Supporting information

Metabolomics uncovers the regulatory pathway of acyl-homoserine lactones-based quorum sensing in anammox consortia

Xi Tang[1,2], Yongzhao Guo[1,3], Shanshan Wu[1,2], Liming Chen[1,2], Huchun Tao[3], Sitong Liu[*1,2]

[1]Key Laboratory of Water and Sediment Sciences, Ministry of Education of China, Beijing 100871, China

[2]College of Environmental Sciences and Engineering, Peking University, Beijing 100871, China

[3]School of Environment and Energy, Peking University Shenzhen Graduate School, Shenzhen 518055, China

*Corresponding author: Sitong Liu

Address: College of Environmental Science and Engineering, Peking University, Yiheyuan Road, No.5, Haidian District, Beijing 100871, China.

Tel/Fax: 0086-10-62754290.

E-mail: liusitong@pku.edu.cn

Document prepared:

Supporting information: 23 pages, 1 Text, 14 Figures, and 2 Tables

Supporting data: 1 file (excel format)

Supplementary materials and methods

Bioreactor operation

The bacteria were mixed inside the reactors using magnetic stirrers at a stirring speed of 80 rpm. The feeding medium was flushed with N_2/CO_2 (95/5%)[1] to remove dissolved oxygen and the pH was maintained at 6.8–7.0. The reactors were operated in a sequencing mode with cycles of 14 min for filling, 30 min for settling and 5 min for discharging. 0.5 L influent was fed to the reactor during the filling stages and 0.5 L treated effluent was discharged during discharging stages. The settling time of the SBR was to set long enough that the biomass concentration in the effluent could not be detected.

Replicated bioreactor operation

Six identical reactors (the same with reactors described in the main article) were operated for 30 days. The operation strategy and the inoculated anammox consortia of this experiment were also the same as the condition in the main article. The concentrations of ammonium, nitrite, and nitrate in the SBRs were measured every two days using APHA standard methods.

Determination of SAA

Biomass taken from the six SBRs were incubated in triplicate in 100 mL serum bottles. The bottles for the treatment bottles were also incubated with 2 μM AHLs. The initial concentrations of NH_4^+-N and NO_2^--N were maintained at 100 mg-N L^{-1} and the biomass concentration was approximately 0.2 gVSS L^{-1}. The concentrations of NH_4^+-N, NO_2^--N, and NO_3^--N were periodically quantified during the incubation. SAA was estimated from the slope of the curve described by the decrease in substrate concentration during the experiment, which was related to the biomass concentration

Extraction of EPS

The biomass was collected by centrifugation at 8,000 *g* for 10 min, and the pellet was washed three times using 0.1 M NaCl (w/v). The sample was re-suspended in 10 mL of 0.1 M NaCl before adding CER at 70 g $gVSS^{-1}$. The mixture was stirred for 3 h at 200 rpm at 4 °C. The EPS solution was

obtained by centrifugation for 20 min at 9,000 *g* at 4 °C, and filtered through a 0.45 μm acetate cellulose membrane

Detection of AHLs by liquid chromatography coupled with mass spectrometry (LC-MS)

. All samples were applied to LC-MS (Q Exactive orbitrap Thermo, CA). 1uL sample was loaded to Extend-C18 (5 μm, 2.1×150 mm) reverse phase column at a flow rate of 0.2 mL min^{-1}, then the sample was eluted to orbitrap mass spectrometer with 100% ACN as eluent. The data with mass range m/z 80-1200 was acquired in positive ion mode using data dependent MSMS acquisition. The full scan and fragment spectra were collected with the resolution of 70,000 and 17,500 respectively. The source parameters are as follows: spray voltage: 3000v; capillary temperature: 320^{o}C; heater temperature: 300^{o}C; sheath gas flow rate: 35; auxiliary gas flow rate: 10. The identification and quantification of AHLs depended on the monitoring profile of five standard (specific LC retention time, appearance of precursor ion m/z and two transition ions, as well as relative intensity of the two transition ions).

16s rRNA sequencing analysis

DNA extraction was conducted using the Fast DNA Spin Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. For each sample, 16S rRNA was amplified by Polymerase Chain Reaction (PCR) and sequenced using the Illumina platform (Illumina Inc., San Diego, CA, USA). Universal bacterial primers were used for amplification, including the 338F forward primer (5'-ACTCCTACGGGAGGCAGCAG-3') and 806R reverse primer (5'-GGACTACHVGGGTWTCTAAT-3')[2]. PCR cycling conditions were as follows: one cycle of 3 min denaturation at 95 °C; 27 cycles of 30 s at 95 °C, 30 s at 55 °C, and 45 s at 72 °C; final extension for 10 min at 72 °C. Sequencing reads were quality-filtered, de-multiplexed, and clustered into Operational Taxonomic Units (OTUs) using Usearch [3] at identity thresholds of 0.97. Finally, the taxonomic assignment of the 343 OTUs was conducted using the RDP Classifier program[4] (0.7 confidence level) and Silva database. The abundance of bacterial genera in each sample was determined using the classified OTUs.

