Supporting Information for:

Direct non-target analysis of dissolved organic matter and disinfection by-products in drinking water with nanoLC-FT-ICR-MS

Limei Han[1], Martin Lohse[1], Maolida Nihemaiti[1], Thorsten Reemtsma[1,2], and Oliver J. Lechtenfeld*[1,3]

[1] Department of Analytical Chemistry, Helmholtz Centre for Environmental Research (UFZ), 04318 Leipzig, Germany

[2] Institute of Analytical Chemistry, University of Leipzig, 04103 Leipzig, Germany

[3] ProVIS – Centre for Chemical Microscopy, Helmholtz Centre for Environmental Research (UFZ), 04318 Leipzig, Germany

* Corresponding Author: oliver.lechtenfeld@ufz.de

Text S1: Drinking water samples

The drinking water treatment plants A and B are operating with 38.000 m^3/d and 25.000 m^3/d capacity, respectively. Both DWTPs receive water from a nearby river via bank filtration. The water treatment process in both plants are, in short: A) bank filtration, mechanical deacidification/aeration, sand filtration, and disinfection with chlorine (Cl_2) gas; B) bank filtration, mechanical deacidification/aeration, flocculation, sand filtration, and disinfection with chlorine dioxide (ClO_2). After disinfection, both DWTPs directly supply the drinking water to consumers.

Table S1. Chemical parameters of the drinking water samples used in this study.

DWTP		DOC (mg/L)	Chlorine (mg/L)
DWTP-A	Before disinfection	2.1	ND*
	After disinfection	2.0	0.2
DWTP-B	Before disinfection	2.5	ND*
	After disinfection	2.7	ND*

* ND: Not detected (below detection limit).

Table S2. Chemical properties and retention time (RT) of model compounds on the C18 column used in this study.

#	Compound	Concentration (ppm)	Formula	Molecular mass (Da)	Log D (pH3)	RT (min)	Company
1	D-Glucuronic acid	0.2	$C_6H_{10}O_7$	194.14	-2.82	7.5-8.0	Sigma-Aldrich
2	2-(4-(2,2-Dicarboxy-ethyl)-2,5-dimethoxy-benzyl)malonic acid	0.04	$C_{16}H_{18}O_{10}$	370.31	0.01	15.6-16.2	Sigma-Aldrich
3	Isoferulic Acid 3-O-ß-D-Glucuronide	0.04	$C_{16}H_{18}O_{10}$	370.31	-0.65	16.9-17.5	Toronto Research Chemicals
4	Vanillic acid	4.7	$C_8H_8O_4$	168.15	1.14	16.3-16.9	Sigma-Aldrich
5	Fraxin	0.04	$C_{16}H_{18}O_{10}$	370.31	-1.25	18.2-18.7	Sigma-Aldrich

Table S3. The gradient program used in the nanoLC system. The flow rate was set to 500 nL/min.

Time (min)	MeOH (v/v %)	ultrapure water (v/v %)
0	1	99
0.5	1	99
10.5	99	1
30.5	99	1
65	1	99

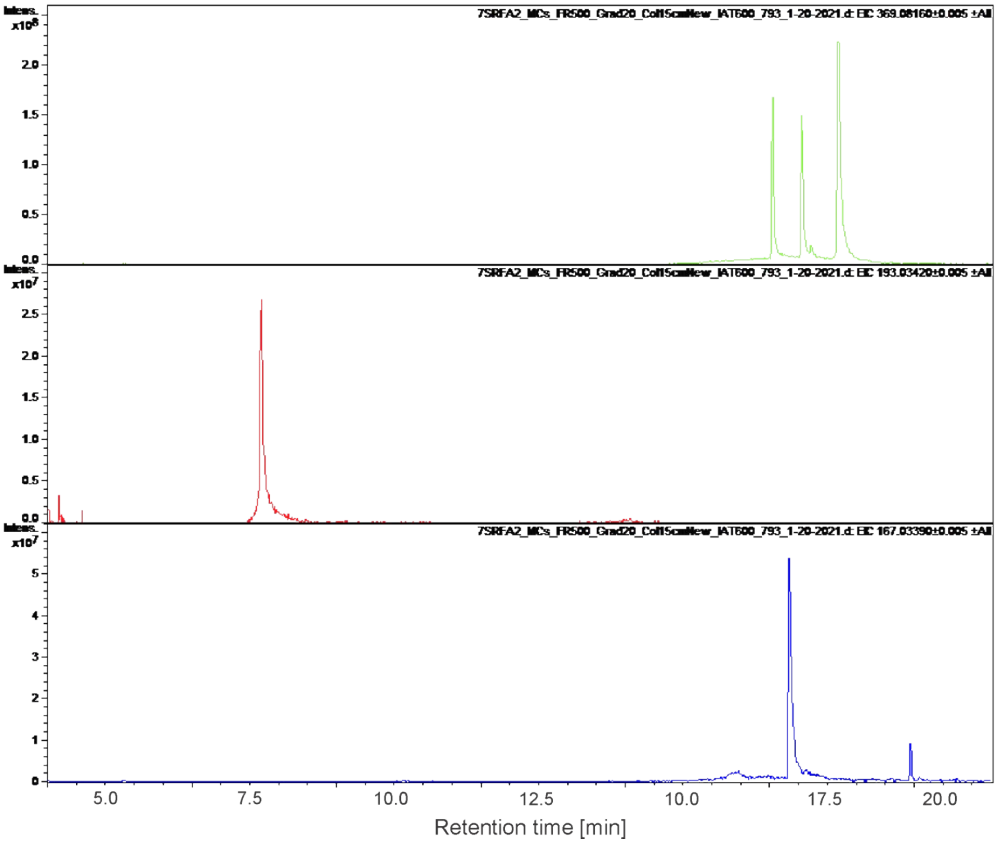


Figure S1. Extracted ion chromatograms (EICs) of model compounds D-Glucuronic acid (red), Vanillic acid (blue), and 3 isomeric compounds (green: 2-(4-(2,2-Dicarboxy-ethyl)-2,5-dimethoxy-benzyl)malonic acid, Isoferulic Acid 3-O-ß-D-Glucuronide, and Fraxin) using the nanoLC method.

Table S4. Comparison of the chromatographic parameters for three isomeric compounds using HPLC ([1]) and nanoLC (this study) methods.

Compounds	k-value* on nanoLC column[#1]	k-value* on RPLC column[#2]
2-(4-(2,2-Dicarboxy-ethyl)-2,5-dimethoxy-benzyl)malonic acid	1.1	2.3
Isoferulic Acid 3-O-ß-D-Glucuronide	1.3	2.4
Fraxin	1.5	2.7

* Retention factor (k-values) are calculated by: $k = (t_R - t_0)/ t_0$ with t_R: retention time of the analyte and t_0: column dead time.

[#1] Acclaim PepMap, 75 µm x 150 mm, 3 µm, 100 Å, Thermo Fisher Scientific, USA; flow rate 500 nL/min.

[#2] ACQUITY HSS T3, 1.8 µm, 100 Å, 150x3mm, Waters, USA; flow rate 0.2 mL/min.

Table S5. Mass resolving power at m/z 400 in samples DW-nLC-A and DW-SPE-nLC-A in different segments.

