60% from 4.0–4.1 min, and 60% A between 4.1–6.0 min. The temperature of the chromatography column was maintained at $40\text{ }^{\circ}\text{C}$ in a column oven. The injection volume was 10 μL in a partial loop using a needle overfill injection mode.

The MS/MS detector was equipped with an ESI interface and operated in positive ionization mode. The quantitative analyses were carried out in multiple reaction monitoring (MRM) mode. The conditions for the MS/MS detector were set as follows: capillary voltage, 3.00 kV; ion source temperature, $150\text{ }^{\circ}\text{C}$; desolvation gas flow, 1,000 L/hr; desolvation gas temperature, $500\text{ }^{\circ}\text{C}$; and cone gas flow, 50 L/hr. Other parameters like quantity ion pairs (m/z), cone voltage and collision energy were listed in [Supplementary Table S1](#) (available in www.besjournal.com).

Urine Metabolomics Analysis

After the morning first-void urine sample was slowly thawed at $4\text{ }^{\circ}\text{C}$, it was added with 100 μL precooled methanol/acetonitrile solution (1:1, v/v), mixed by vortex, stand at $-20\text{ }^{\circ}\text{C}$ for 60 min, and centrifuged at 14,000 g and $4\text{ }^{\circ}\text{C}$ for 20 min. The supernatant was collected and dried under the vacuum. For mass spectrometry, 100 μL acetonitrile aqueous solution (acetonitrile: water = 1:1, v/v) was added for reconstitution, followed by vortex and centrifugation (14,000 g, $4\text{ }^{\circ}\text{C}$, 15 min) to collect the supernatant. To monitor the stability and repeatability of instrument analysis, quality control (QC) samples were prepared by pooling 10 μL of each sample and these were analyzed together with the other samples. The QC samples were inserted regularly and analyzed in every eight samples.

Chromatographic Condition Metabolic profiling of urine samples was performed on an Agilent 1290 Infinity LC system (Agilent Technologies, Santa-Clara, California, USA) coupled with an AB SCIEX Triple TOF 5600 System (AB SCIEX, Framingham, MA, USA) in Shanghai Applied Protein Technology Co., Ltd. (<http://www.aptbiochem.com/>). Chromatographic separation was implemented on ACQUITY UPLC BEH Amide 1.7 μm (2.1 \times 100 mm) columns (Waters Co., Ltd., USA) for both positive and negative models. The column temperature was set at $25\text{ }^{\circ}\text{C}$. The delivery flow rate was 0.3 mL/min, and 2 μL aliquot of each sample was injected onto the

Supplementary Table S1. Differences in OH-PAH levels between high and low exposure groups ($\mu\text{g/g Cr}$)

OH-PAHs	High exposure group (GM \pm GSD, n = 18)	Low exposure group (GM \pm GSD, n = 19)	P
1-OHNap	1.55 \pm 6.18	0.32 \pm 4.76	0.008
2-OHNap	8.70 \pm 2.57	1.69 \pm 2.70	< 0.001
2-OHFlu	6.74 \pm 5.31	0.39 \pm 2.82	< 0.001
2(3)-OHPhe	0.47 \pm 5.94	0.31 \pm 3.55	0.436
1(9)-OHPhe	3.26 \pm 5.80	0.12 \pm 4.46	< 0.001
4-OHPhe	0.11 \pm 4.53	0.07 \pm 2.86	0.372
1-OHPy	0.30 \pm 3.38	0.16 \pm 4.02	0.143
6-OHChr	0.08 \pm 4.43	0.04 \pm 3.08	0.153
Total OH-PAHs	30.63 \pm 2.30	3.83 \pm 2.54	< 0.001

Note. OH-PAH, hydroxyl polycyclic aromatic hydrocarbon; Cr, creatinine; GM, geometric mean; GSD, geometric standard deviation; OHNap, hydroxynaphthalene; OHFlu, hydroxyfluorene; OHPhe, hydroxyphenanthrene; OHPy, hydroxypyrene; OHChr, hydroxychrysene.

column. The mobile phase A (25 mmol/L ammonium acetate and 25 mmol/L ammonia in water) and B (acetonitrile) was used. The gradient elution procedure was as follows: 0–1 min, 95% B; 1–14 min, B varied linearly from 95% to 65%; 14–16 min, B varied linearly from 65% to 40%; 16–18 min, B maintained at 40%; 18–18.1 min, B varied linearly from 40% to 95%; 18.1–23 min, B maintained at 95%; during the whole analysis process, the samples were placed in a 4 °C automatic sampler. To avoid the influence caused by the fluctuation of the instrument detection signal, the random sequence was used for the continuous analysis of samples.

Q-TOF Mass Spectrum Condition Q-TOF/MS was performed on positive ion mode and negative ion mode. Electrospray ionization (ESI) source conditions on Triple TOF were set as follows: ion source gas 1 (Gas1), 60 psi; ion source gas 2 (Gas2), 60 psi; curtain gas (CUR) 30 psi; source temperature, 600 °C; ionsapary voltage floating (ISVF), 5,500 V (+) and –5,500 V (–); TOF/MS scan m/z range, 60–1,000 Da; product ion scan m/z range, 25–1,000 Da; TOF/MS scan accumulation time, 0.20 s/spectra; product ion scan accumulation time 0.05 s/spectra. The secondary MS was obtained using information-dependent acquisition (IDA) in the high sensitivity mode. The parameters were set as follows: declustering potential, 60 V (+) and –60 V (–); collision energy, 35 eV (+) and –15 eV (–); exclude isotopes within 4 Da, candidate ions to monitor per cycle: 6.

Data Processing The raw UPLC-Q-TOF/MS data were converted into the mzXML files by ProteoWizard MS convert tool, and then the XCMS program was used for peak alignment, retention time correction, and peak area extraction. For metabolite structure identification, accurate mass number matching (< 25 ppm) and secondary spectrum matching were used to retrieve the database built by the laboratory.