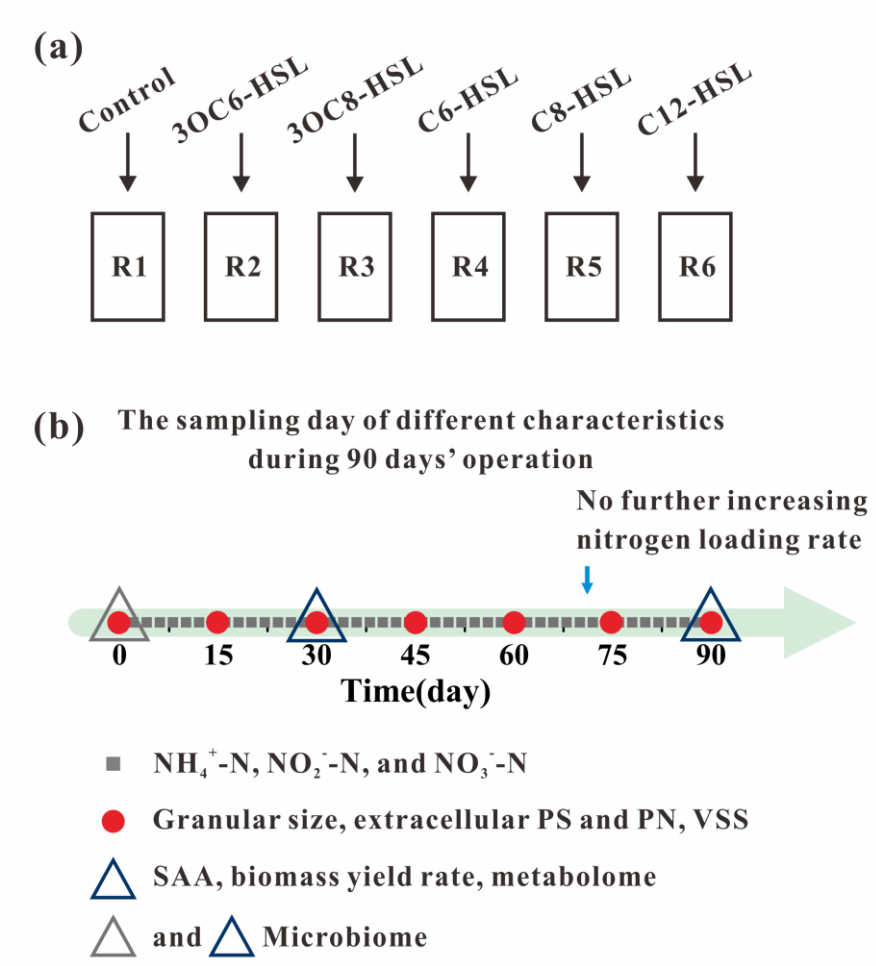
Profiling of metabolites

Mass spectrometry data were collected using a Q Exactive mass spectrometer (Thermo Fisher, Waltham, MA, USA). The extracts were reconstituted in 80% methanol. One aliquot was analyzed in positive-ion mode and the other in negative-ion mode. The sample was applied to a normal phase chromatography column with 50% acetonitrile containing 10 mM ammonium formate as the eluant. The data collected over the mass range m/z 70–1050 were acquired using data-dependent mass spectrometry (MS)/MS acquisition. The full scan and fragment spectra were collected at a resolution of 70,000 and 17,500, respectively. The source parameters were as follows: spray voltage: 3000 V; capillary temperature: 320 °C; heater temperature: 300 °C. A Trace finder search with a home-built database containing 740 compounds was used to identify metabolites. Overrepresented pathways were found by assigning pathways identified from metabolites according to the KEGG pathways of the metabolites**.**

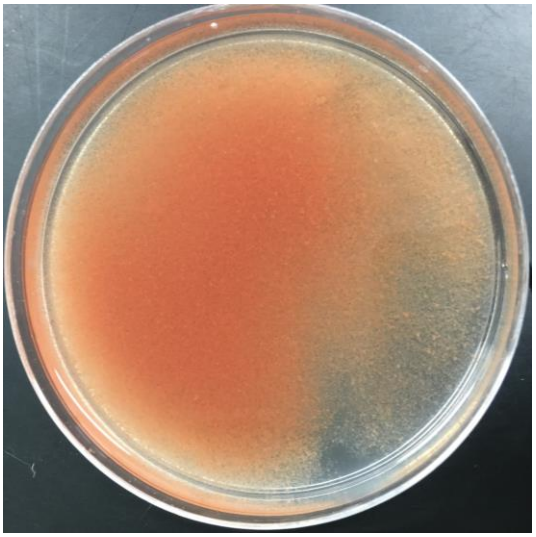
Text S1

Replicated bioreactor operation

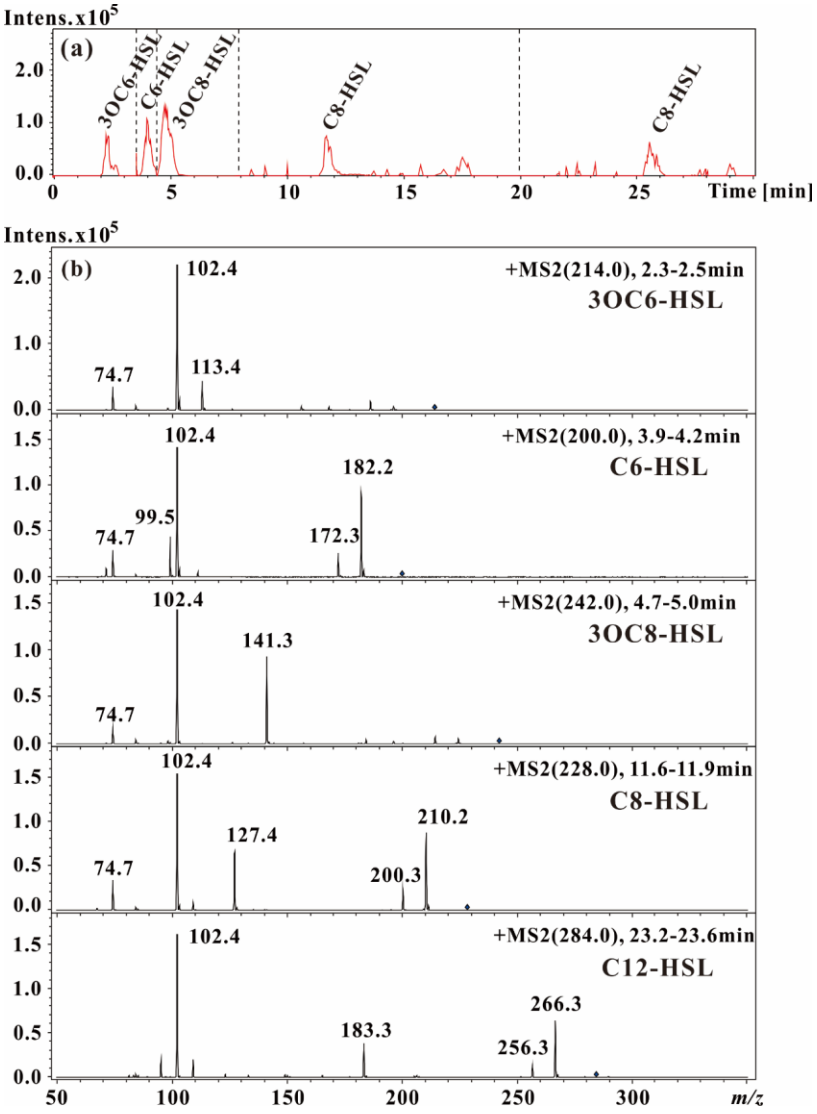
The trend of fold-changes in the NRRs of all treatment reactors compared to that of the control without AHLs adding were similar to the results in the main article. The addition of AHLs except for C12-HSL did not significantly affect the NRRs during phase I. Although the increased fold-changes was little different from that in the main article, the addition of 3OC6-HSL, C6-HSL and C8-HSL still significantly increased the NRRs compared to the control during phase II.