	DW-nLC-A	DW-SPE-nLC-A
Segment (min)	Mass resolving power at m/z 400	
11-11.5	278963	NA
11.5-12	287997	359793
12-12.5	253787	288076
12.5-13	270668	273120
13-13.5	291577	261661
13.5-14	288383	272381
14-14.5	278164	265052
14.5-15	268308	263016
15-15.5	267877	278540
15.5-16	275574	280936
16-16.5	263983	268291
16.5-17	266621	279894
17-17.5	275722	276978
17.5-18	273465	277444
18-18.5	270342	266645
18.5-19	270685	260556
19-19.5	268027	275142
19.5-20	256972	274119
20-20.5	262094	260818
20.5-21	274556	271424

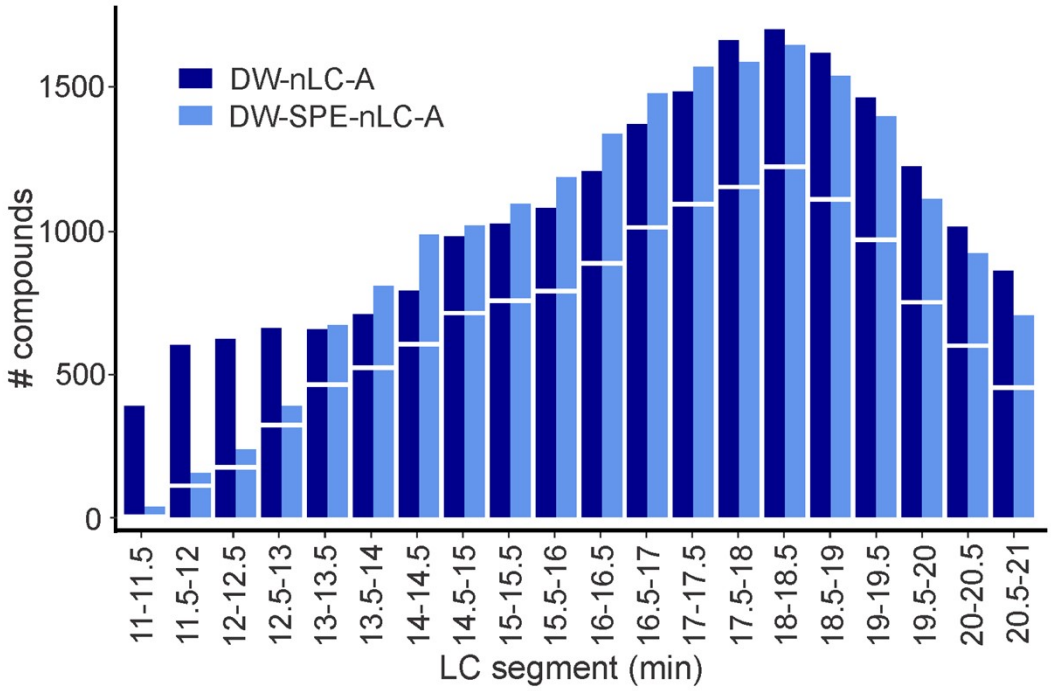


Figure S2. Total number of DOM compounds detected in DW-nLC-A and DW-SPE-nLC-A. The white lines represent the number of shared compounds between both samples in the respective segment.

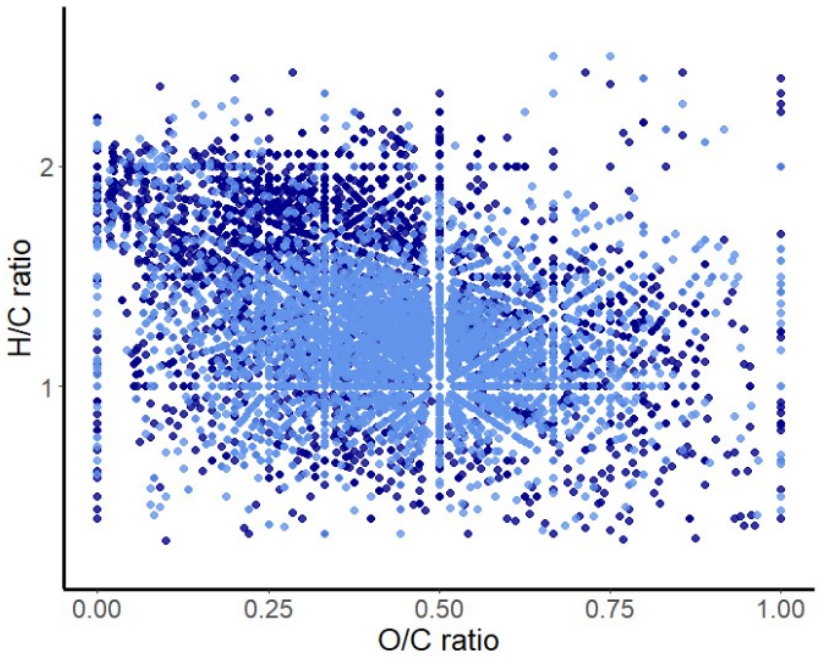


Figure S3. All uniquely detected DOM compounds in SPE (DW-SPE-nLC-A, blue) and non-SPE (DW-nLC-A, dark blue) processed samples.

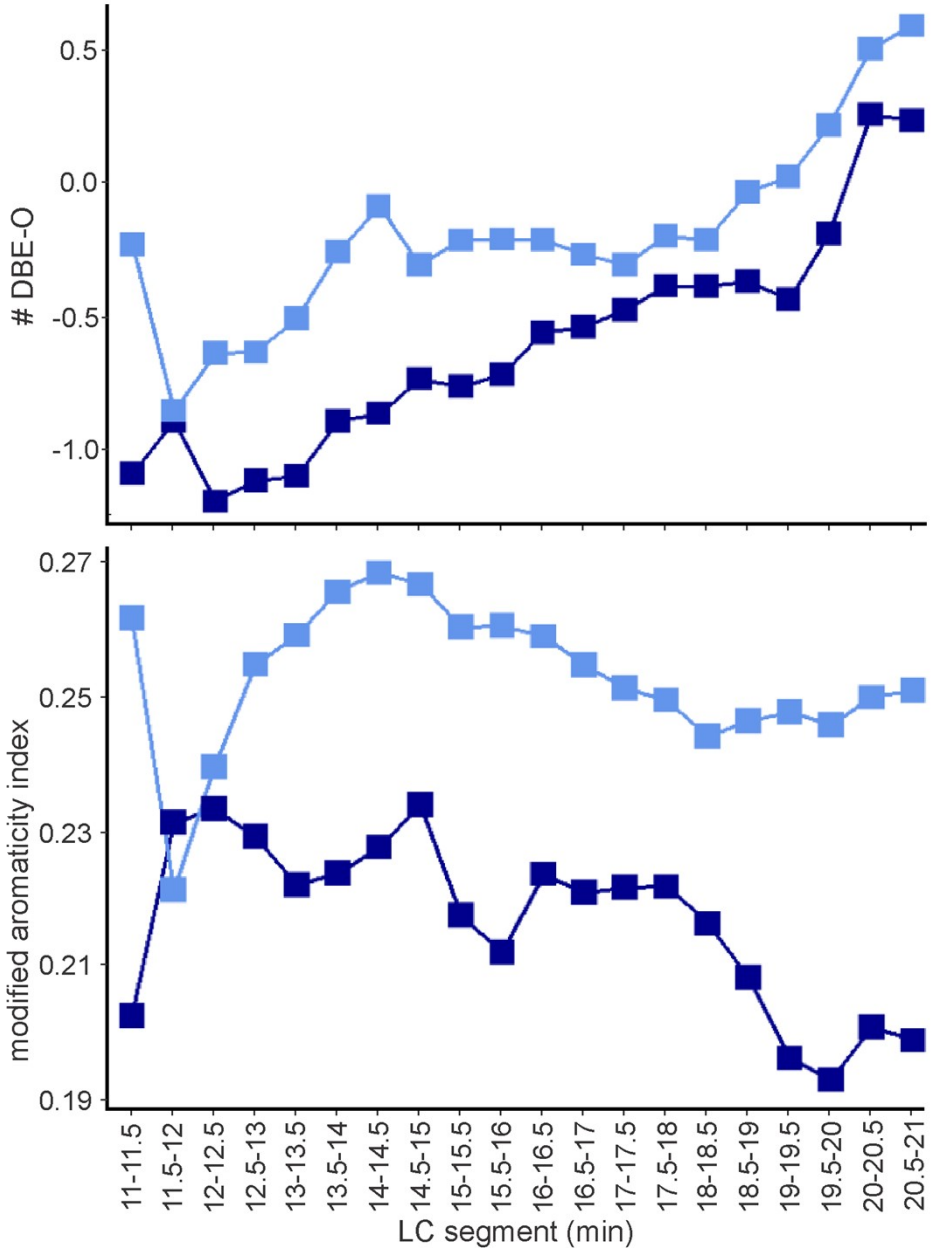


Figure S4. Intensity weighted average modified aromaticity index (AI_{mod}) (a) and DBE-O (b) of DOM compounds in different segments in SPE (DW-SPE-nLC-A, blue) and non-SPE (DW-nLC-A, dark blue) processed samples.

