Supporting Information

Molecular Characteristics of Dissolved Organic Nitrogen and Its Interaction with Microbial Communities in Pre-chlorinated Raw Water Distribution System

Hang Xu, [†,#] *Chenshuo Lin,* [*,†,#] *Zhen Shen,* [†,#] *Li Gao,*[§] *Tao Lin,* [*,†,#] *Hui Tao,* [†,#] *Wei Chen,* [†,#] *Jian Luo,* [⊥] *and Chunhui Lu* [‡]

† Key Laboratory of Integrated Regulation and Resource Development on Shallow Lake of Ministry of Education, College of Environment, Hohai University, Nanjing 210098, China

#Hohai University, College of Environment, Nanjing 210098, China

§Future Water Strategy Group, South East Water, PO BOX 2268, Seaford, Victoria 3198, Australia

‡State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China
⊥School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
Chenshuo Lin and Hang Xu contributed equally to this work.
*Corresponding Author (Chenshuo Lin e-mail: Lincs23@126.com)
(Tao Lin e-mail: hit_lintao@163.com)

Number of pages: 11

Number of tables: 2

Number of figures: 5

Supporting Materials and Methods

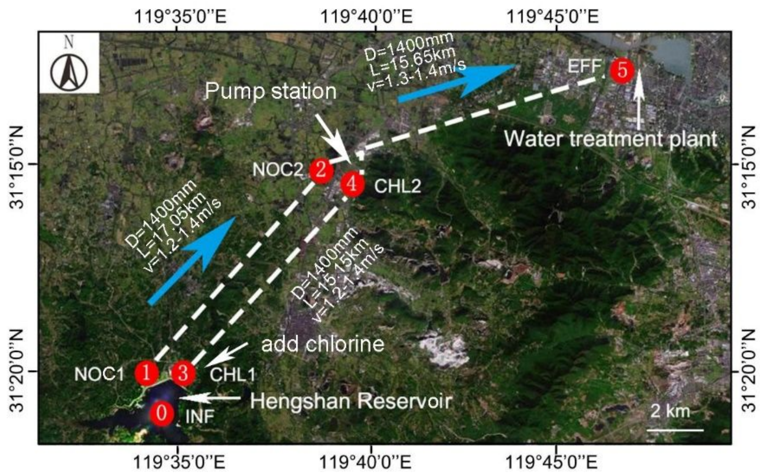


Figure S1. Locations of sampling sites in Yixing, China: Hengshan Reservoir (INF), inlet and outlet of chlorinated pipe (CHL1 and CHL2, respectively), inlet and outlet of unchlorinated pipe (NOC1 and NOC2, respectively), and combined effluent (EFF).

S1. Sampling process

DOM was extracted by solid-phase extraction (SPE) using the Bond Elut PPL SPE cartridges (500mg, 6ml; Agilent, USA), based on the method described by Dittmar et al.[1] These SPE cartridges were firstly rinsed and activated using methanol and 0.01mol L^{-1} formic acid. Then the water sample (100 mL) was adjusted to pH 2 by 0.1mol L^{-1} formic acid. The water sample passed through the cartridges with the flowrate of 5 mL min^{-1}. Afterwards, the cartridges were rinsed with 50 mL acidified LC-MS water to remove the interference of inorganic salts on FT-ICR-MS analysis. Analytes were subsequently eluted with methanol (1mL) and were stocked at 4°C until further analysis. Three DOM replicates were prepared for each site and mixed to one composite sample to mitigate artifacts.

High resolution mass spectrometry analysis was performed using a Apex QE 12 T (Bruker) equipped with an electrospray ionization (ESI, Bruker) source set in negative mode of ionization. Before FT-ICR-MS operation, formic acid (~98%) was used to adjust the samples' pH and rinse the SPE cartridges. After an internal calibration, the mass measurement accuracy was typically within 0.2 p.p.m. for singly charged ions across a broad m/z range (200–700m/z), which can be used to trace the molecular fate [2].