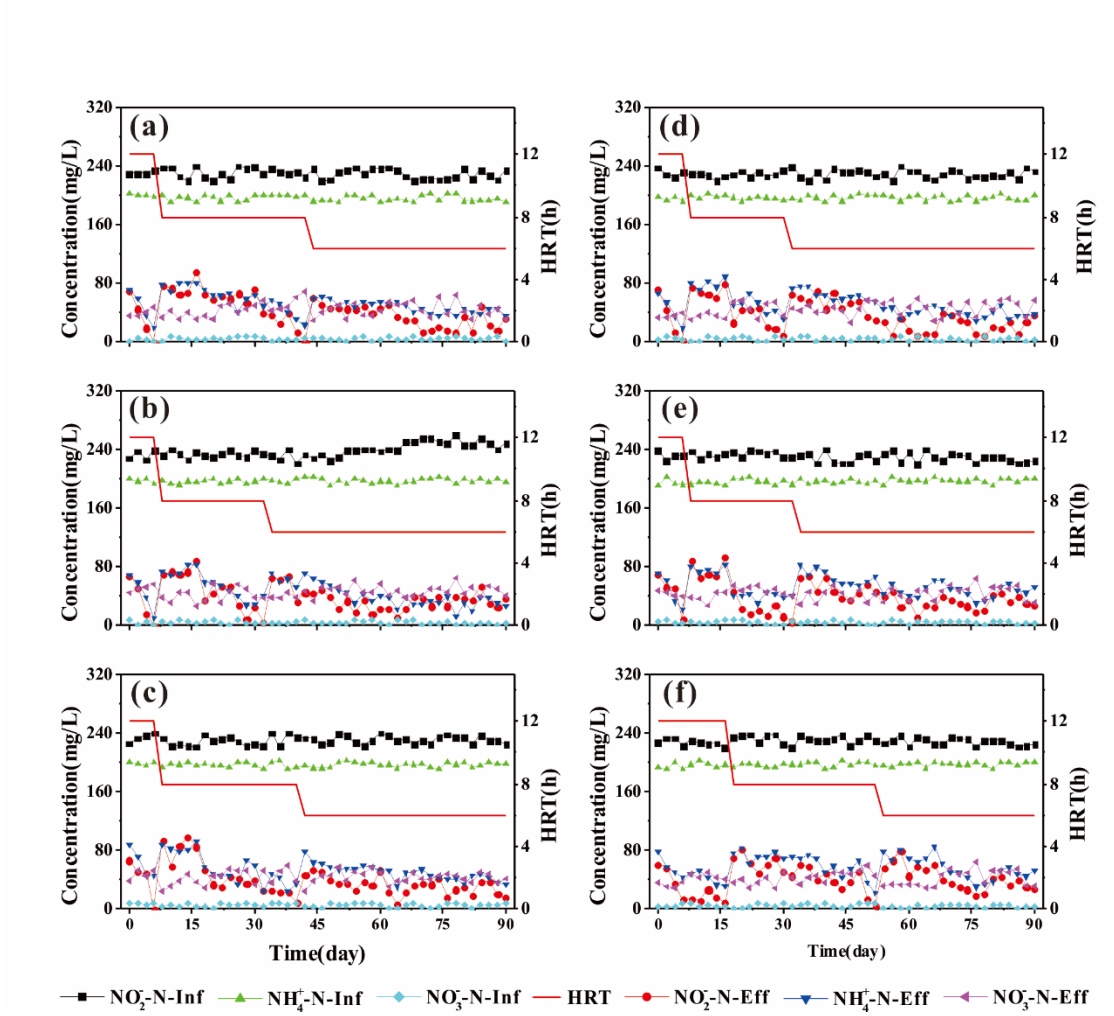
Supplementary Figure S1 Overview of the experimental operation on six identical reactors (a) and the sampling day for different characteristics (b), including nitrogen concentration, flocs size, extracellular PS (Polysaccharide) and PN (Protein), VSS (Volatile Suspended Solids), SAA (Specific Anammox Activities), biomass yield rate, metabolome, microbiome.