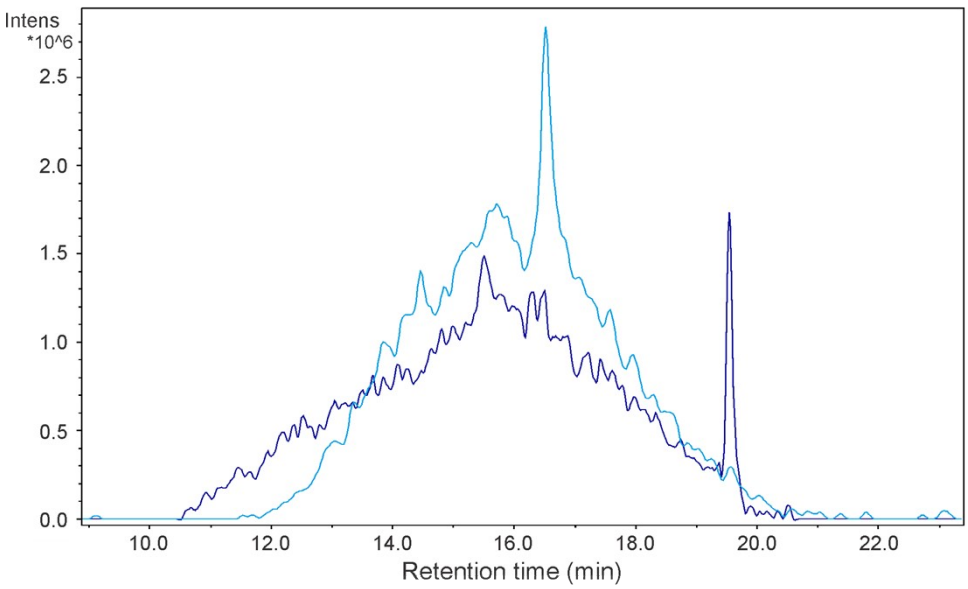


Figure S5. EIC of *m/z* 313.0929, $C_{14}H_{18}O_8$ in DW-nLC-A (dark blue, without extraction), DW-SPE-nLC-A (light blue, extracted with SPE).

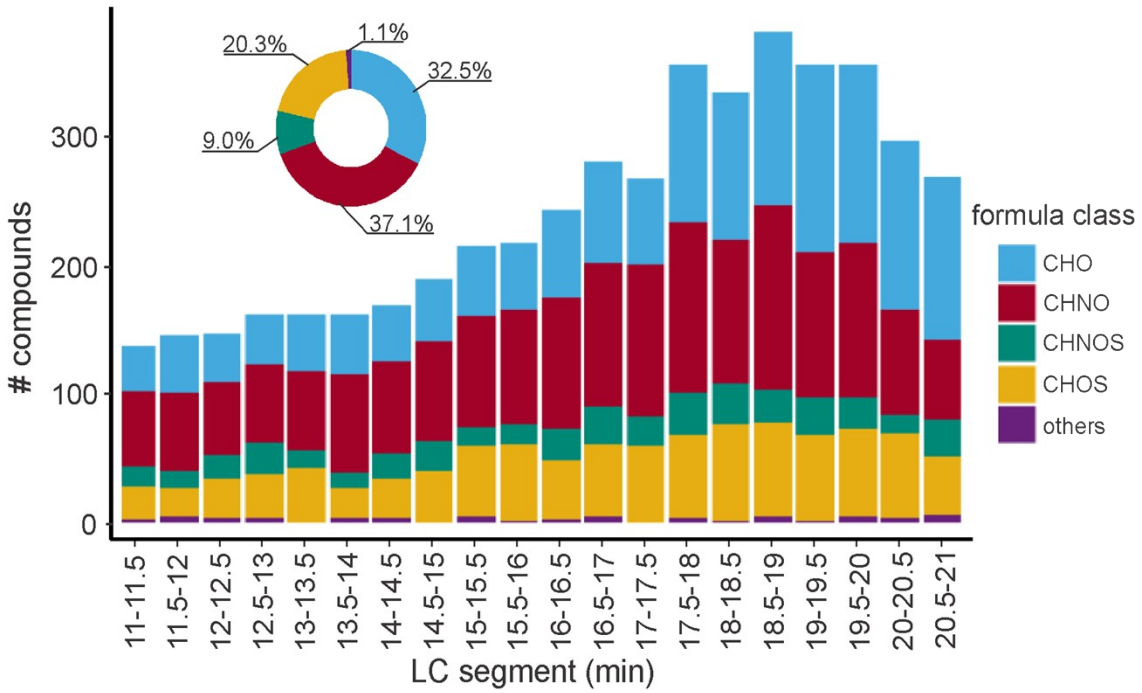


Figure S6. Formula class distribution of non-halogenated DBP compounds in sample DW-nLC-B. The bars indicate the total number of non-halogenated DBPs detected in each segment distinguished by formula classes (CHO, CHNO, CHOS, CHNOS, and others). The inlay summarizes the percentage of formula classes of all non-halogenated DBPs (all segments combined).