REFERENCES

(1) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnol. Oceanogr.: Methods 2008, 6, 230−235.

(2) Maizel, A.C.; Remucal, C.K. The effect of advanced secondary municipal wastewater treatment on the molecular composition of dissolved organic matter. Water Res. 2017, 122, 42–52.

S2. 16S rRNA sequencing

The PCR reactions for V4-V5 region of the 16S rRNA gene were carried out with primers 515F (5′ -GTGCCAGCMGCCGCGG-3′) and (5′ -CCGTCAATTCMTTTRAGTTT-3′) [3] and were performed in triplicate 20 μL mixture containing 4 μL of 5×FastPfu Buffer, 2 μL of 2.5 mM dNTPs, 0.8 μL of each primer (5 μM), 0.4 μL of FastPfu polymerase, and 10 ng of template DNA by PCR (94 °C for 3 min, 25 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s, and finishing with a final extension at 72 °C for 5 min). Then electrophoresis was used on 2% agarose gel and amplicons were purified with Agencourt AMPure XP beads (Beckman, USA) for pyrosequencing on an Illumina MiSeq platform (Shanghai Sangon).

REFERENCES

(3) Caporaso, J. G.; Lauber, C. L,; Walters, W. A.; Berg-Lyons, D.; Lozupone, C. A.; Turnbaugh, P.J.; Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A. 2011, 108:4516–22.

Table S1. DON molecular formula list for each sample in PRWDSs

Table S2. Physiochemical parameters in PRWDSs

		Temperature (℃)	DO (mg/L)	Turbidity (NTU)	TOC (mg/L)	pH	Free chlorine (mg/L)	NH_4^+-N (mg/L)	NO_2^--N (mg/L)	NO_3^--N (mg/L)
	April	18.3-20.4	8.43-9.25	4.02-4.21	2.88-3.55	7.56-8.33		0.168-0.208	0.008-0.019	1.372-1.403
INF	June	21.6-24.9	6.38-7.11	2.82-3.15	2.47-3.33	7.48-8.04	0	0.157-0.194	0.009-0.018	1.130-1.151
	August	25.8-28.7	5.03-5.27	2.00-2.27	1.80-2.05	7.72-7.97		0.151-0.197	0.006-0.015	1.269-1.313

	April	19.1-20.6	7.75-8.51	3.41-3.86	2.75-3.01	7.71-8.45		0.140-0.175	0.007-0.013	1.501-1.629
NOC1	June	20.5-23.3	5.83-6.55	2.71-3.02	2.17-2.53	7.66-8.03	0	0.144-0.170	0.007-0.012	1.523-1.642
	August	26.1-28.4	4.91-5.14	1.80-2.15	1.97-2.08	7.82-8.79		0.135-0.163	0.006-0.013	1.558-1.673
	April	18.3-19.8	6.81-7.89	2.89-3.36	2.48-2.75	8.37-8.87		0.129-0.151	0.004-0.006	1.717-1.769
NOC2	June	19.5-23.7	5.64-6.01	2.45-2.77	2.08-2.35	8.04-8.55	0	0.132-0.152	0.005-0.007	1.788-1.812
	August	25.7-27.1	4.26-4.73	1.25-1.61	1.88-2.15	7.82-8.43		0.123-0.144	0.003-0.005	1.697-1.831
	April	19.3-21.6	7.88-8.31	2.85-3.13	2.93-3.22	8.03-8.51	0.85-1.01	0.139-0.170	0.010-0.015	1.298-1.322
CHL1	June	20.8-23.1	5.69-6.33	2.42-2.59	2.80-3.10	7.86-8.24	0.82-0.98	0.132-0.164	0.012-0.014	1.198-1.240
	August	25.0-27.9	4.49-4.87	1.93-2.13	2.33-2.52	7.90-8.76	0.84-0.93	0.125-0.167	0.009-0.015	1.266-1.384
	April	18.1-20.2	6.63-7.47	2.92-3.30	3.01-3.36	7.79-8.39	0.57-0.78	0.129-0.170	0.006-0.011	1.347-1.469
CHL2	June	19.0-22.5	5.29-5.77	2.50-2.72	2.62-3.04	7.87-8.70	0.51-0.69	0.131-0.163	0.005-0.012	1.496-1.531
	August	24.1-26.7	4.12-4.53	1.83-2.06	2.47-2.74	7.54-8.56	0.53-0.66	0.123-0.164	0.004-0.009	1.508-1.646
	April	17.5-18.8	5.24-5.45	2.88-3.48	3.14-3.50	7.88-8.26	0.37-0.62	0.112-0.128	0.003-0.006	1.847-1.901
EFF	June	19.4-21.9	3.68-3.91	2.39-2.55	2.56-2.91	8.02-8.43	0.33-0.52	0.115-0.132	0.002-0.004	1.761-1.809
	August	24.3-26.7	2.83-3.04	1.66-2.00	2.35-2.58	7.96-8.32	0.30-0.57	0.096-0.120	0.002-0.005	1.781-1.827