For the data extracted by XCMSs, the ion peaks with missing values > 50% in the group were deleted. The software SIMCA-P 14.1 (Umetrics, Umea, Sweden) was used for pattern recognition. After the data were preprocessed by Pareto-scaling, multidimensional statistical analysis orthogonal partial least squares discriminant analysis (OPLS-DA) was carried out.

Quality Control and Assurance Spectrogram comparison and Principal Component Analysis (PCA) analysis for QC samples were used for quality control and assurance.

Comparison of Total Ion Chromatograph (TIC) of QC Samples Spectrogram overlap comparison was performed for TICs of QC samples (Figure S1-1). The response intensity and retention time of each chromatographic peak basically overlap, indicating the variation caused by instrument error was small in the whole experiment.

PCA Analysis for all Samples The peaks obtained from all the samples and QC samples were extracted and then PCA analysis was carried out after Pareto-scaling, as shown in Figure S1-2. The QC samples in positive and negative ion modes were closely clustered together, which indicated that the experiment of this project has good repeatability.

Supplementary Table S2. Pathways and biological functions related to the metabolites

Metabolites	Pathway and biological function	Categorized group
D-Glucuronate, L-Gulonic gamma-lactone, L-Threonate	Ascorbate and aldarate metabolism	
Creatine, Urea, Creatinine	Arginine and proline metabolism, Arginine biosynthesis, Glycine, serine and threonine metabolism	
N-Acetyl-L-aspartic acid	Alanine, aspartate and glutamate metabolism	
L-Lysine	Lysine degradation	Amino acid metabolism
Hippuric acid	Phenylalanine metabolism	
Gentisic acid	Tyrosine metabolism	
Citraconic acid	Valine, leucine and isoleucine biosynthesis	
D-Glucuronate	Amino sugar and nucleotide sugar metabolism, Inositol phosphate metabolism, Pentose and glucuronate interconversions	
Isocitrate	Citrate cycle (TCA cycle), Glyoxylate and dicarboxylate metabolism	Carbohydrate metabolism
Acetyl phosphate	Pyruvate metabolism	
D-gluconate, D-Glucono-1,5-lactone	Pentose phosphate pathway	
Glycerophosphocholine	Ether lipid metabolism, Glycerophospholipid metabolism	Lipid metabolism
Adenosine	Sphingolipid signaling pathway	
Adenosine, Deoxyinosine, Hypoxanthine, Urea, Xanthine	Purine metabolism	Nucleotide metabolism
Pseudouridine, Urea	Pyrimidine metabolism	
Pantothenate	beta-Alanine metabolism	Metabolism of other amino acids
Acetyl phosphate	Taurine and hypotaurine metabolism	
L-Lysine, Urea	Biotin metabolism	
Pantothenate	Pantothenate and CoA biosynthesis	Metabolism of cofactors and vitamins
7,8-Dihydrobiopterin	Folate biosynthesis	
4-Pyridoxic acid	Vitamin B6 metabolism	
7-Methylxanthine, Xanthine	Caffeine metabolism	Biosynthesis of other secondary metabolites
Isocitrate	Glucagon signaling pathway	
L-Lysine	Protein digestion and absorption	
Adenosine	Regulation of lipolysis in adipocytes, Renin secretion, Vascular smooth muscle contraction	Organismal Systems
Pantothenate	Vitamin digestion and absorption	
L-Lysine, Urea	ABC transporters	
Adenosine	cAMP signaling pathway, cGMP-PKG signaling pathway, Neuroactive ligand-receptor interaction	Environmental Information Processing
L-Lysine	Aminoacyl-tRNA biosynthesis	Genetic Information Processing
Adenosine	Alcoholism	
Isocitrate	Central carbon metabolism in cancer	
Glycerophosphocholine	Choline metabolism in cancer	Human Diseases
Urea	Epithelial cell signaling in Helicobacter pylori infection	
Adenosine	Morphine addiction, Parkinson disease	

Supplementary Table S3. The quantity ion pairs, cone voltage and collision energy of ten hydroxyl metabolites of PAHs

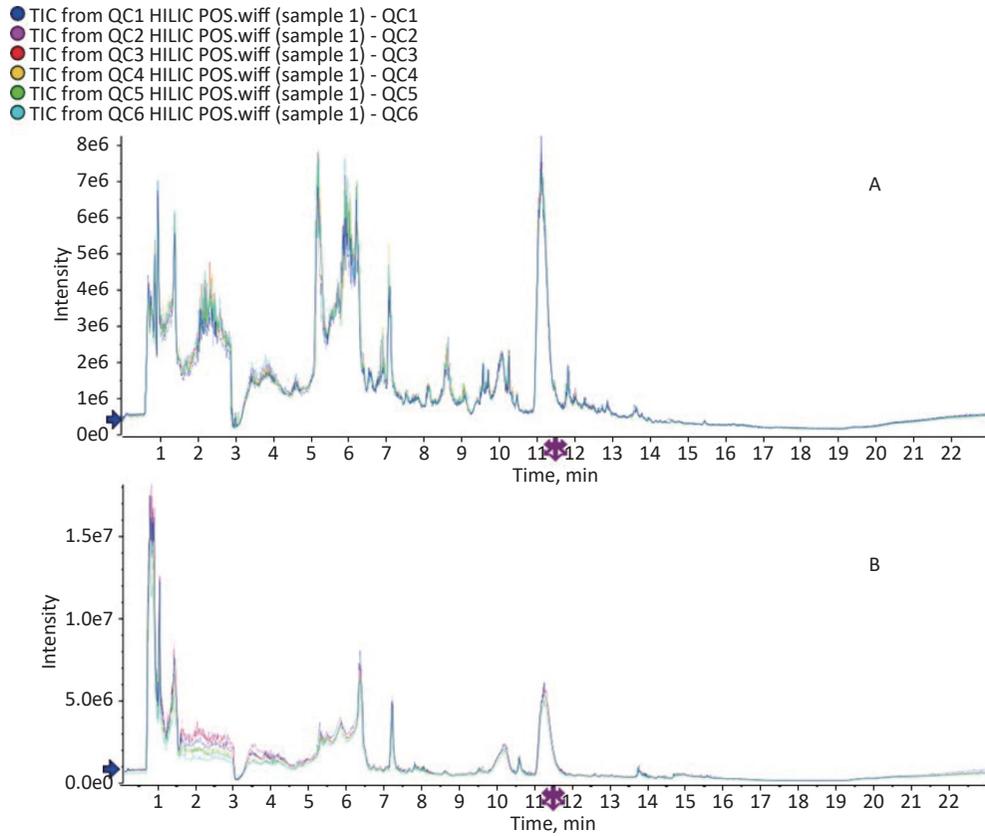
Metabolites	Retention time (min)	Quantity ion pairs (m/z)	Cone voltage/V	Collision energy/eV
1-OHNap	1.45	143/119	25	32
2-OHNap	1.60	143/119	25	32
2-OHFlu	1.98	181/153	21	41
2(3)-OHPh	2.12	193/165	28	35
1(9)-OHPh	2.25	193/165	28	35
4-OHPh	2.35	193/165	28	35
1-OHPy	2.53	217/189	33	30
6-OHChr	2.78	243/215	32	40

