Supporting Information for

Spectroscopic and molecular-level characteristics of dissolved organic matter in a highly polluted urban river in South China

Quanhui Ye[1,2], Zi-Ting Zhang[1,2], Yu-Chen Liu[1,2], Ying-Hui Wang[1,2], Song Zhang[1,2], Chen He[3], Quan Shi[3], Hai-Xiang Zeng[1,2], Jun-Jian Wang[1,2,*]

[1] Guangdong Key Laboratory of Soil and Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology of China, Shenzhen, Guangdong 518055, China;

[2] State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China;

[3] State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing 102249, China;

* Corresponding Author: Jun-Jian Wang (Mailing Address: 1088 Xueyuan Road, Xili, Nanshan, Shenzhen, Guangdong 518055, China; Email: wangjj@sustech.edu.cn)

Tables

Table S1. General water quality of the mainstream, tributary, pond, reservoir waters and two wastewater (Mean±SD)[a].

Waters	n	pH	EC (mS/cm)	DOC (mg/L)	TDN (mg/L)
Average and statistical differences (mainstream versus tributary versus pond versus reservoir waters)					
Mainstream (M)	22	7.4±0.2c	2.1±3.4a	4.01±0.76bc	15.42±1.95a
Tributary (T)	43	7.7±0.5b	3.1±4.9a	6.87±6.45b	12.40±10.50ab
Pond (P)	11	8.3±0.3a	0.8±0.5a	11.97±6.00a	14.60±26.93bc
Reservoir (R)	6	7.3±0.1c	0.1±0.03b	3.21±0.95c	0.86±0.74c
M vs T vs P vs R		P>T>M≈R	M≈T≈P>R	P>T≈M≈R	M>P≈T>R
Average and statistical differences (Disturbed versus reservoir waters)					
Disturbed waters (DW)	76	7.7±0.3a	2.4±4.3a	6.78±5.95a	13.6±13.14a
Reservoir (R)	6	7.3±0.1b	0.1±0.03b	3.21±0.95b	0.86±0.74b
DW vs R		DW>R	DW>R	DW>R	DW>R
W1	-	6.8	2.4	3.52	14.08
W2	-	6.3	1.7	10.59	41.08

[a]EC: electrical conductivity; DOC: dissolved organic carbon; TDN: total dissolved nitrogen. ">" suggests significant inter-group differences ($P < 0.05$) in the post hoc analyses following one-way ANOVA or nonparametric Kruskal-Wallis test between disturbed and reservoir waters, and among mainstream, tributary, pond and reservoir waters.

Table S2. Comparison of the general water quality in this study with previously reported urban streams.

Rivers	City/State	Country	pH	EC (mS/cm)	DOC (mg/L)	References
Maozhou mainstream	Shenzhen	China	6.7-7.5	0.5-15.8	3.3-5.8	This study
Maozhou mainstream	Shenzhen	China	6.8-7.5	0.5-2.2	4.2-11.2	Liao et al.[1]
Zhujiang River	Guangzhou	China	-	0.4-30	5.1-15.2	Meng et al.[2]
Nanming River	Guiyang	China	7.5-8.0	-	0.5-5.6	Mostofa et al.[3]
Headwater streams	Maine	USA	>7.0	0.5-1.8	5.1-8.1	Parr et al.[4]
Daliao River	Shenyang	China	-	-	4.3-7.5	He et al.[5]

-not available

Table S3. Spectroscopic indices and percentages of four fluorescent components in the mainstream, tributary, pond, reservoir waters and two wastewater (Mean±SD)[a].