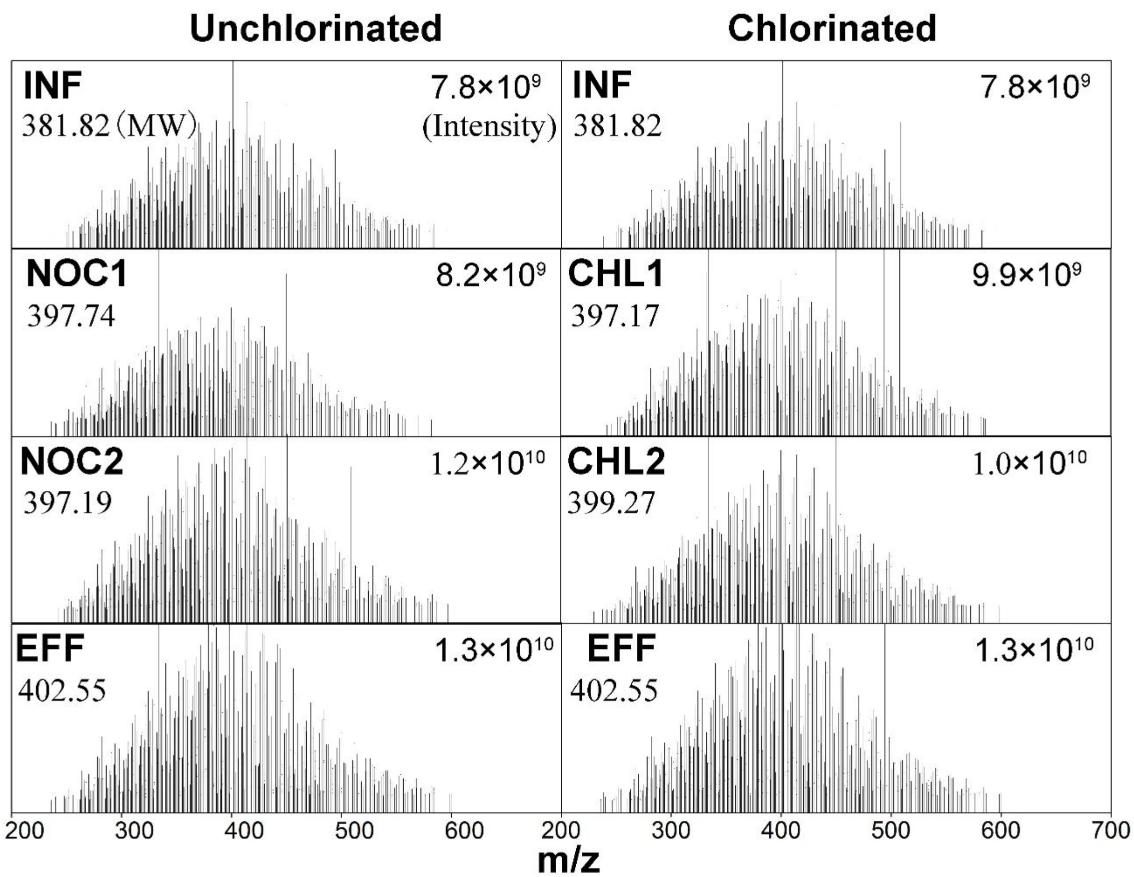
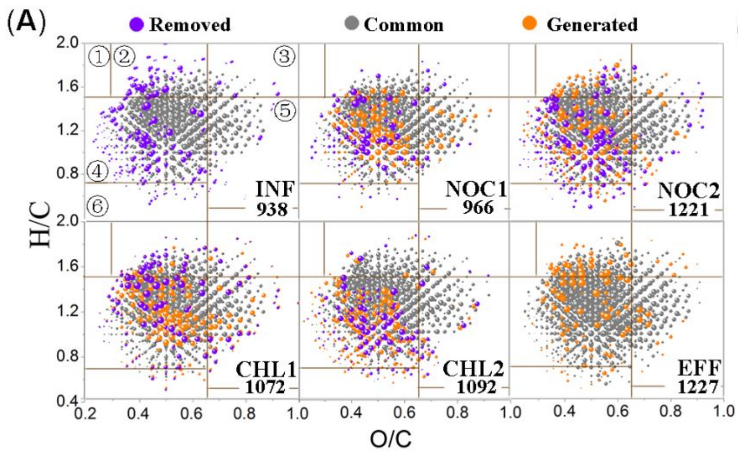