Note. OHNap, hydroxynaphthalene; OHFlu, hydroxyfluorene; OHPh, hydroxyphenanthrene; OHPy, hydroxypyrene; OHChr, hydroxychrysene.

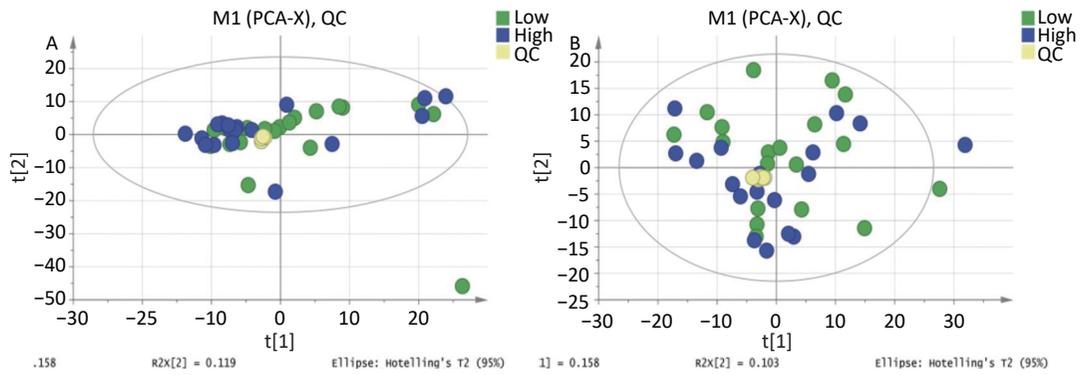
Supplementary Table S4. Detection rate and accuracy test ($n = 6$)

Metabolites	Detection limit ($\mu\text{g/L}$)	Quantity limit ($\mu\text{g/L}$)	Linear range ($\mu\text{g/L}$)	Regression coefficient	Standard concentration ($\mu\text{g/L}$)	Average measured concentration ($\mu\text{g/L}$, mean \pm SD)	Average recovery rate (% \pm CV)
1-OHNap	0.0056	0.0187	0.01–50	$r = 0.9997$	5.00	4.16 ± 0.05	83.2 ± 11.3
2-OHNap	0.0076	0.0253	0.01–50	$r = 0.9998$	5.00	4.32 ± 0.03	86.4 ± 10.4
2-OHFlu	0.0270	0.0900	0.05–50	$r = 0.9996$	5.00	4.07 ± 0.02	81.4 ± 12.5
2(3)-OHPh	0.0088	0.0293	0.01–50	$r = 0.9998$	10.00	11.36 ± 0.04	114 ± 11.2
1(9)-OHPh	0.0053	0.0177	0.01–50	$r = 0.9999$	10.00	11.10 ± 0.10	111.0 ± 11.3
4-OHPh	0.0080	0.0267	0.01–50	$r = 0.9998$	5.00	5.04 ± 0.09	100.0 ± 12.1
1-OHPy	0.0120	0.0400	0.02–50	$r = 0.9997$	5.00	3.80 ± 0.08	76.0 ± 14.3
6-OHChr	0.0010	0.0033	0.002–50	$r = 0.9996$	5.00	3.60 ± 0.10	72.0 ± 15.2