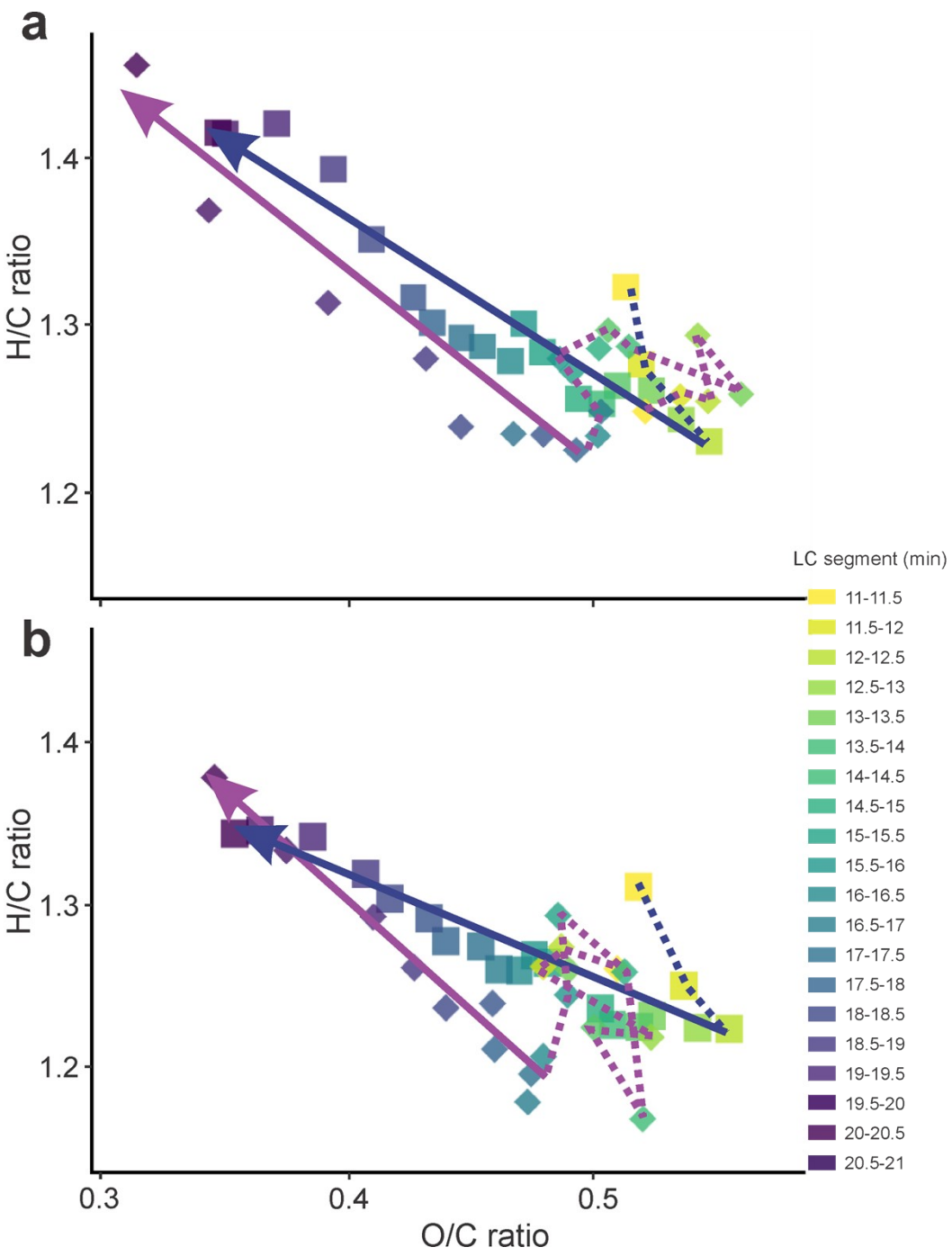
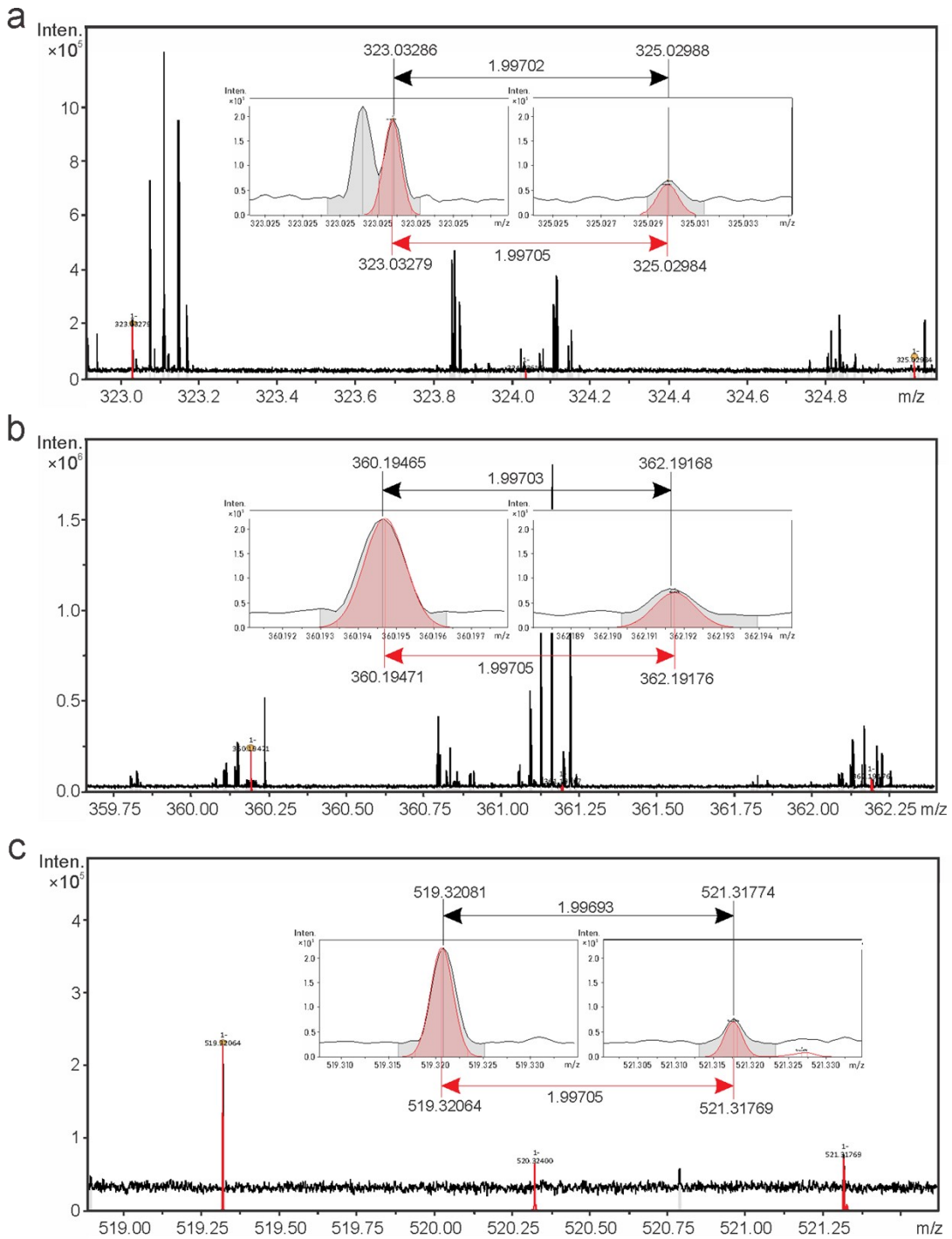


Figure S7. Intensity weighted average hydrogen-to-carbon (H/C) vs oxygen-to-carbon (O/C) ratios of all segments of DOM (square, connected with blue line) and non-halogenated DBPs (diamond, connected with purple line) MFs detected in a) DW-nLC-A and b) DW-nLC-B. Color represents mean retention time (min) of segments. Lines connecting dots are used as visual aid.

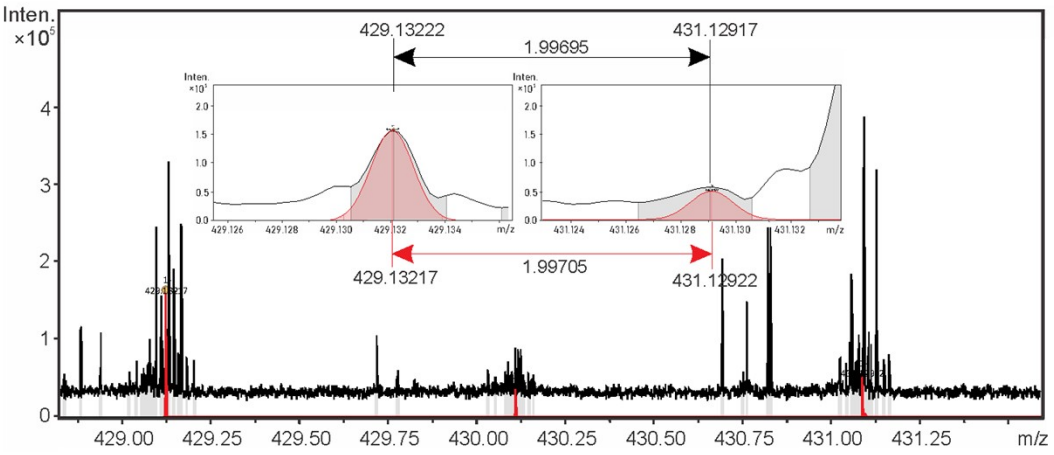
Table S6. Peak mass to charge ratio (*m/z*), signal to noise ratio (S/N), molecular formula (MF), and MF hydrogen to carbon ratio (H/C), oxygen to carbon ratio (O/C), double bond equivalent (DBE), double bond equivalent minus oxygen (DBE-O), and retention time (RT) of corresponding segment of detected chlorinated DBPS (Cl-DBPs) in sample DW-nLC-A and DW-nLC-B.