Figure S2. Mass spectra of the PRWDS water samples from INF, NOC1, NOC2, CHL1, CHL2 and EFF. The average MW and intensity of LMW-DON were showed under and on the right of the sample name, respectively.



(B)

Removed	INF	NOC1	NOC2	CHL1	CHL2	EFF
②proteins	14	C	6	9	4	
③carbohydrates,	57	10	10	31	4	
④lignin-like	176	70	120	76	95	
⑤tannin	16	3	16	19	10	
⑥condensed aromatic	15	4	23	11	12	
Generated	INF	NOC1	NOC2	CHL1	CHL2	EFF
②proteins		2	9	9	3	5
③carbohydrates,		13	37	26	7	23
④lignin-like		153	214	162	140	78
⑤tannin		13	30	32	11	10
⑥condensed aromatic		5	31	15	9	9

Figure S3. (A) Van Krevelen diagrams of the DON masses removed (Unique molecules compared to the subsequent sample are in purple), in common (Common molecules compared to both the subsequent and former samples are in grey) and formed (Unique molecules compared to the former sample are in orange) compared to the corresponding influent (①lipids, ②proteins, ③carbohydrates, ④lignin-like, ⑤tannin, and ⑥condensed aromatic, respectively). Bubble sizes of the dots represent the m/z of each formula. (B) Summary of removed (Unique molecules compared to the subsequent sample are in purple), and generated DON molecular formulas (Unique molecules compared to the former sample are in orange) sorted by the different classes. Lipids were not detected and measured in any samples.