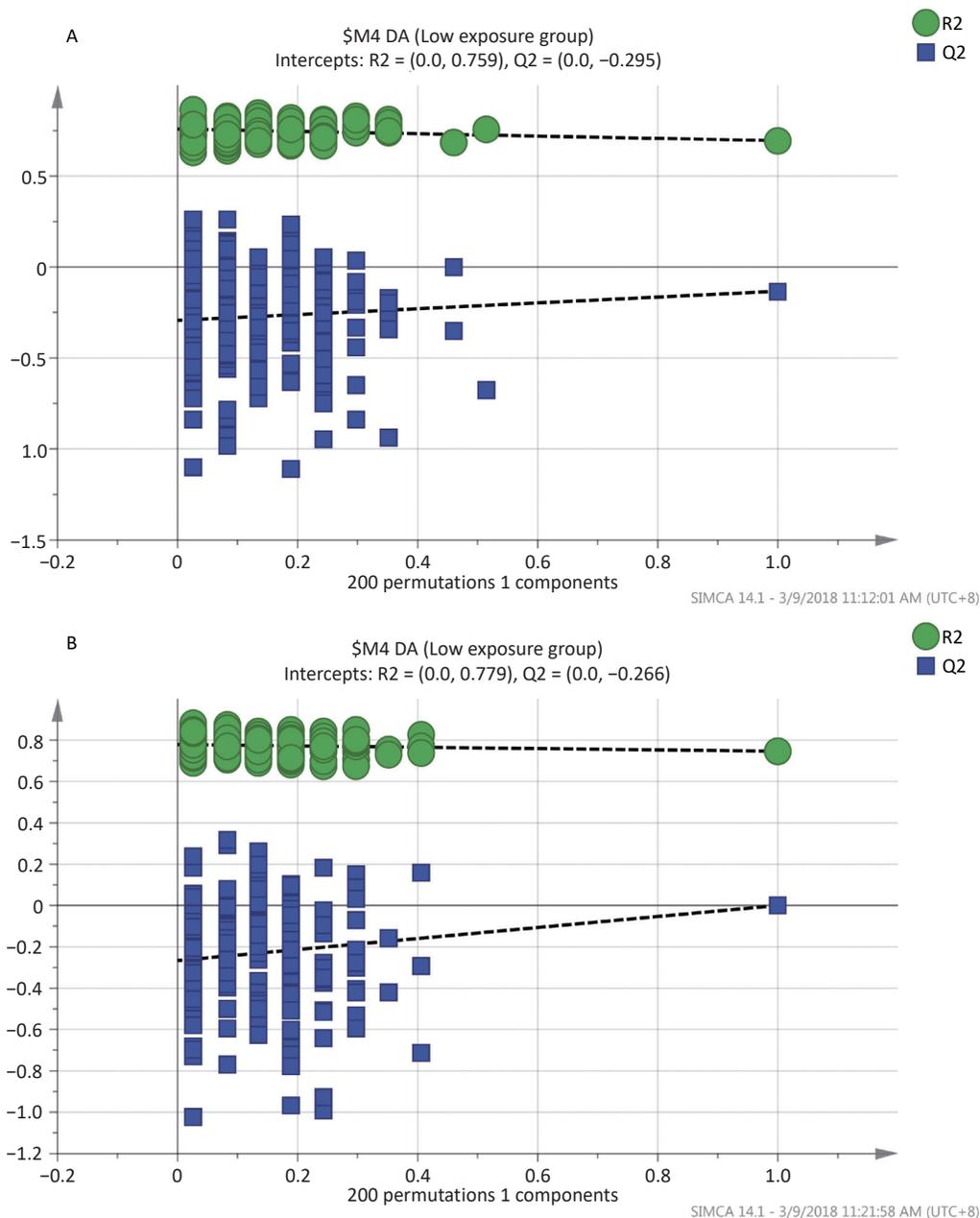
Note. SD, standard deviation; CV, coefficient of variation; OHNap, hydroxynaphthalene; OHFlu, hydroxyfluorene; OHPh, hydroxyphenanthrene; OHPy, hydroxypyrene; OHChr, hydroxychrysene.



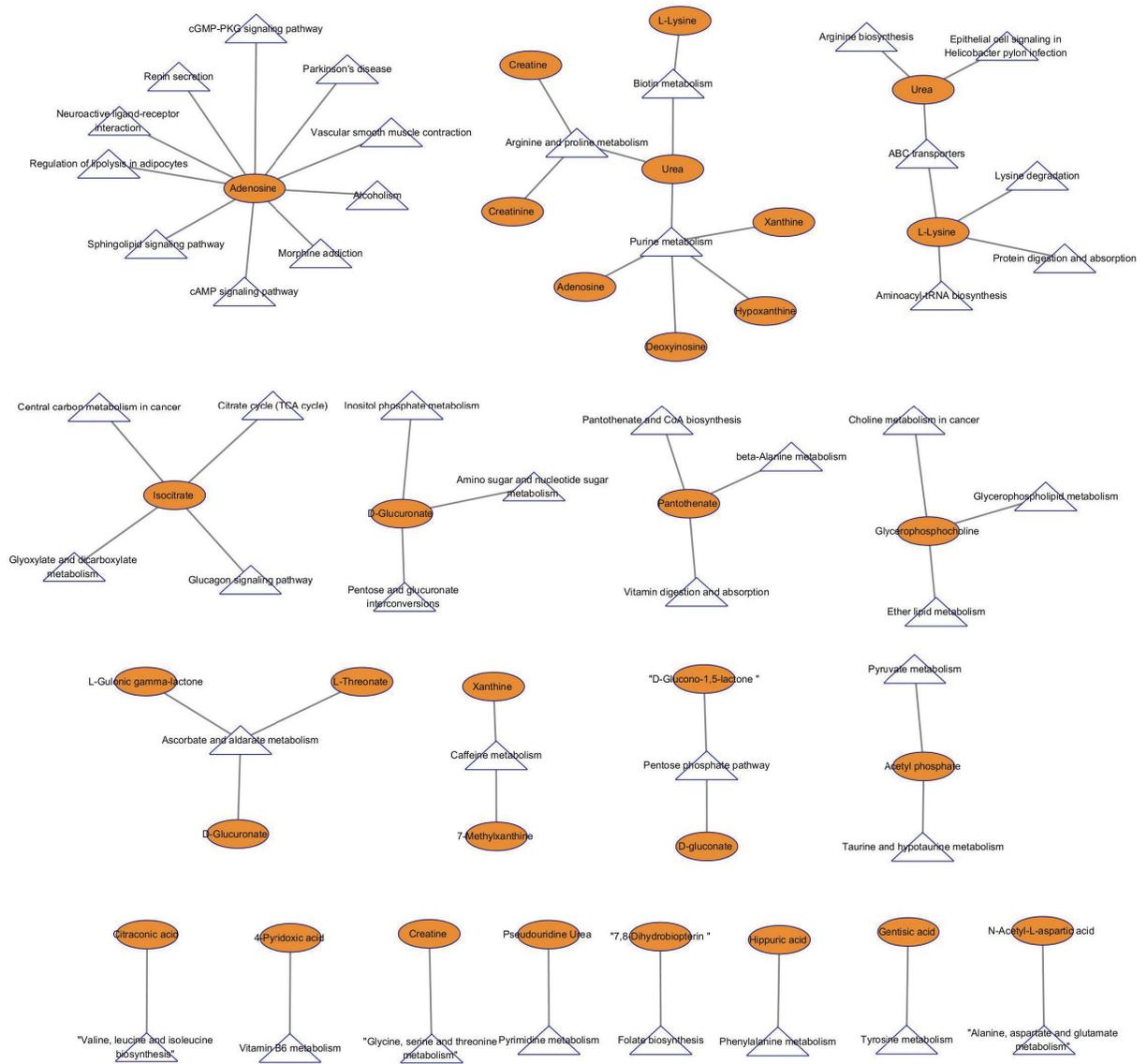
Supplementary Figure S1. TIC spectrogram overlap comparison. The above was for positive ion mode and below was for negative ion mode.