#	peak m/z	peak S/N	Molecular formula	H/C	O/C	DBE	DBE-O	RT (min)	sample	Figure S8
1	323.03286	47	C15 H13 Cl1 O6	0.87	0.40	9	3	16-16.5	DW-nLC-A	a
2	360.19465	52	C18 H32 Cl1 N1 O4	1.78	0.22	3	-1	18-18.5	DW-nLC-A	b
3	519.32081	60	C26 H49 Cl1 N2 O6	1.89	0.23	3	-3	11.5-12	DW-nLC-B	c
4	429.13223	33	C20 H27 Cl1 O8	1.35	0.40	7	-1	16.5-17	DW-nLC-B	d
5	329.06992	48	C17 H15 Cl1 N2 O3	0.89	0.18	11	8	17-17.5	DW-nLC-B	e
6	413.13738	37	C20 H27 Cl1 O7	1.35	0.35	7	0	17-17.5	DW-nLC-B	f
7	378.20526	54	C18 H34 Cl1 N1 O5	1.89	0.28	2	-3	18-18.5	DW-nLC-B	g
8	396.21586	50	C18 H36 Cl1 N1 O6	2.00	0.33	1	-5	18-18.5	DW-nLC-B	h
9	396.21586	59	C18 H36 Cl1 N1 O6	2.00	0.33	1	-5	18.5-19	DW-nLC-B	i
10	280.98919	67	C9 H11 Cl1 O6 S1	1.22	0.67	4	-2	18.5-19	DW-nLC-B	j
11	336.19467	53	C16 H32 Cl1 N1 O4	2.00	0.25	1	-3	18.5-19	DW-nLC-B	k
12	415.14322	38	C22 H25 Cl1 N2 O4	1.14	0.18	11	7	19-19.5	DW-nLC-B	l
13	377.02334	13	C15 H17 Cl3 N2 O3	1.13	0.20	7	4	19.5-20	DW-nLC-B	m
14	382.23658	27	C18 H38 Cl1 N1 O5	2.11	0.28	0	-5	19.5-20	DW-nLC-B	n
15	348.23106	95	C18 H36 Cl1 N1 O3	2.00	0.17	1	-2	20-20.5	DW-nLC-B	o
16	344.19981	101	C18 H32 Cl1 N1 O3	1.78	0.17	3	0	20-20.5	DW-nLC-B	p

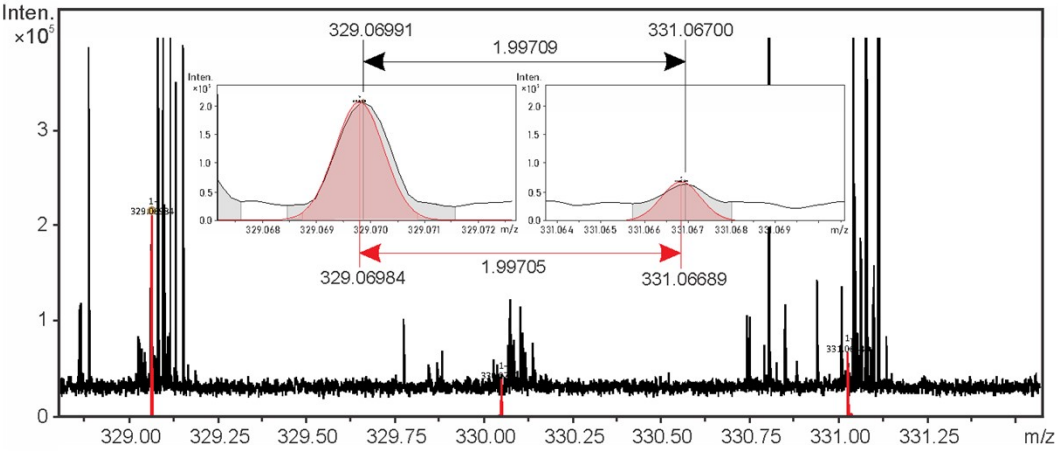

a
Inten.
×10^5
323.03286
1.99702
325.02988
323.03279
1.99705
325.02984
m/z
b
Inten.
×10^8
360.19465
1.99703
362.19168
360.19471
1.99705
362.19176
m/z
c
Inten.
×10^5
519.32081
1.99693
521.31774
519.32064
1.99705
521.31769
520.32400
m/z