Waters	n	$SUVA_{254}$ (L/mg C/m)	E2/E3	FI	HIX	BIX	C1 (%)	C2 (%)	C3 (%)	C4 (%)
Average and statistical differences (mainstream versus tributary versus pond versus reservoir waters)										
Mainstream (M)	22	1.87±0.16a	6.26±0.39a	2.08±0.03a	2.74±0.59b	1.03±0.02a	21±4b	21±5b	26±5a	32±5a
Tributary (T)	43	2.00±0.51a	5.42±1.02b	1.98±0.14b	2.39±0.59c	0.98±0.08b	18±6b	18±5b	28±5a	34±5a
Pond (P)	11	1.39±0.41b	6.20±1.24a	1.85±0.15c	3.16±0.61ab	0.91±0.05c	29±5a	26±5a	24±5ab	21±5b
Reservoir (R)	6	2.01±0.13a	5.34±0.21b	1.71±0.06c	3.98±0.66a	0.90±0.01c	35±5a	30±5a	17±5b	18±5b
Average and statistical differences (Disturbed versus reservoir waters)										
Disturbed waters (DW)	76	1.87±0.46b	5.66±0.46a	1.98±0.14a	2.61±0.66b	0.98±0.07a	21±7b	20±6b	27±8a	32±8a
Reservoir (R)	6	2.01±0.13a	5.34±0.21b	1.71±0.06b	3.98±0.64a	0.90±0.01b	35±5a	30±5a	17±5b	18±5b
DW vs R		R>DW	DW>R	DW>R	R>DW	DW>R	R>DW	R>DW	DW>R	DW>R
W1	-	1.76	8.25	2.10	2.39	1.02	19	21	24	36
W2	-	1.20	7.66	2.13	2.20	1.02	17	19	30	34

[a]$SUVA_{254}$: specific ultraviolet absorbance at 254 nm; FI = fluorescence index; HIX: humification index; BIX = biological index. ">" suggests significant inter-group differences ($P < 0.05$) and the statistical approaches are the same as in Table S1.

Table S4. Comparison of the four fluorescent components identified by PARAFAC with previously reported components.

Component	Excitation max (nm)	Emission max (nm)	Description and likely structure	Reported component in the references
C1	<260 (336)	450	UVC humic-like	Shutova et al.[6]: C1 Murphy et al.[7]: C2
C2	<260 (309)	404	UVC reprocessed humic-like, Low molecular weight, common in marine environments associated with biological activity	Shutova et al.[6]: C2 Wünsch et al.[8]: C3 Kulkarni et al.[9]: C3
C3	275	310	Tyrosine-like, Amino acids, free or bound in proteins, fluorescence resembles free tyrosine	Hambly et al.[10]: C5 Podgorski et al.[11]:C5
C4	<240	350	Tryptophan-like, Amino acids, free or bound in proteins, fluorescence resembles free tryptophan	Osburn et al.[12]: C6 Wünsch et al.[8]: C1 Walker et al.[13]: C4

Table S5. Loadings of water parameters on the first two principal components of principal component analysis.

Water parameter	Principal component 1	Principal component 2
pH	-0.344	0.427
Electronic conductivity (EC)	0.465	0.226
Dissolved organic carbon (DOC)	0.266	**0.905**
Total dissolved nitrogen (TDN)	**0.613**	**0.524**
DOC/TDN	**-0.621**	0.298
Specific ultraviolet absorbance at 254 nm ($SUVA_{254}$)	-0.004	**-0.739**
E2/E3	-0.123	-0.267
Fluorescence index (FI)	**0.776**	-0.104
Humification index (HIX)	**-0.922**	-0.114
Biological index (BIX)	**0.633**	**-0.532**
Fluorescent component C1	**-0.955**	0.045
Fluorescent component C2	**-0.919**	0.044
Fluorescent component C3	**0.835**	0.413
Fluorescent component C4	**0.698**	-0.458

Table S6. Abundance changes of some representative CHOS molecules from M6 to M13.

Neutral Mass	Class	Formula	Percentage at M6	Percentage at M9	Percentage at M13	Abundance change from M6 to M9	Abundance change from M9 to M13
312.10	O5S1	$C_{15}H_{20}O_5S_1$	1.8%	2.8%	2.1%	68.1%	-29.9%
298.09	O5S1	$C_{14}H_{18}O_5S_1$	1.1%	1.8%	1.3%	63.9%	-29.0%
284.07	O5S1	$C_{13}H_{16}O_5S_1$	0.6%	0.8%	0.6%	55.5%	-27.5%
326.12	O5S1	$C_{16}H_{22}O_5S_1$	0.7%	1.0%	0.8%	50.0%	-25.5%
282.13	O3S1	$C_{15}H_{22}O_3S_1$	0.1%	0.3%	0.1%	149.2%	-50.9%
314.05	O7S1	$C_{13}H_{14}O_7S_1$	0.1%	0.3%	0.2%	109.1%	-39.0%
274.05	O6S1	$C_{11}H_{14}O_6S_1$	0.1%	0.2%	0.1%	105.8%	-42.1%
296.14	O3S1	$C_{16}H_{24}O_3S_1$	0.2%	0.3%	0.2%	104.3%	-41.6%
328.06	O7S1	$C_{14}H_{16}O_7S_1$	0.2%	0.3%	0.2%	100.6%	-35.8%
310.16	O3S1	$C_{17}H_{26}O_3S_1$	0.2%	0.3%	0.2%	77.5%	-33.5%
342.08	O7S1	$C_{15}H_{18}O_7S_1$	0.2%	0.3%	0.2%	75.7%	-31.3%