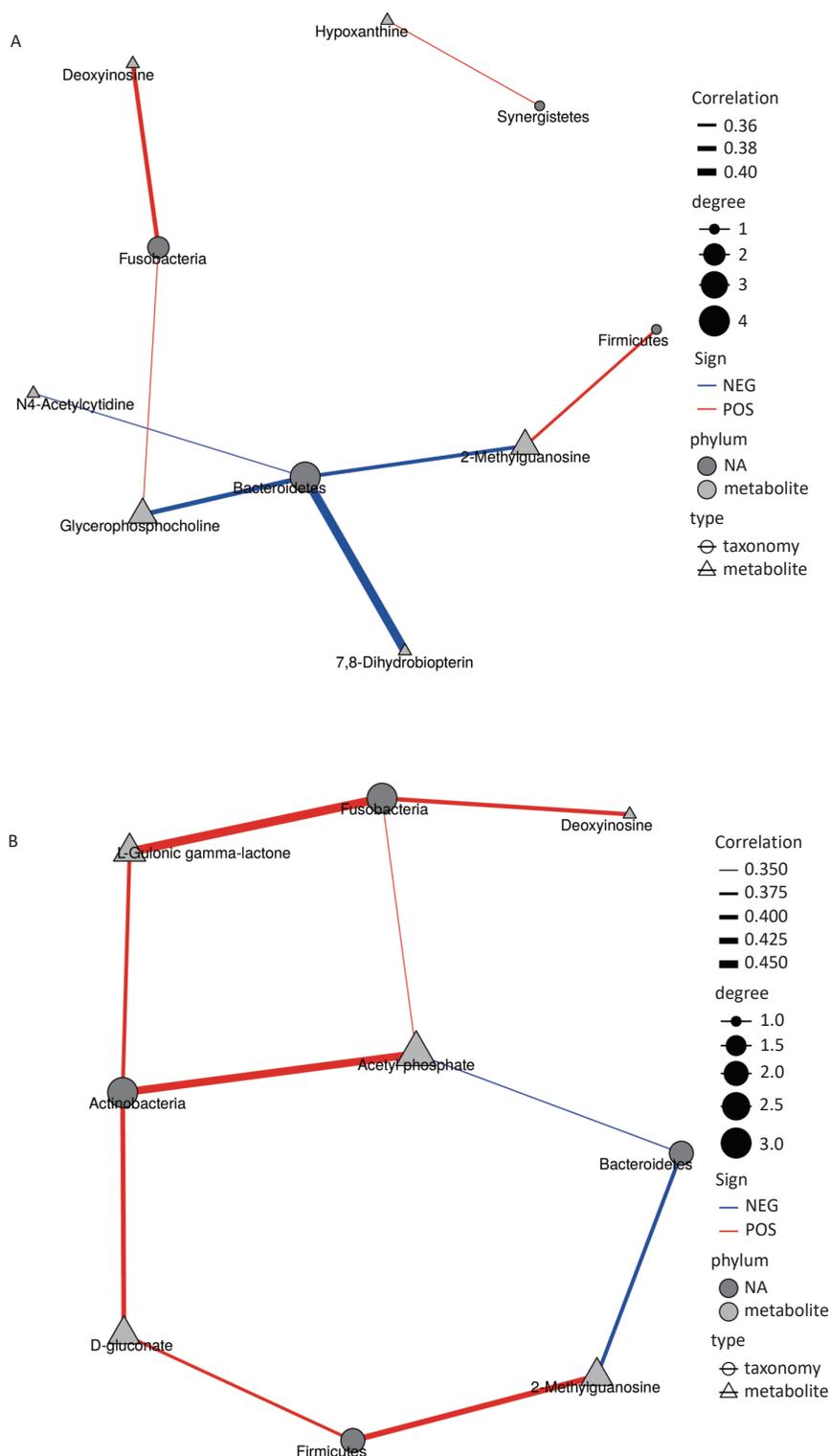
Supplementary Figure S2. PCA score plot for metabolomics analysis of urine samples from the high and low PAHs exposure groups and QC samples. (A) Positive ion mode; (B) Negative ion mode.