91

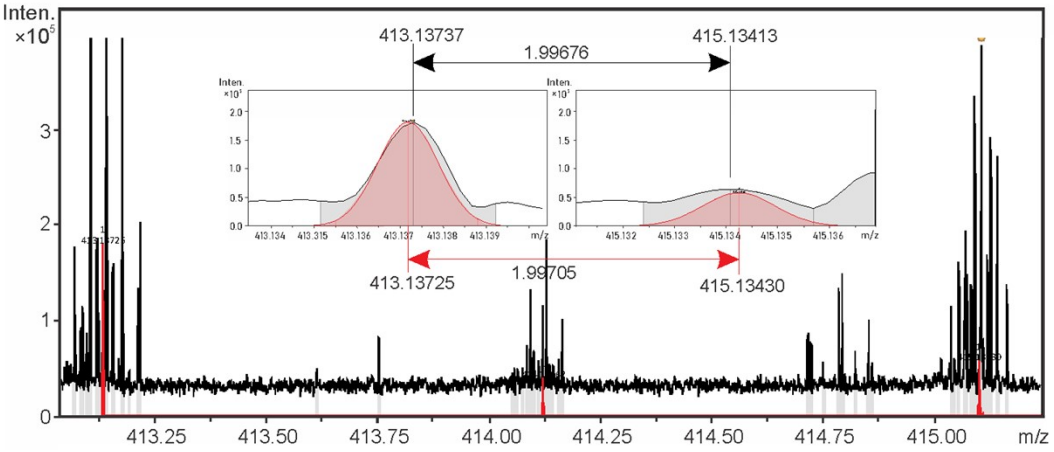
d


Inten.
×10⁵
429.13222
1.99695
431.12917
429.13217
1.99705
431.12922
m/z

e

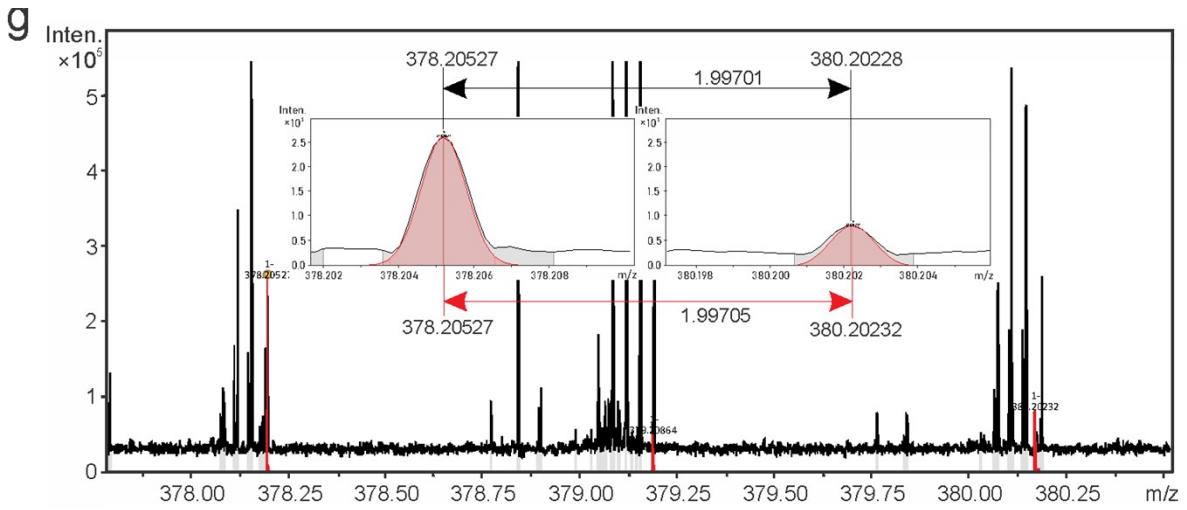

Inten.
×10⁵
329.06991
1.99709
331.06700
329.06984
1.99705
331.06689
m/z

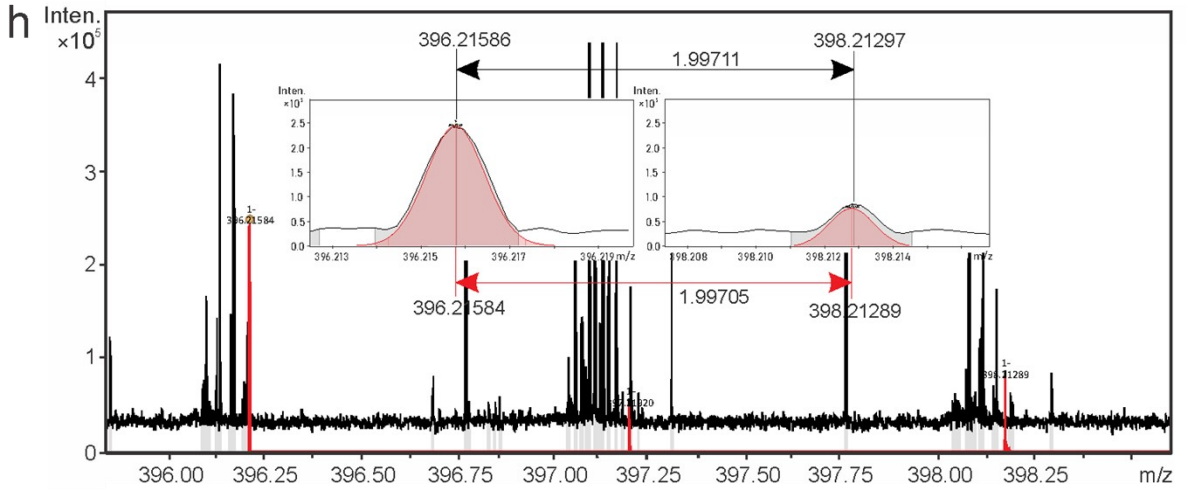
f

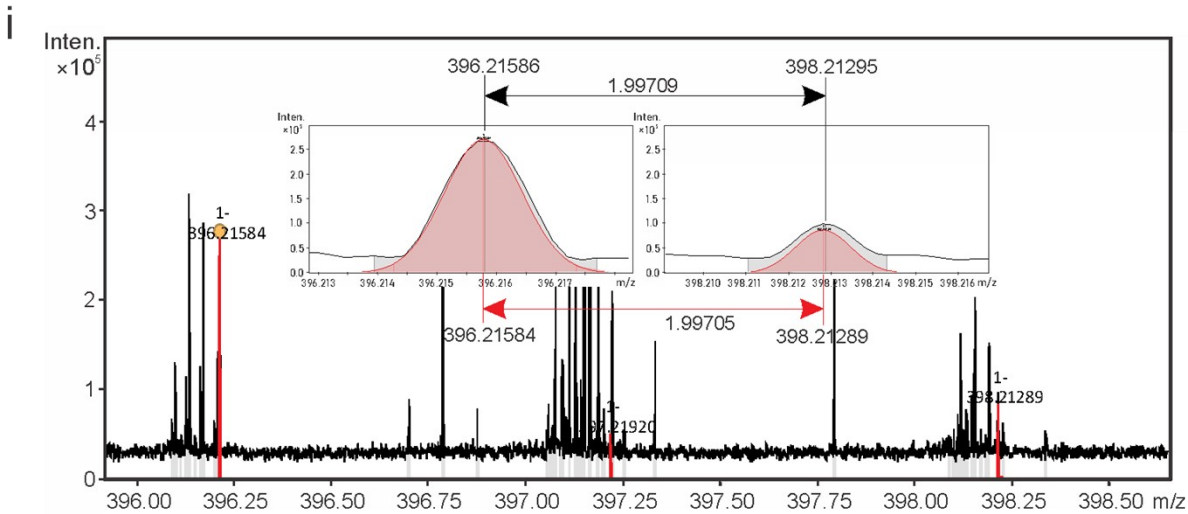

Inten.
×10⁵
413.13737
1.99676
415.13413
413.13725
1.99705
415.13430
m/z

92

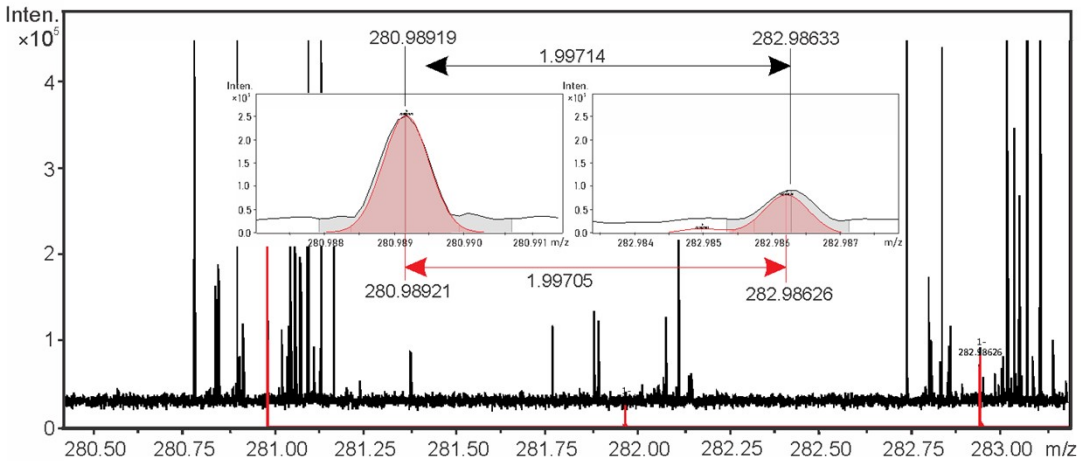
93


g
Inten.
×10^5
378.20527
1.99701
380.20228
Inten.
×10^3
378.20527
1.99705
380.20232
m/z