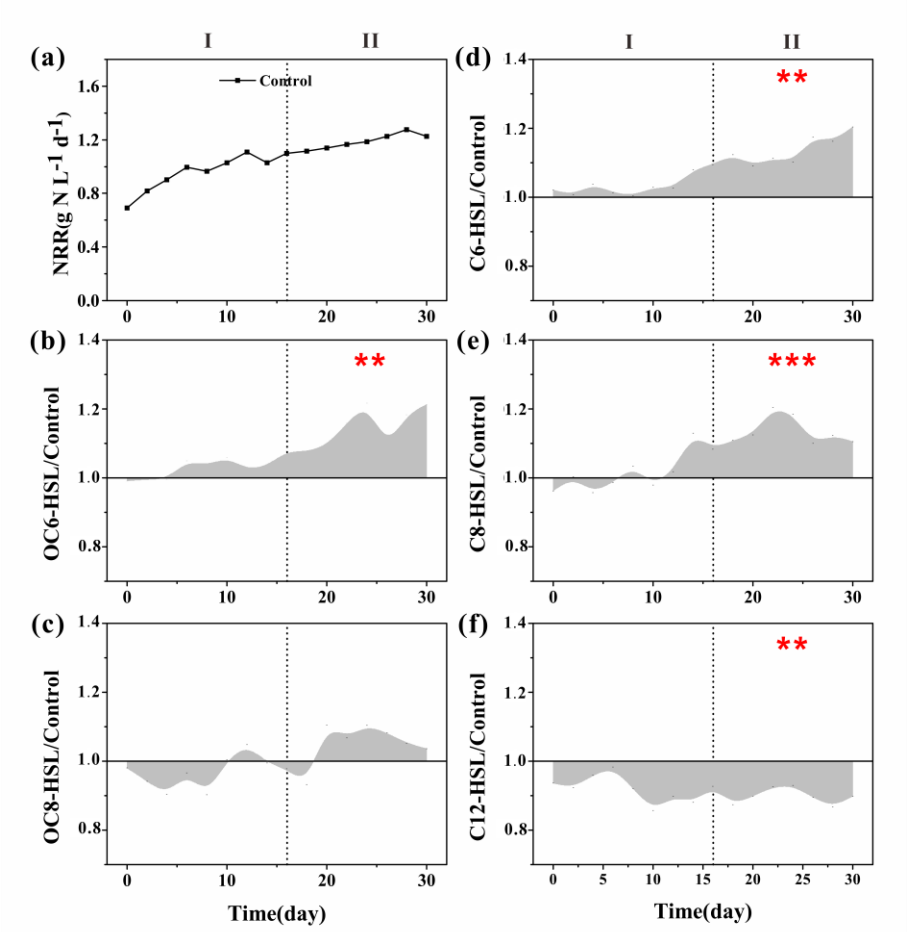
Supplementary Figure S2 Picture of the seeding sludge.



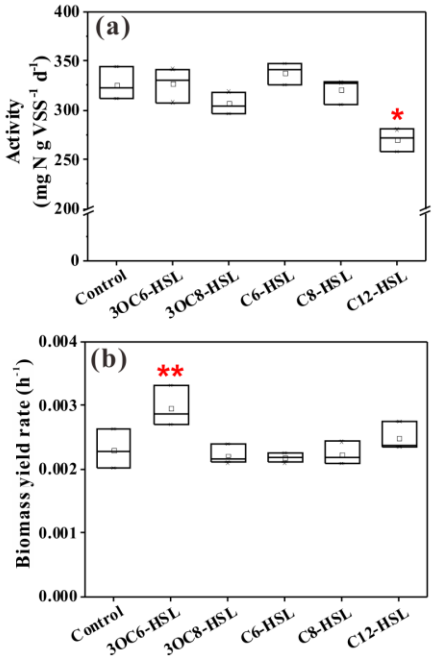
Supplementary Figure S3 LC-MS profiling of AHLs extracted from the anammox consortia supernatant. MRM mode was used to sufficiently identity five AHLs in complex supernatant extracts. In Chromatogram (a), isolated peaks eluted at 2.30, 4.01, 4.75, 11.61, and 25.27 min respectively, which were the same as the standards (data not show). Isolated peak was subjected to fragmentation to obtain a MS/MS spectrum (b). The characteristic ion fragment of the lactone ring (m/z=102) was verified in all MS/MS spectrum, and relative ion fragments of the acyl side chain were also identified to confirm the existence of specific AHLs.



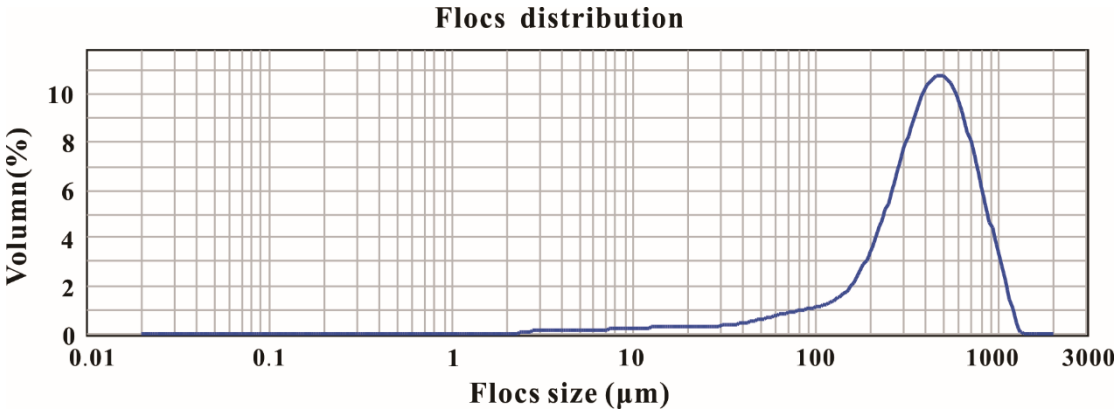
Supplementary Figure S4 Profile of influent and effluent nitrogen concentrations during the 90days' operation in six reactors. The control reactor (a) and the treatment reactor with addition of 2 μM 3OC6-HSL, 3OC8-HSL, C6-HSL, C8-HSL and C12-HSL respectively (b-f) improved NRR performance by progressively shorting HRT from 12 h to 6 h.