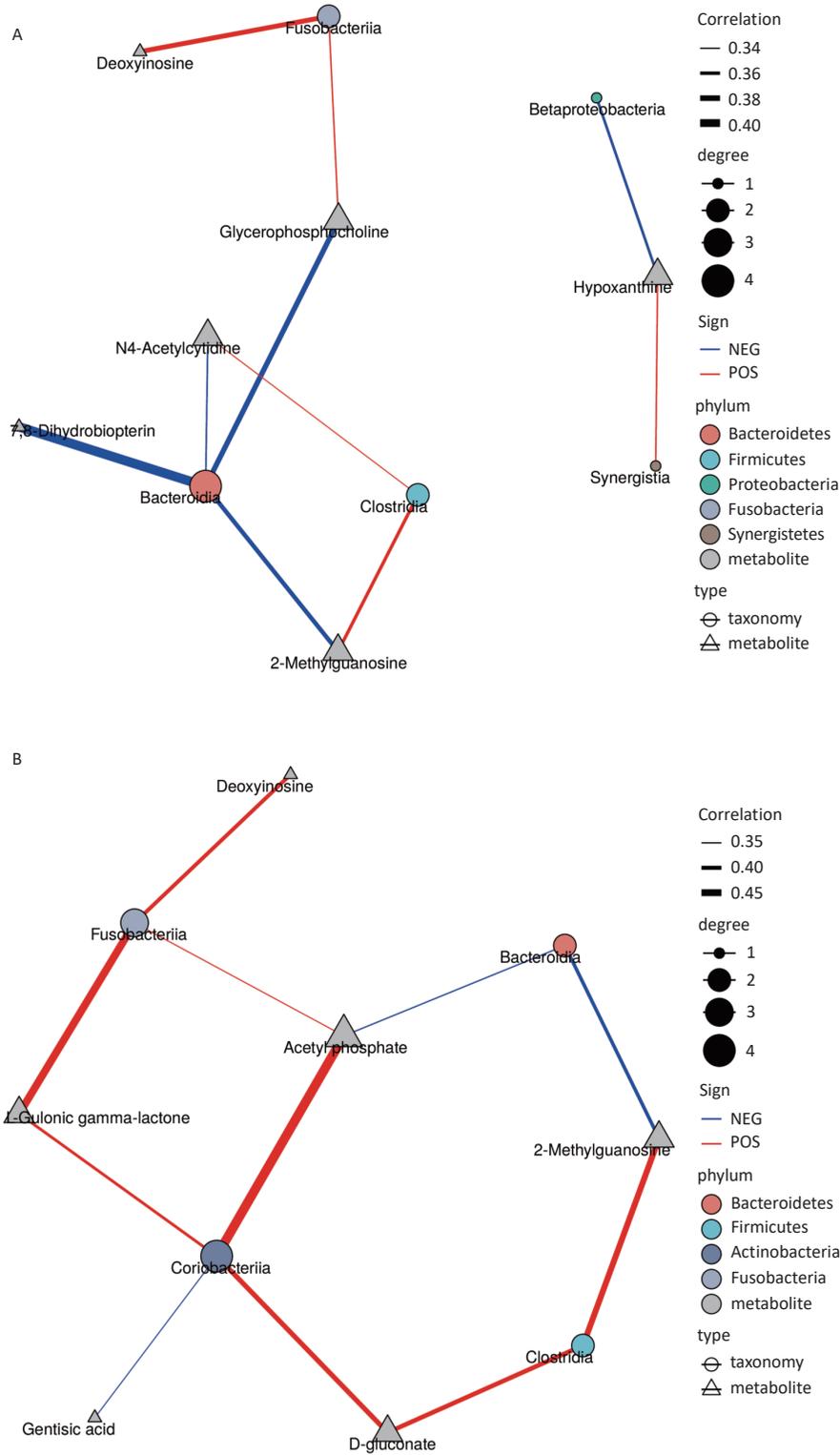
Supplementary Figure S3. Permutation test results for OPLS-DA models. (A) Positive ion mode; (B) Negative ion mode.



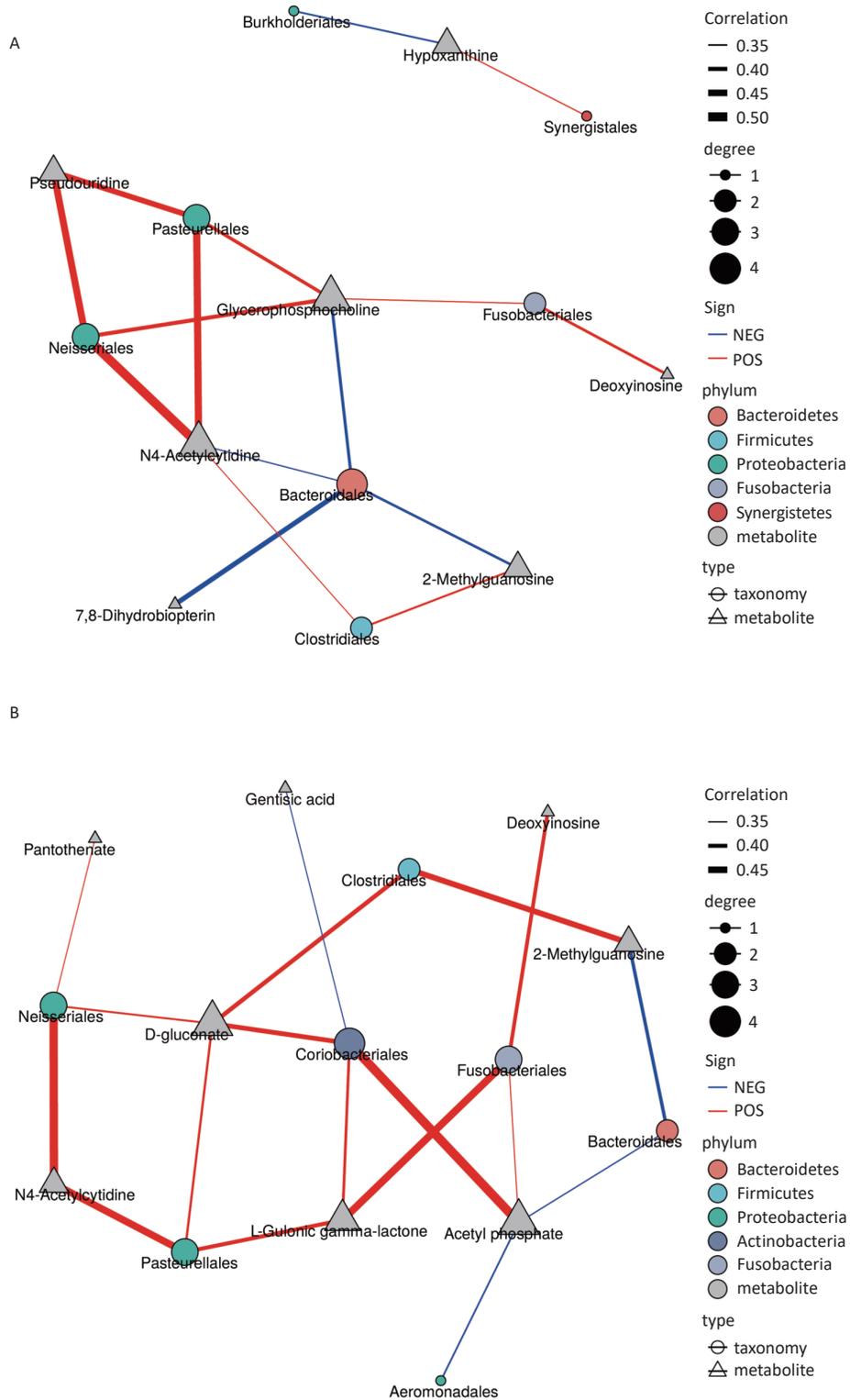
Supplementary Figure S4. Networks for the metabolites and related pathways and biological functions. Orange ellipse represents the metabolite and triangle the pathway or biological function.