Abundance change from M6 to M9 = $(P_{M9}\times DOC_{M9}-P_{M6}\times DOC_{M6}) / P_{M6}\times DOC_{M6}$, and abundance change from M9 to M13 = $(P_{M13}\times DOC_{M13}-P_{M9}\times DOC_{M9}) / P_{M9}\times DOC_{M9}$, where P_i = the percentage of single CHOS compound at site *i* and DOC_i = dissolved organic carbon concentration at site *i*.

Figures

Figure S1. Photos of some representative sites: (A) reservoir water at R6, (B) upper mainstream site at M2, (C) middle mainstream site at M14, (D) lower mainstream site at M22, (E) upper tributary site at T1-3, (F) lower tributary site at T20-4, (G) pond water at P11, (H) wastewater site at W1.

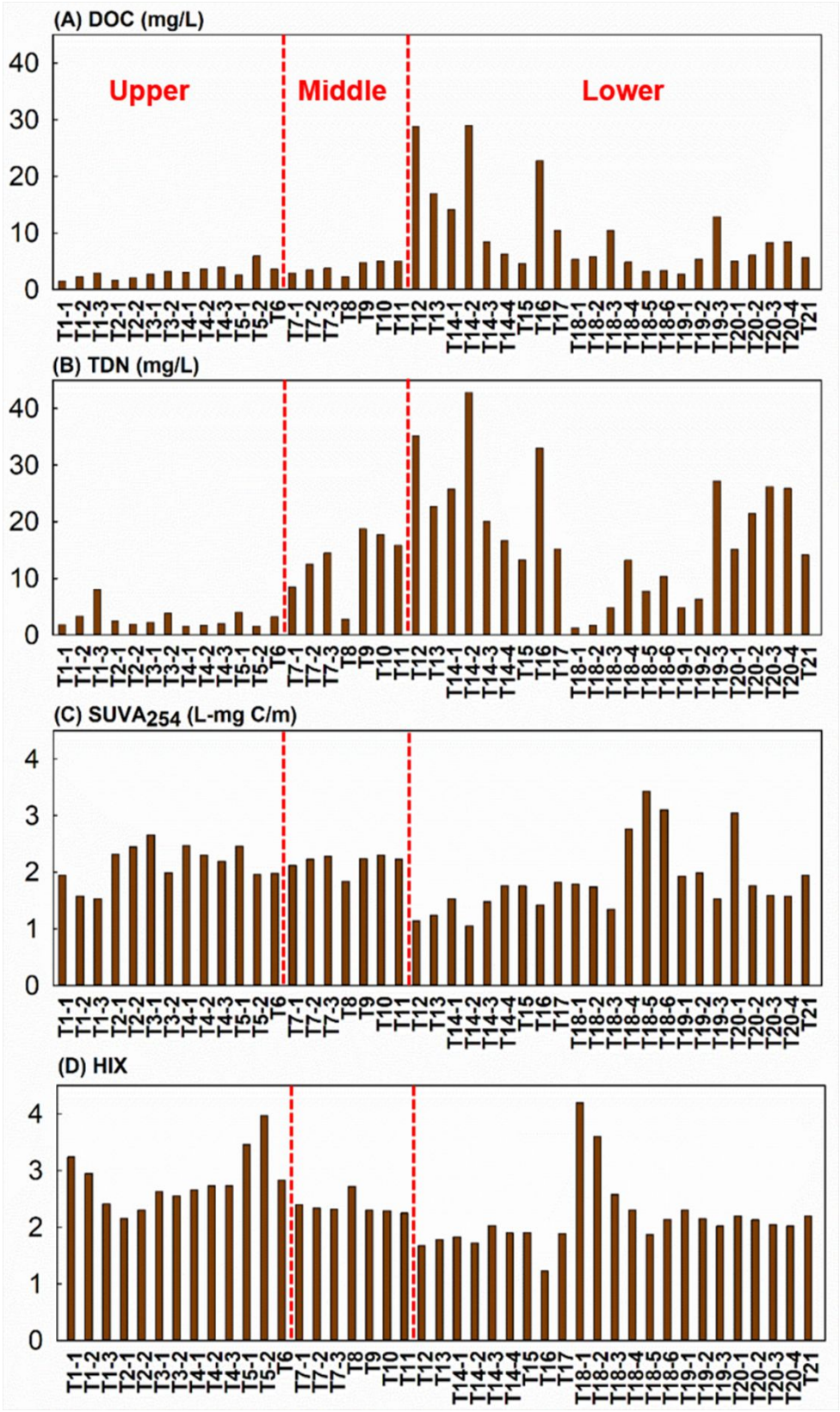

Figure S2. Increased DOC and TDN levels and decreased optical indices ($SUVA_{254}$ and HIX) from upper to lower tributary. (A) DOC: dissolved organic carbon; (B) TDN: total dissolved nitrogen; (C) $SUVA_{254}$: specific ultraviolet absorbance at 254 nm; (D) HIX: humification index.

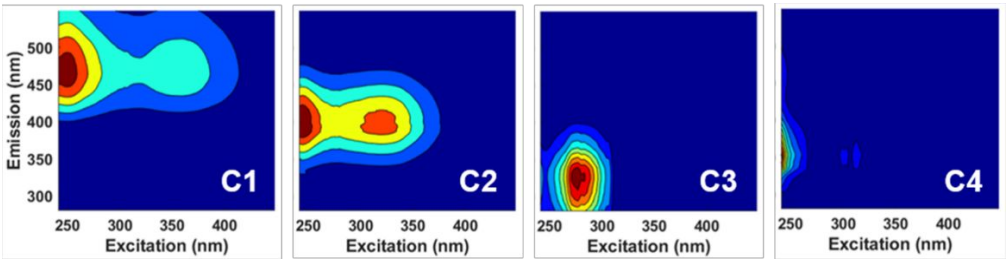

Figure S3. Loadings of the four fluorescent dissolved organic matter (DOM) components on the excitation and emission spectra.

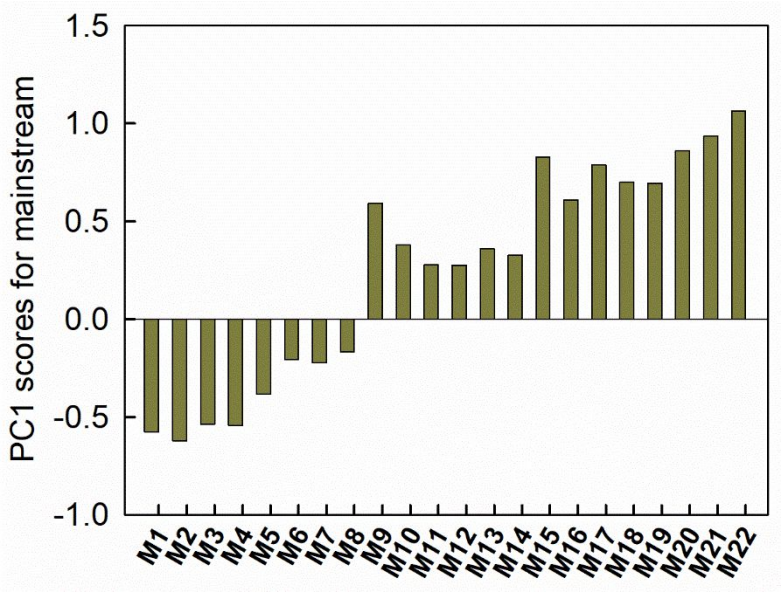

Figure S4. Scores of principal component 1 (PC1: 42.7%) from mainstream sites M1 to M22.