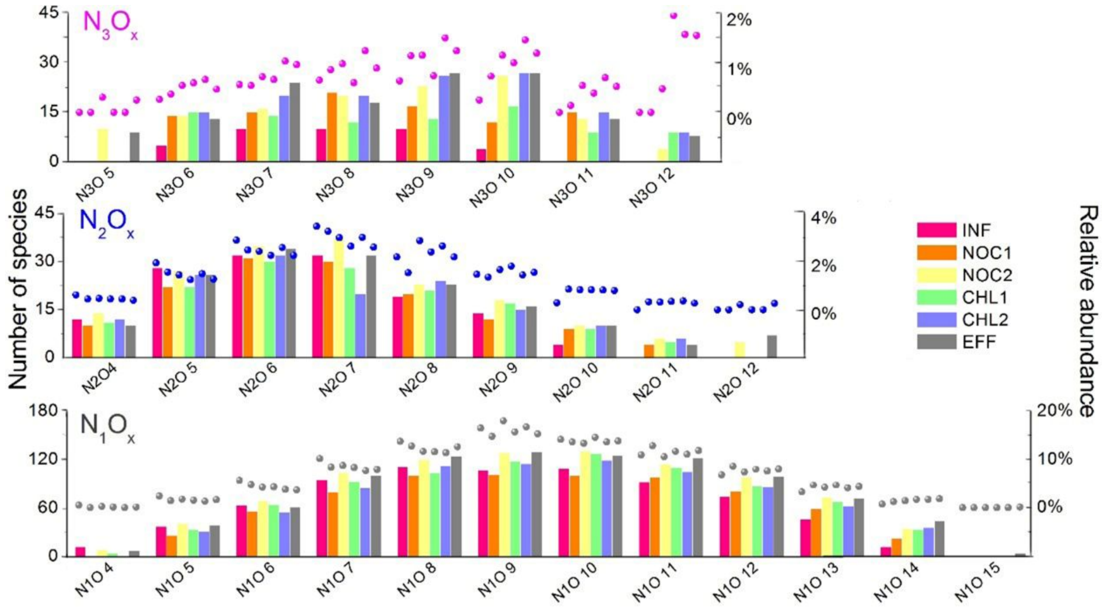


Figure S4. Classification of DON species into different subgroups according to the numbers of N and O atoms in their molecules in PRWDS. The histogram indicates the number of molecular formulae of each subgroup. And the scatter plots with right y-axis show the relative abundance of each subgroup based on the total intensity of all samples. The N_1O_p, N_2O_p, and N_3O_p indicated one, two, and three N atoms in the DON molecule, respectively.

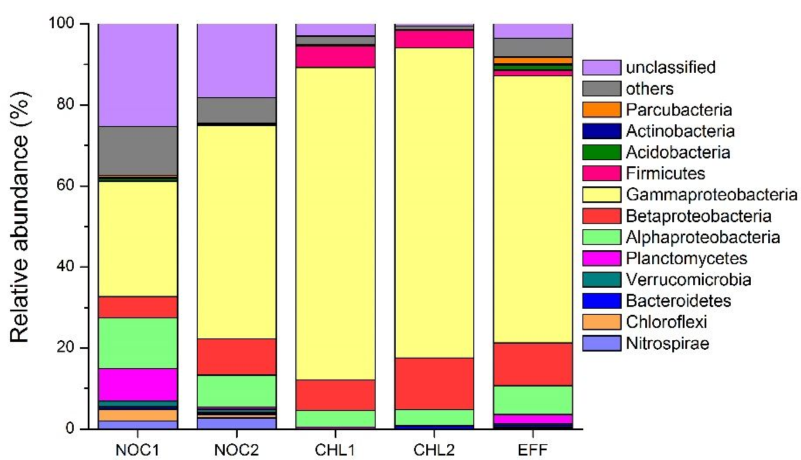


Figure S5. Relative abundance of the most abundant phyla from biofilm samples in PRWDS