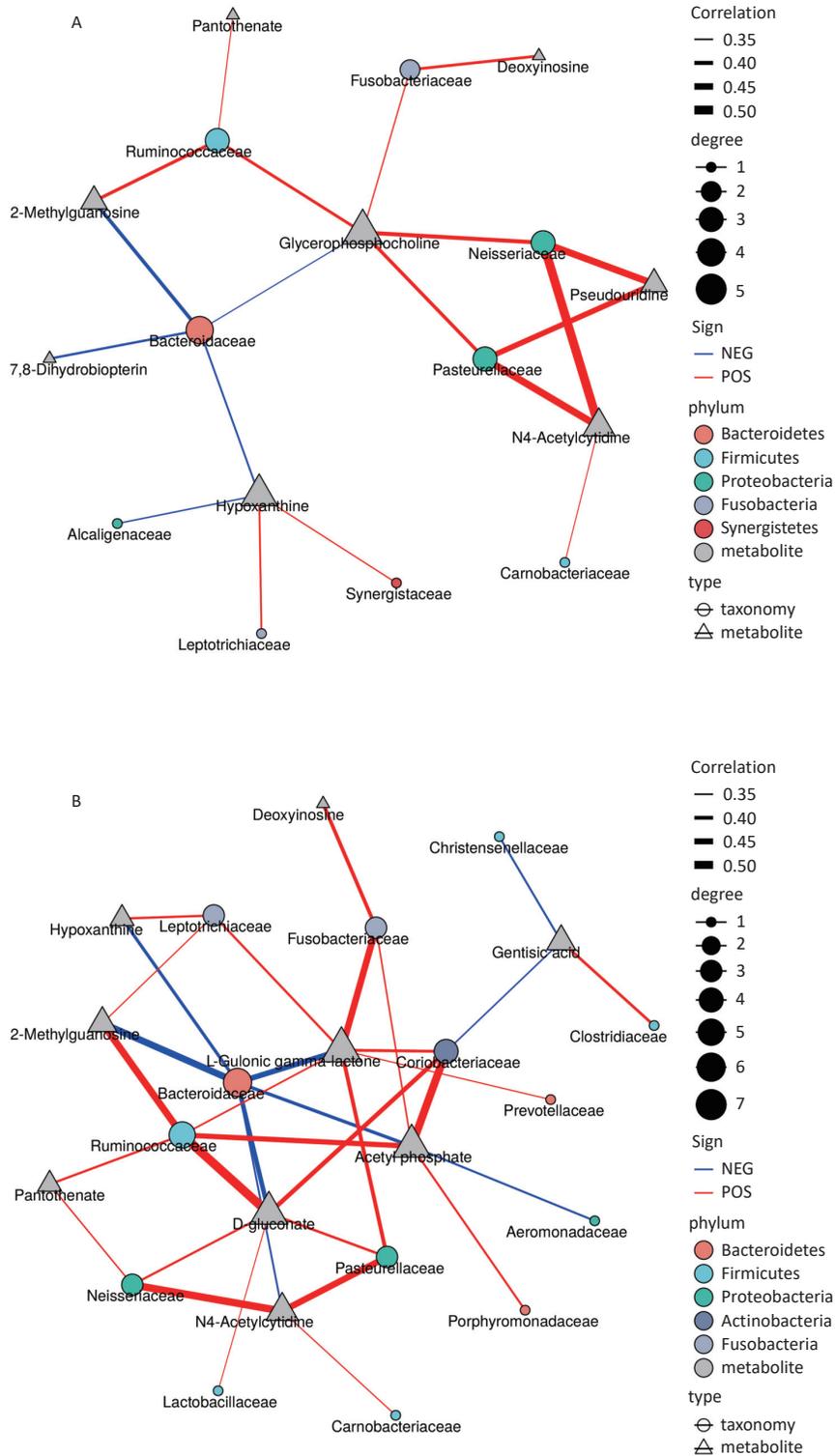
Supplementary Figure S5. Microbiota (phylum)-metabolites correlation network based on Spearman’s correlation coefficients. Each node represents one genus (circle) or metabolite (triangle), and two nodes are linked if the correlation was significant (two-sided pseudo $P < 0.05$) and more than 0.3. Lines between nodes show positive correlations (red lines) or negative correlations (blue lines). The node size is proportional to the number of correlated metabolites. (A) positive ion mode; (B) negative ion mode.