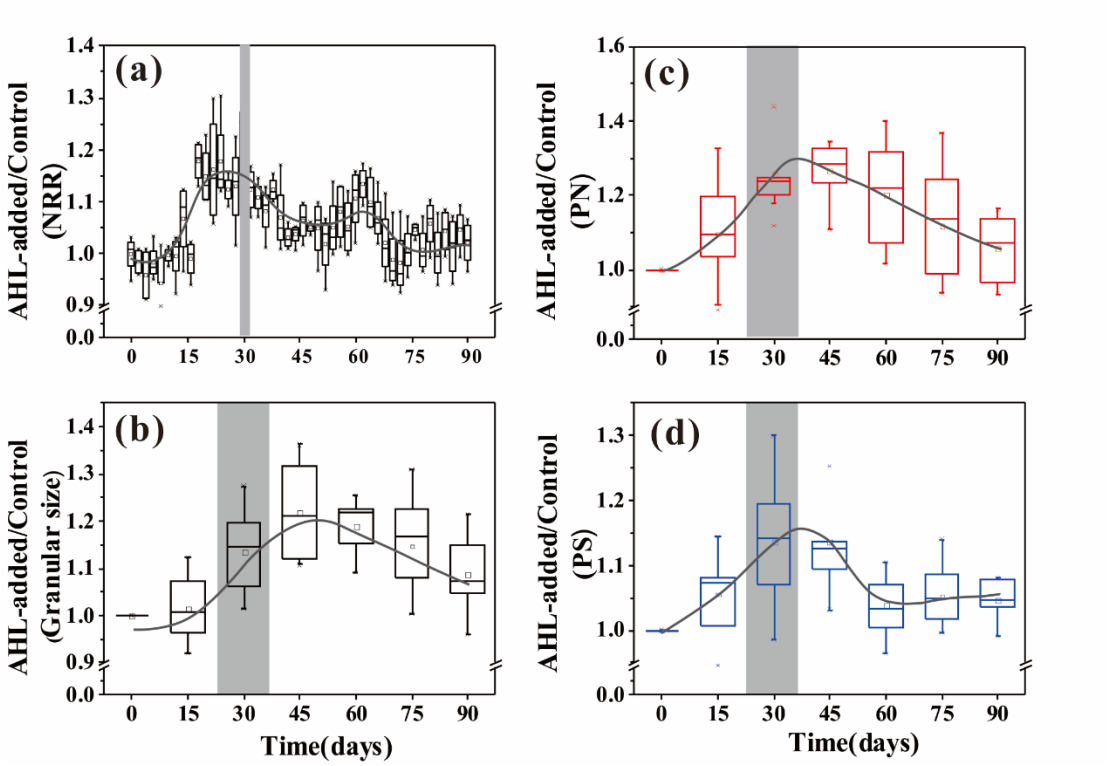
Supplementary Figure S5 Effects of exogenous 3OC6-HSL, 3OC8-HSL, C6-HSL, C8-HSL and C12-HSL on the nitrogen removal rate (NRR). The NRR profile of the control reactor during the 30 days operation is presented in (a). The fold-changes of the NRRs between treatment reactor and the control are illustrated in (b-f). Two distinct phases of NRR performance are separated by the dotted line (phases I, II). One-way ANOVA with post hoc test by Dunnett's multiple-comparison was conducted to compare the treatment reactors to the control reactor, where significant differences are indicated as follows: * $p<0.05$, ** $p<0.01$.



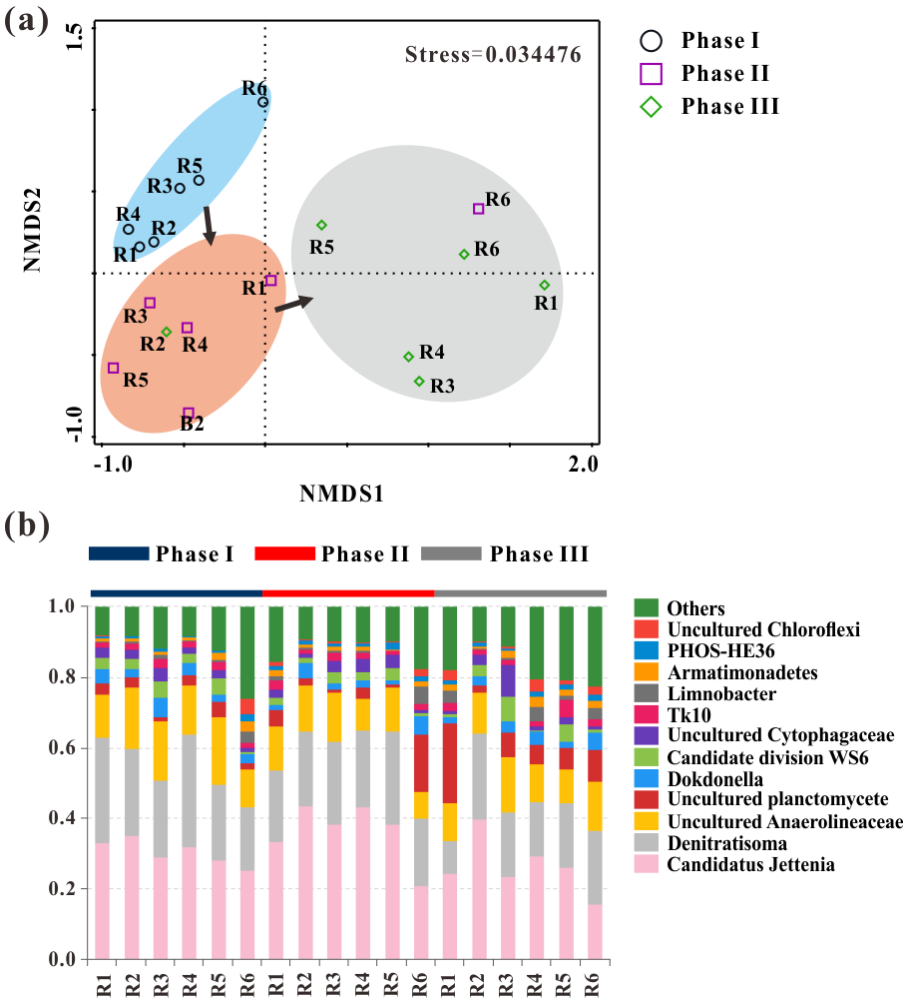
Supplementary Figure S6 Effects of exogenous 3OC6-HSL, 3OC8-HSL, C6-HSL, C8-HSL and C12-HSL on the activity (a) and biomass yield rate (b) of anammox consortia during phase III. Significant differences were conducted by One-way ANOVA with post hoc test by Dunnett's multiple-comparison between the treatment reactors and the control (*$p < 0.05$ and **$p < 0.001$).