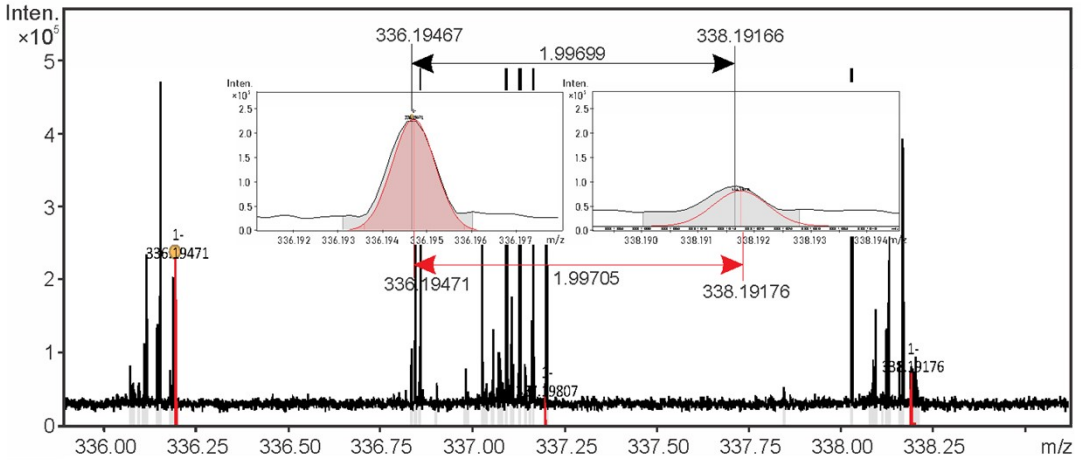

h
Inten.
×10^5
396.21586
1.99711
398.21297
Inten.
×10^3
396.21584
1.99705
398.21289
m/z


i
Inten.
×10^5
396.21586
1.99709
398.21295
Inten.
×10^3
396.21584
1.99705
398.21289
m/z

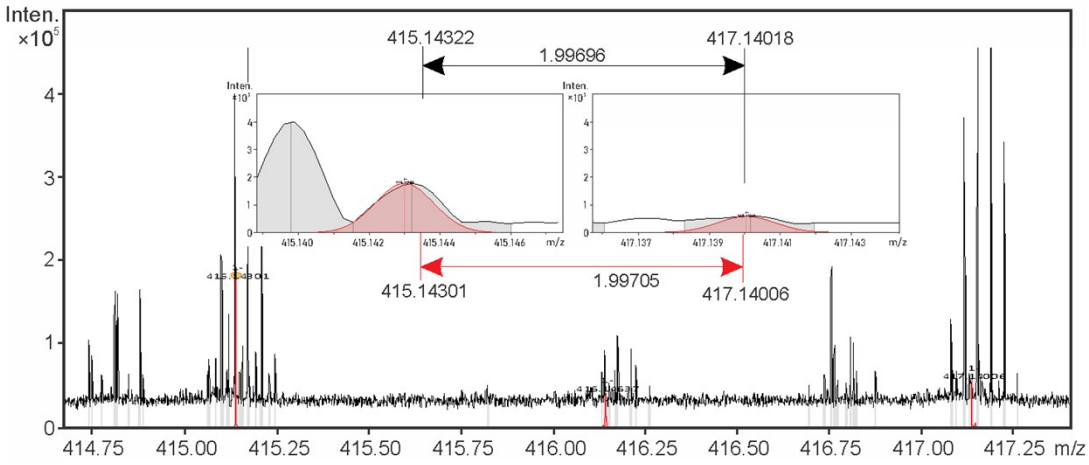
j


Inten.
×10^5
280.98919
1.99714
282.98633
280.98921
1.99705
282.98626
280.50 280.75 281.00 281.25 281.50 281.75 282.00 282.25 282.50 282.75 283.00 m/z

k

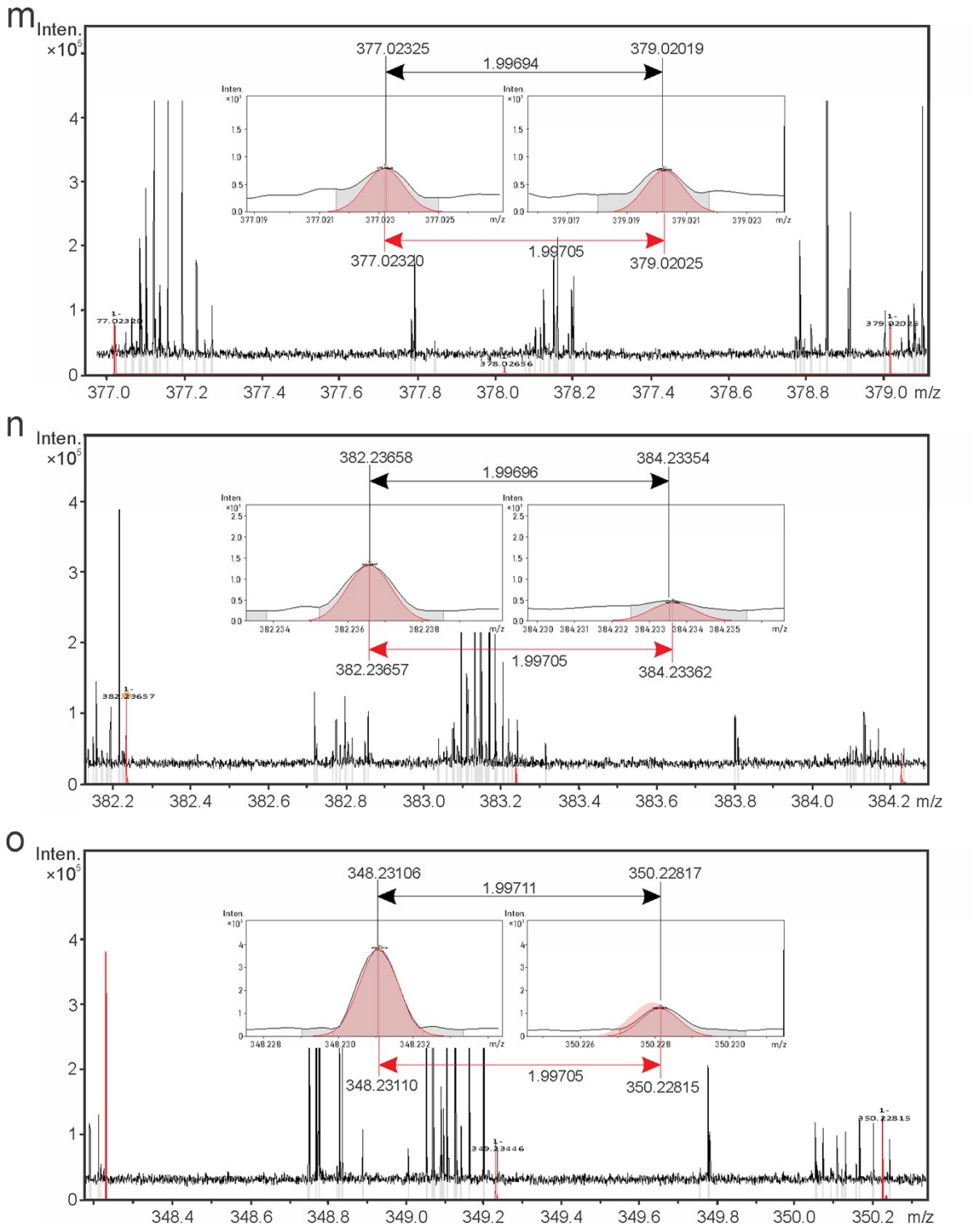

Inten.
×10^5
336.19467
1.99699
338.19166
336.19471
1.99705
338.19176
336.00 336.25 336.50 336.75 337.00 337.25 337.50 337.75 338.00 338.25 m/z

l


Inten.
×10^5
415.14322
1.99696
417.14018
415.14301
1.99705
417.14006
414.75 415.00 415.25 415.50 415.75 416.00 416.25 416.50 416.75 417.00 417.25 m/z