Table S3. Bacterial abundance and diversity of biofilm samples in the PRWDS

Sample	Seq_num	OTU_num	Shannon	ACE	Chao1	Coverage	Simpson
NOC1-Apr.	58280	3484	7.590	3826.756	3503.886	0.933	0.012
NOC1-Jun.	59852	3502	7.597	3900.253	3521.932	0.923	0.010
NOC1-Aug.	62037	3566	7.602	3963.920	3572.451	0.917	0.009
NOC2-Apr.	54173	2252	5.806	2433.935	2681.981	0.994	0.016
NOC2-Jun.	56491	2304	5.812	2566.125	2753.395	0.992	0.015
NOC2-Aug.	58064	2330	5.814	2701.047	2953.943	0.912	0.014
CHL1-Apr.	41528	899	3.222	1337.773	1352.042	0.990	0.075
CHL1-Jun.	42155	905	3.229	1455.284	1469.375	0.928	0.072
CHL1-Aug.	44687	924	3.231	1503.354	1521.240	0.944	0.070
CHL2-Apr.	39904	948	3.314	1578.828	1557.412	0.949	0.050
CHL2-Jun.	40689	991	3.317	1632.944	1651.062	0.956	0.048
CHL2-Aug.	43008	1015	3.322	1679.903	1677.284	0.937	0.047
EFF-Apr.	42604	1306	4.087	1773.792	1738.205	0.997	0.041

EFF-Jun.	42978	1332	4.092	1794.329	1745.974	0.935	0.038
EFF-Aug.	45088	1380	4.102	1812.046	1798.095	0.963	0.036

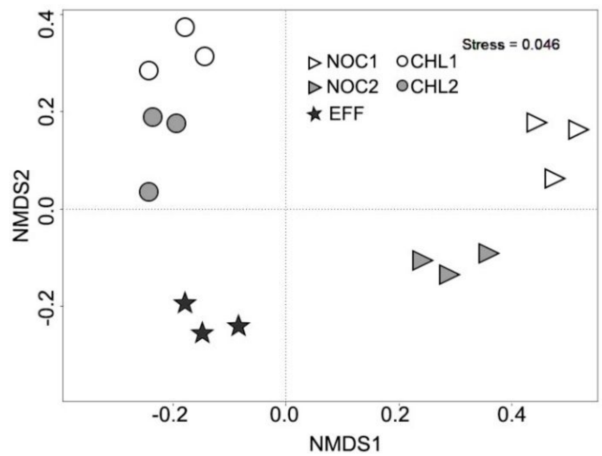


Figure S6. NMDS analysis of BMC based on the OTUs detected from biofilm samples in PRWDS.

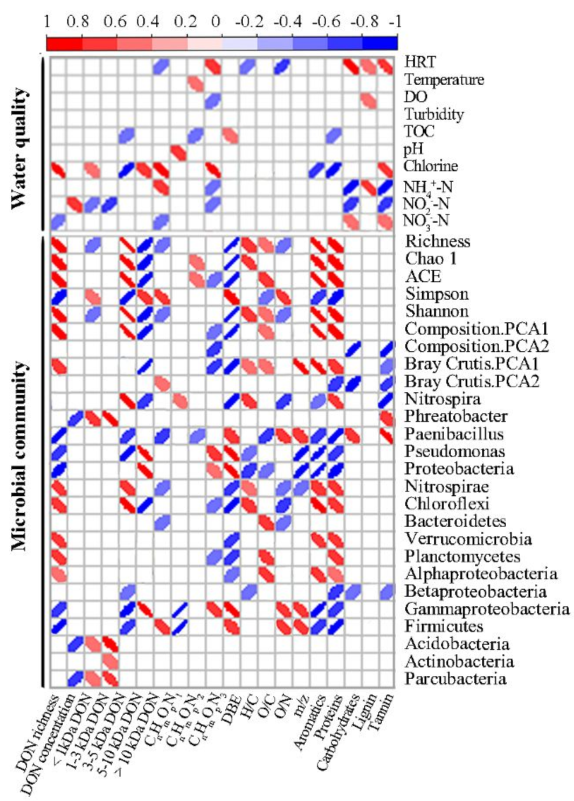


Figure S7. Spearman correlations coefficients between DON characteristics, water quality along the pipe and microbial properties in the PRWDS with $p < 0.05$.