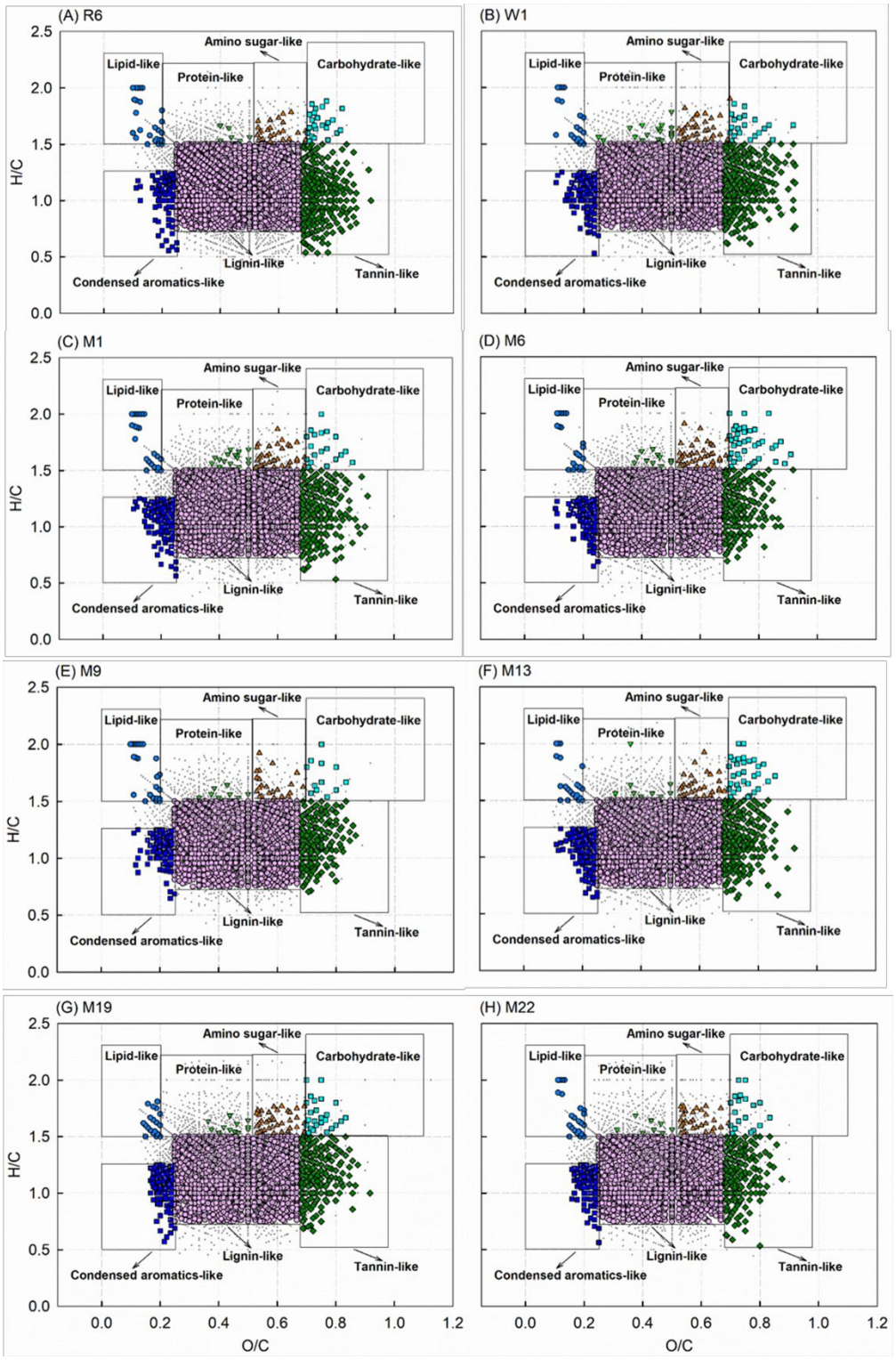

Figure S5. Representative van Krevelen diagrams of all the compound classes in all 8 water samples (A) R6, (B) W1, (C) M1, (D) M6, (E) M9, (F) M13, (G) M19, (H) M22 based on the method in Wu et al.[14]. The gray dots represent the unclassified group.

REFERENCES

(1) Liao, H.; Yu, K.; Duan, Y.; Ning, Z.; Li, B.; He, L.; Liu, C., Profiling microbial communities in a watershed undergoing intensive anthropogenic activities. *Sci. Total Environ.* **2019,** *647*, 1137-1147.