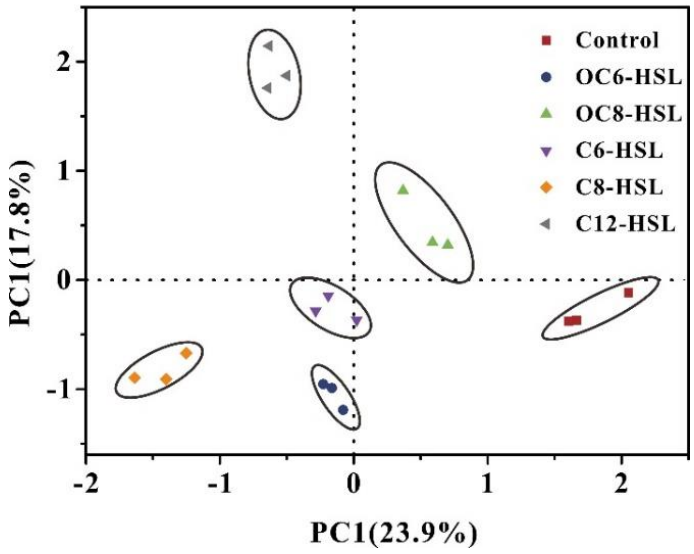
Supplementary Figure S7 Flocs distribution of sample obtained on the day 30 of reactor operation



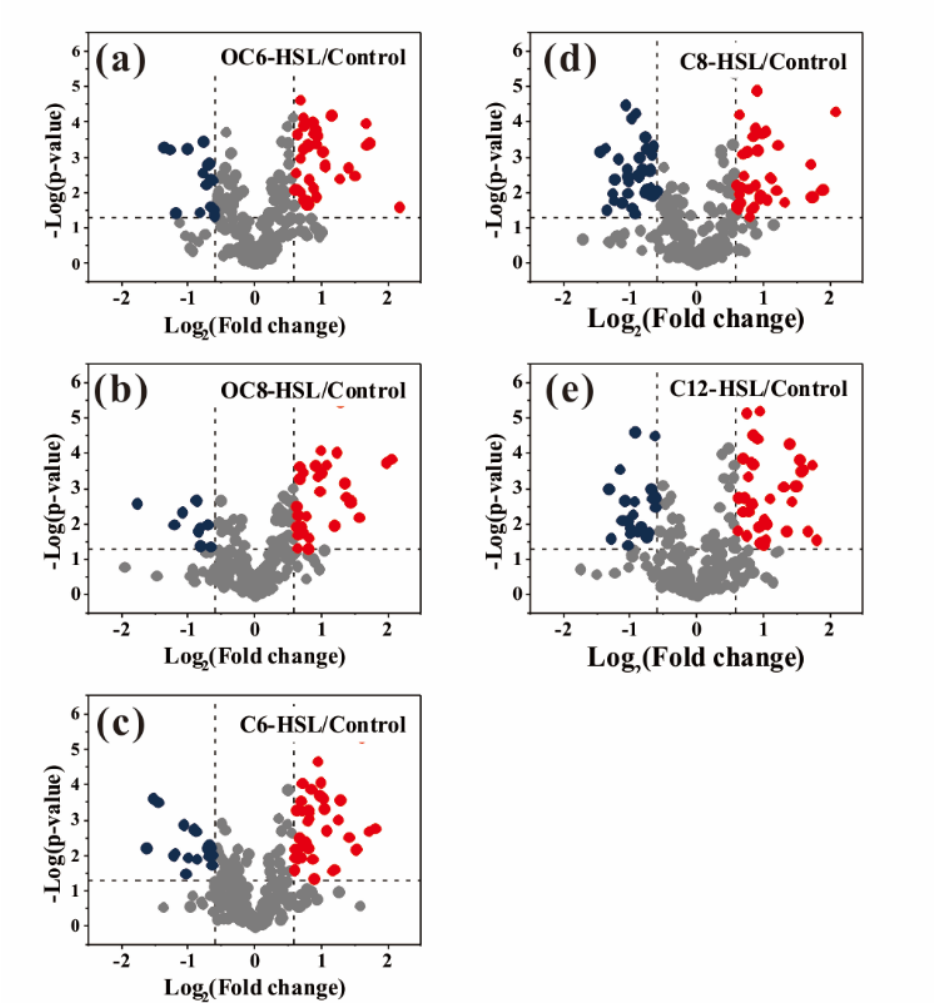
Supplementary Figure S8 Fold-changes in NRR (a), granular size (b), PN (c) and PS (d) content between the reactors dosing AHLs and the control during the 90 days' operation. Comparisons of these phenotypic data between the reactors dosing AHLs and the control at each sampling time are described as box plots. At around day 30 of reactors operation, NRR, granular size, PN and PS of the he reactors dosing AHLs were remarkably different from those of the control.