94

95


m
Inten.
×10^5
377.02325
1.99694
379.02019
377.02320
1.99705
379.02025
377.0 377.2 377.4 377.6 377.8 378.0 378.2 377.4 378.6 378.8 379.0 m/z
n
Inten.
×10^5
382.23658
1.99696
384.23354
382.23657
1.99705
384.23362
382.2 382.4 382.6 382.8 383.0 383.2 383.4 383.6 383.8 384.0 384.2 m/z
o
Inten.
×10^5
348.23106
1.99711
350.22817
348.23110
1.99705
350.22815
348.4 348.6 348.8 349.0 349.2 349.4 349.6 349.8 350.0 350.2 m/z

96

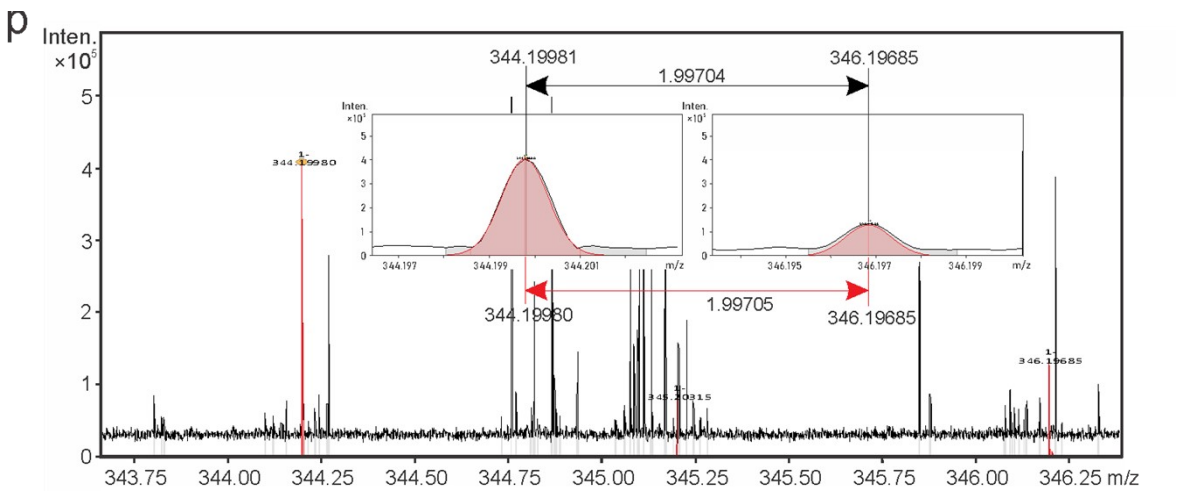


Figure S8. Raw spectra (black) and theoretical isotopologue intensity pattern matching for 16 Cl-DBPs (ref. Table S6) detected in DW-nLC-A (a and b) and DW-nLC-B (c to p). The mass differences of isotopologue peaks were indicated with black (raw spectra) and red (theoretical isotopologue) arrows.

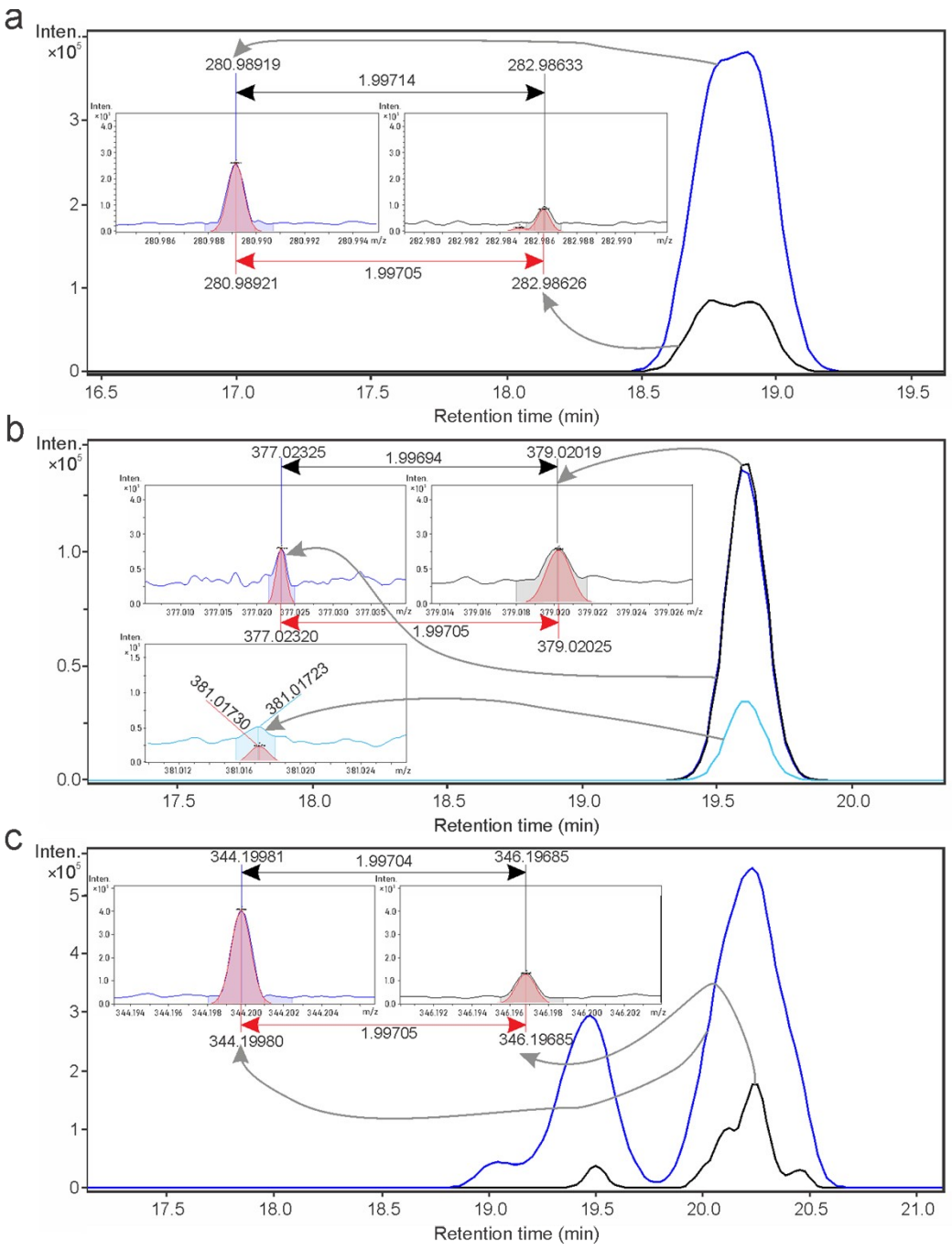


Figure S9. EICs of molecular formula a) C9 H11 Cl1 O6 S1, b) C15 H17 Cl3 N2 O3, and c) C18 H32 Cl1 N1 O3. For mono-chlorinated MF: $^{35}Cl_1$ (blue), $^{37}Cl_1$ (black); for tri-chlorinated MF: $^{35}Cl_3$ (blue), $^{35}Cl_2{}^{37}Cl_1$ (black), $^{35}Cl_1{}^{37}Cl_2$ (light-blue). Inside figures: the theoretical isotopologue peaks (red) and measured peaks (blue, black, and light-blue) in raw spectra in DW-nLC-B. The mass differences of isotopologue peaks were indicated with black (raw spectra) and red (theoretical isotopologue) arrows.

109

Additional References:

(1) Han, L.; Kaesler, J.; Peng, C.; Reemtsma, T.; Lechtenfeld, O. J. Online Counter Gradient LC-FT-ICR-MS Enables Detection of Highly Polar Natural Organic Matter Fractions. *Analytical Chemistry* **2021**, *93* (3), 1740-1748. DOI: 10.1021/acs.analchem.0c04426.