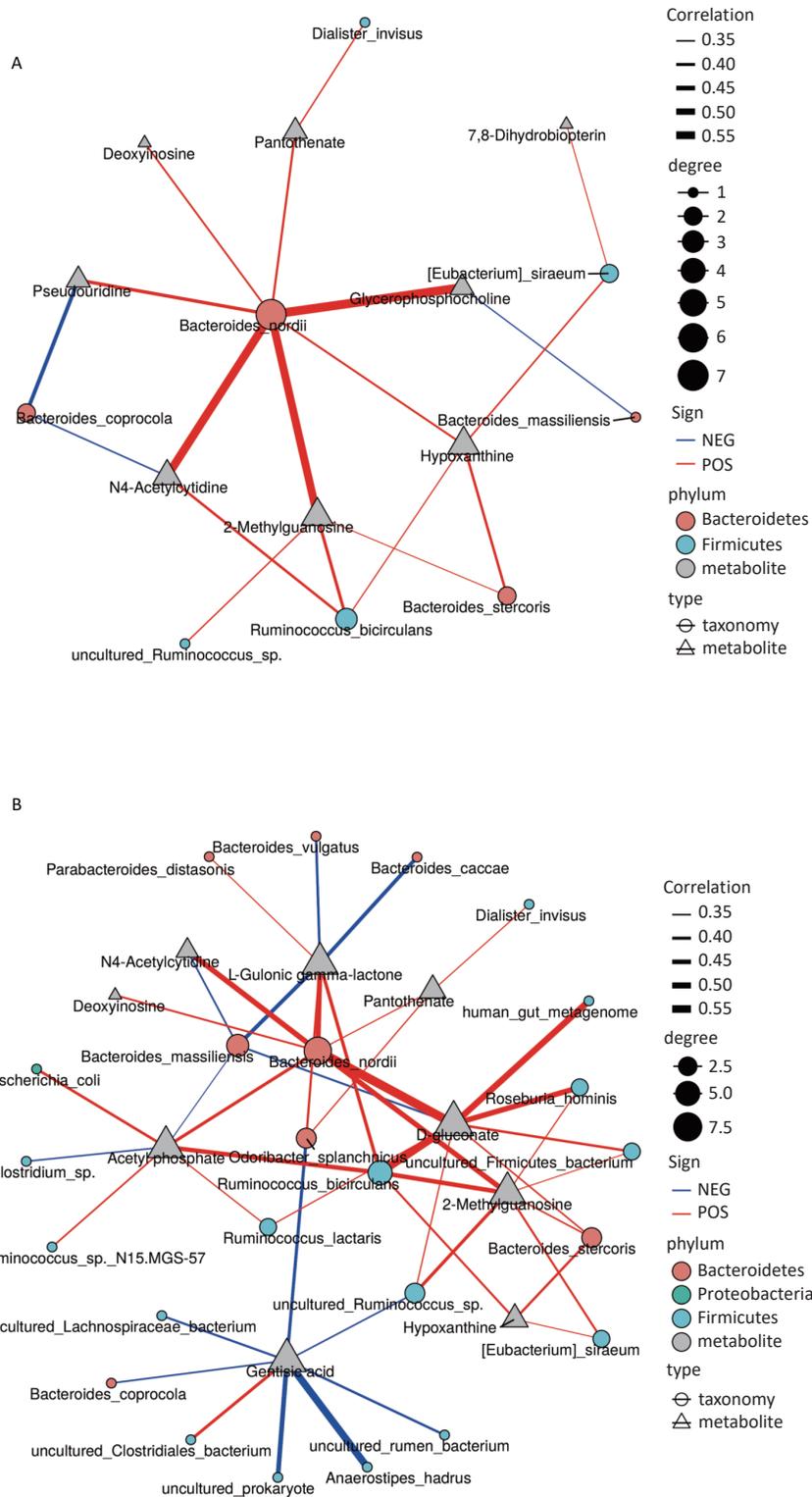
Supplementary Figure S6. Microbiota (class)-metabolites correlation network based on Spearman’s correlation coefficients. Each node represents one genus (circle) or metabolite (triangle), and two nodes are linked if the correlation was significant (two-sided pseudo $P < 0.05$) and more than 0.3. Lines between nodes show positive correlations (red lines) or negative correlations (blue lines). The node size is proportional to the number of correlated metabolites. (A) positive ion mode; (B) negative ion mode.



Supplementary Figure S7. Microbiota (order)-metabolites correlation network based on Spearman’s correlation coefficients. Each node represents one genus (circle) or metabolite (triangle), and two nodes are linked if the correlation was significant (two-sided pseudo $P < 0.05$) and more than 0.3. Lines between nodes show positive correlations (red lines) or negative correlations (blue lines). The node size is proportional to the number of correlated metabolites. (A) positive ion mode; (B) negative ion mode.



Supplementary Figure S8. Microbiota (family)-metabolites correlation network based on Spearman's correlation coefficients. Each node represents one genus (circle) or metabolite (triangle), and two nodes are linked if the correlation was significant (two-sided pseudo $P < 0.05$) and more than 0.3. Lines between nodes show positive correlations (red lines) or negative correlations (blue lines). The node size is proportional to the number of correlated metabolites. (A) positive ion mode; (B) negative ion mode.



Supplementary Figure S9. Microbiota (species)-metabolites correlation network based on Spearman’s correlation coefficients. Each node represents one genus (circle) or metabolite (triangle), and two nodes are linked if the correlation was significant (two-sided pseudo $P < 0.05$) and more than 0.3. Lines between nodes show positive correlations (red lines) or negative correlations (blue lines). The node size is proportional to the number of correlated metabolites. (A) positive ion mode; (B) negative ion mode.