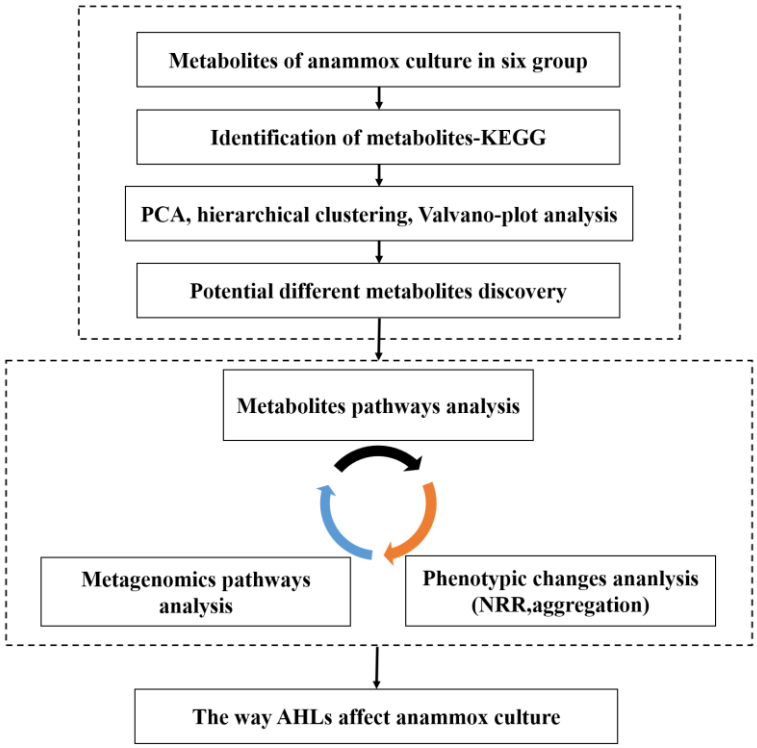
Supplementary Figure S9 Reactor operation time drives microbiological community perturbance. NMDS analysis showed the community differences and similarities among all 18 samples (six reactors (control (R1), 3OC6-HSL (R2), 3OC6-HSL (R3), C6-HSL (R4), C8-HSL (R5), C12-HSL (R6)) during three different phases (I, II, III)) (a). Different phases are labeled with different color as indicated. Arrows illustrate the shift in the microbial community during the three phases. In order to obtain a more accurate composition of the microbial community, relative abundances were depicted in a 100% stacked column (b). The most abundant genera (relative abundance >1%) were selected.



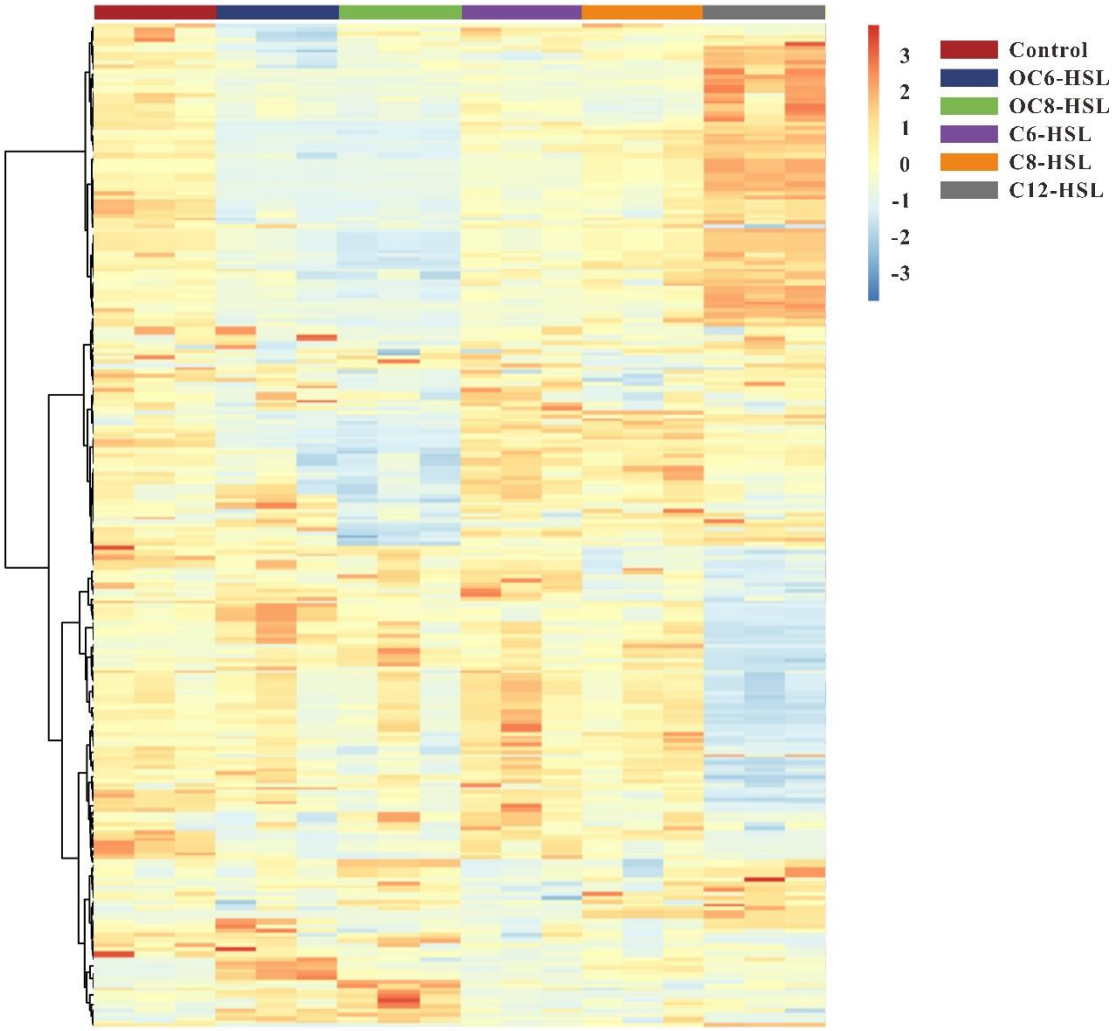
Supplementary Figure S10 Exogenous AHLs drive the perturbation of metabolites. Principal Component Analysis (PCA) scatter plots derived from anammox consortia metabolites of the six reactors during phase II. Biomass samples for metabolite analysis were collected in triplicate. Six groups (control, 3OC6-HSL, 3OC8-HSL, C6-HSL, C8-HSLand C12-HSL) are completely separated into six clusters.