(2) Meng, F.; Huang, G.; Yang, X.; Li, Z.; Li, J.; Cao, J.; Wang, Z.; Sun, L., Identifying the sources and fate of anthropogenically impacted dissolved organic matter (DOM) in urbanized rivers. *Water Res.* **2013,** *47*, (14), 5027-5039.

(3) Mostofa, K. M. G.; Wu, F.; Liu, C.-Q.; Fang, W. L.; Yuan, J.; Ying, W. L.; Wen, L.; Yi, M., Characterization of Nanming River (southwestern China) sewerage-impacted pollution using an excitation-emission matrix and PARAFAC. *Limnology* **2010,** *11*, (3), 217-231.

(4) Parr, T. B.; Cronan, C. S.; Ohno, T.; Findlay, S. E. G.; Smith, S. M. C.; Simon, K. S., Urbanization changes the composition and bioavailability of dissolved organic matter in headwater streams. *Limnol. Oceanogr.* **2015,** *60*, (3), 885-900.

(5) He, D.; He, C.; Li, P.; Zhang, X.; Shi, Q.; Sun, Y., Optical and Molecular Signatures of Dissolved Organic Matter Reflect Anthropogenic Influence in a Coastal River, Northeast China. *J. Environ. Qual.* **2019,** *48*, (3), 603-613.

(6) Shutova, Y.; Baker, A.; Bridgeman, J.; Henderson, R. K., Spectroscopic characterisation of dissolved organic matter changes in drinking water treatment: From PARAFAC analysis to online monitoring wavelengths. *Water Res.* **2014,** *54*, 159-169.

(7) Murphy, K. R.; Hambly, A.; Singh, S.; Henderson, R. K.; Baker, A.; Stuetz, R.; Khan, S. J., Organic Matter Fluorescence in Municipal Water Recycling Schemes: Toward a Unified PARAFAC Model. *Environ. Sci. Technol.* **2011,** *45*, (7), 2909-2916.

(8) Wünsch, U. J.; Murphy, K. R.; Stedmon, C. A., The One-Sample PARAFAC Approach Reveals Molecular Size Distributions of Fluorescent Components in Dissolved Organic Matter. *Environ. Sci. Technol.* **2017,** *51*, (20), 11900-11908.

(9) Kulkarni, H. V.; Mladenov, N.; Datta, S.; Chatterjee, D., Influence of monsoonal recharge on arsenic and dissolved organic matter in the Holocene and Pleistocene aquifers of the Bengal Basin. *Sci. Total Environ.* **2018,** *637*, 588-599.

(10) Hambly, A. C.; Arvin, E.; Pedersen, L. F.; Pedersen, P. B.; Seredyńska-Sobecka, B.; Stedmon, C. A., Characterising organic matter in recirculating aquaculture systems with fluorescence EEM spectroscopy. *Water Res.* **2015,** *83*, 112-120.

(11) Podgorski, D. C.; Zito, P.; McGuire, J. T.; Martinovic-Weigelt, D.; Cozzarelli, I. M.; Bekins, B. A.; Spencer, R. G. M., Examining Natural Attenuation and Acute Toxicity of Petroleum-Derived Dissolved Organic Matter with Optical Spectroscopy. *Environ. Sci. Technol.* **2018,** *52*, (11), 6157-6166.

(12) Osburn, C. L.; Anderson, N. J.; Stedmon, C. A.; Giles, M. E.; Whiteford, E. J.; McGenity, T. J.; Dumbrell, A. J.; Underwood, G. J. C., Shifts in the Source and Composition of Dissolved Organic Matter in Southwest Greenland Lakes Along a Regional Hydro-climatic Gradient. *J. Geophys. Res. Biogeosci.* **2017,** *122*, (12), 3431-3445.

(13) Walker, S. A.; Amon, R. M. W.; Stedmon, C. A., Variations in high-latitude riverine fluorescent dissolved organic matter: A comparison of large Arctic rivers. *J. Geophys. Res. Biogeosci.* **2013,** *118*, (4), 1689-1702.

(14) Wu, X.; Wu, L.; Liu, Y.; Zhang, P.; Li, Q.; Zhou, J.; Hess, N. J.; Hazen, T. C.; Yang, W.; Chakraborty, R., Microbial Interactions With Dissolved Organic Matter Drive Carbon Dynamics and Community Succession. *Front. Microbiol.* **2018,** *9*, 1234.