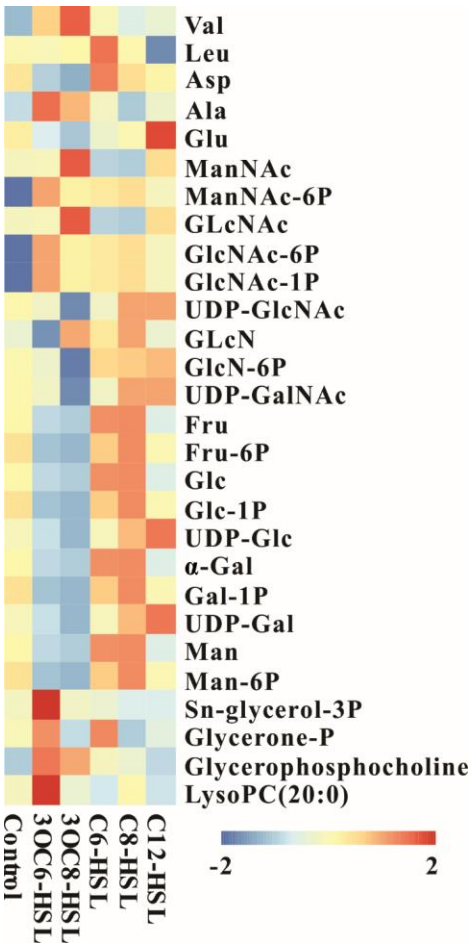
Supplementary Figure S11 Metabolic changes induced by exogenous AHLs. Volcano plots with log_2 (fold-change) on the x axis and –log (p-value) on the y axis were generated. Changes in metabolites between the consortia dosing AHLs (3OC6-HSL, 3OC8-HSL, C6-HSL, C8-HSL, and C12-HSL) and the control were conducted in (a-e). Metabolites that increased by at least 1.5-fold and had $p<0.05$ are shown in red, whereas those that decreased by at least 1.5-fold and had a $p < 0.05$ are shown in blue.



Supplementary Figure S12 Overview of this study to identify the metabolic pathway induced by AHLs in anammox consortia. Untargeted metabolomics technique was performed to identify metabolites of anammox consortia. Statistical and bioinformatics methods were used to identify significantly different metabolites. Analyses of metabolic pathways, metagenomics pathways and relative phenotypes were then applied. The annotation of metagenomic pathways was analysis using the metagenomic data of the same anammox consortia, which had been deposited to the Metagenomics RAST (MG- RAST) server with accession numbers **mgm4713028.3 and mgm4713025.3**. The genes with KEGG annotation were assigned into KEGG pathways. Finally, the regulatory pathway of AHLs in anammox consortia was discovered.



Supplementary Figure S13 Hierarchical clustering analysis of metabolites in samples from the control and the consortia with the addition of 3OC6-HSL, 3OC8-HSL, C6-HSL, C8-HSL and C12-HSL during phase III. Clustering was performed using Pearson correlation as the distance metric. The cluster tree shows how the samples and metabolites divide. Across the top are the samples, labeled by their group in different color, and along the side are the metabolites.



Supplementary Figure S14 Heatmap of metabolites related to the growth and extracellular PN of anammox consortia in the control and the consortia with the addition of 3OC6-HSL, 3OC8-HSL, C6-HSL, C8-HSL and C12-HSL during phase III.

Supplementary Table S1 The content of extracellular PN and intracellular amino acid used in multiple linear regression model from 15 different samples of anammox consortia.

Supplementary Table S2 The result of multiple linear regression using the content of extracellular PN and intracellular amino acid.

Adjuster R^2	p-value of the model		p-value of the variables	Standardized Coefficients
0.850	0.000	Aspartic	.000	0.540
		Valine	.001	0.492
		Glutamic acid	.003	0.411

Supplementary Table S3 Abbreviations for metabolites related to NRR and aggregation.

Metabolites	Abbreviations
2-Isopropylmalic acid	2-Isopropylmalic acid
Valine	Val
Leucine	Leu
Alanine	Ala
Aspartate	Asp
(S)-Malate	(S)-Malate
Fumarate	Fumarate
Succinate	Succinate
2-Oxo-glutarate	2-Oxo-glutarate
cis-Aconitinate	cis-Aconitinate
Glutamate	Glu
Citrate	Citrate
Oxoloacetate	Oxoloacetate
N-Acetylmannosamine	ManNAc
N-Acetyl-D-mannosamine 6-phosphate	ManNAc-6P
N-Acetyl-D-glucosamine	GLcNAc
N-Acetyl-D-Glucosamine 6-Phosphate	GlcNAc-6P
N-Acetyl-alpha-D-glucosamine 1-phosphate	GlcNAc-1P
Uridine diphosphate-N-acetylglucosamine	UDP-GlcNAc
Glucosamine	GLcN
Glucosamine 6-phosphate	GlcN-6P
Uridine diphosphate-N-acetylgalactosamine	UDP-GalNAc
Fructose	Fru
Fructose 6-phosphate	Fru-6P
ADP-glucose	ADP-Glc

UDP-D-Xylose	UDP-D-Xyl
D-Glucose	Glc
Glucose 1-phosphate	Glc-1P
Uridine diphosphate glucose	UDP-Glc
D-Galactose	α-Gal
Galactose 1-phosphate	Gal-1P
Uridine diphosphategalactose	UDP-Gal
D-Mannose	Man
Mannose 6-phosphate	Man-6P
sn-Glycerol 3-phosphate	Sn-glycerol-3P
Glycerone phosphate	Glycerone-P
Glycerophosphocholine	Glycerophosphocholine
Lysophosphatidylcholine(20:0)	LysoPC(20:0)
Nicotinamide adenine dinucleotide	NAD

Supplementary references

(1) van Dongen, U.; Jetten, M. S. M.; Van Loosdrecht, M. C. M. The SHARON-Anammox process for treatment of ammonium rich wastewater. *Water Sci. Technol.* **2001**, *44* (1), 153–160.

(2) Srinivasan, S.; Hoffman, N. G.; Morgan, M. T.; Matsen, F. A.; Fiedler, T. L.; Hall, R. W.; Ross, F. J.; McCoy, C. O.; Bumgarner, R.; Marrazzo, J. M. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. *PLoS One* **2012**, *7* (6), e37818.

(3) Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26* (19), 2460–2461.

(4) Cole, J. R.; Chai, B.; Farris, R. J.; Wang, Q.; Kulam-Syed-Mohideen, A. S.; McGarrell, D. M.; Bandela, A. M.; Cardenas, E.; Garrity, G. M.; Tiedje, J. M. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Res.* **2007**, *35* (suppl 1), D169–D172.