Suspect and non-target screening for contaminants of emerging concern in an urban estuary

Zhenyu Tian[1,2]*, Katherine T. Peter[1,2], Alex D. Gipe[1,2], Haoqi Zhao[3], Fan Hou[3], David A. Wark[1,2], Tarang Khangaonkar[4], Edward P. Kolodziej[1,2,3], C. Andrew James[1,2]*

[1]Center for Urban Waters, Tacoma, WA, 98421, USA

[2]Interdisciplinary Arts and Sciences, University of Washington Tacoma, Tacoma, WA, 98421, USA

[3]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, 98195, USA

[4]Pacific Northwest National Laboratories, 1100 Dexter Avenue N, Seattle, WA 98011, USA

*Corresponding author contact information

Email: tianzy@uw.edu, jamesca@uw.edu

Supporting information

Pages: 20

Tables: 8

Figure: 6

List of reference standards

Most reference standards for identification and semi-quantification were purchased from Sigma-Aldrich (St. Louis, MO, USA). The name, purity, and CAS registration number were listed below: triclopyr (98%, CAS 55335-06-3), 2,4-dinitrophenol (98%, CAS 51-28-5), chlorothalonil (98%, CAS 28343-61-5), bisphenol S (98%, CAS 80-09-1), venlafaxine hydrochloride (98%, CAS 99300-78-4), O-desvenlafaxine (93413-62-8), lidocaine (CAS 137-58-6), tebuthiuron (34014-18-1), 2,6-dichlorobenzamide (97%, 2008-58-4), sucralose (98%, 56038-13-2), 4-nitrophenol (100-02-7), hexa(methoxymethyl)melamine (HMMM; 95%, CAS 3089-11-0), 1,3-diphenylguanidine (97%, CAS 102-06-7), 1,3-dicyclohexylurea (98%, CAS 2387-23-7), N-butylbenzenesulfonate (99%, CAS 3622-84-2), tris(2-chloroethyl)phosphate (95%, CAS 115-96-8), tris(2-butoxyethyl)phosphate (95%, CAS 78-51-3), fluridone (CAS 59756-60-4), diuron (CAS 330-54-1), diisononyl phthalate (99%, CAS 28553-12-0), 1-cyclohexyl-2-pyrrolidone (99%, CAS 6837-24-7), 5-methyl-1-H-benzotriazole (98%, CAS 136-85-6), N,N-dimethyldodecylamine (99% CAS 1643-20-5), 2-benzothiazolesulfonic acid potassium salt (CAS 941-57-1), salicylic acid (CAS 69-72-7), 4-hydroxyquinoline (98%, CAS 611-36-9), 2-mercaptobenzothiazole (97%, CAS 149-30-4), benzotriazole (98%, CAS 95-14-7), triisopropanolamine (95%, CAS 122-20-3), caffeine (CAS 58-08-2), propamocarb (CAS 24579-73-5), metsulfuron-methyl (CAS 74223-64-6), rimsulfuron (CAS 122931-48-0), vanillin (CAS 121-33-5), metoprolol tartrate (98%, CAS 56392-17-7). Methamphetamine (CAS 7632-10-2) was purchased from Cerilliant (98%, Round Rock, TX, USA 78665). PFOS (CAS 1763-23-1), PFHxS (CAS 355-46-4), PFBS (CAS 375-73-5), PFOA (CAS 335-67-1), PFHpA (CAS 375-85-9), PFHxA (CAS 307-24-4), PFPeA (CAS 2706-90-3) were purchased from Wellington Laboratories (Guelph, ON, Canada).

Semi-quantification using external calibration curve and internal standards

CECs confirmed by reference standards were quantified based on 4-point external calibration curves (1, 5, 50, 200 μg/L) on the UHPLC-qTOF, using the same instrumental methods as were used for analyzing the marine water samples. The concentration range of the calibration curve was based on the estimated concentration range in marine water (0-100 ng/L) and the 1000x concentration from solid phase extraction processing, resulting in an in-vial concentrations of 0-100 μg/L. Concentrations were first measured based on the external calibration curves. Next, for each internal standard (IS, 11 compounds in ESI+ and 8 compounds in ESI-, **Table S2**) spiked in the samples and the calibrants, IS response ratios were calculated based on their peak areas (Equation 1), and the median response ratio was calculated amongst all of the IS. Then, the normalized concentrations were calculated by dividing measured concentrations with median response ratio (Equation 2), to account for differences between calibrants and samples.

$$\text{IS response ratio} = \frac{\text{IS peak area in sample}}{\text{IS peak area in calibration curve}} \quad (1)$$

$$\text{normalized concentration} = \text{measured concentration} / \text{median IS response ratio} \quad (2)$$

Although the loss during sample extraction processes was not be taken into consideration, the variation from instrumental sensitivity and matrix effects were taken into account by the IS correction method.

Supporting Information Tables (Table S4 and S5 in a separate Excel file)

Table S1. Information for marine water sampling sites and date sampled.

No.	Site Name	Latitude	Longitude	Apr	May	Jun	Aug	Oct
1	Port Townsend Water Street	48.10917	-122.76630		√	√	√	√
2	Point No Point	47.90860	-122.52670	√	√	√	√	√
3	Silverdale, Dyes Inlet	47.64279	-122.69671	√	√	√	√	√
4	Hood Canal Holly	47.57058	-122.97154	√	√	√	√	√
5	Evergreen Rotary Park	47.57617	-122.62899	√	√	√	√	√
6	Commencement Bay Skookum	47.28931	-122.40957		√	√	√	√
7	Hammersley Inlet-Arcadia Point	47.19903	-122.93953	√	√	√	√	√
8	Budd Inlet, West Bay	47.06035	-122.91503	√	√	√	√	√
9	Saltar's Point	47.16998	-122.61066	√	√	√	√	√
10	Salmon Beach	47.29181	-122.52806	√	√	√	√	√
11	Commencement Bay Thea Foss	47.25919	-122.43469		√	√	√	√
12	Edmonds Ferry	47.81418	-122.38215		√	√	√	√
13	Everett Boat Launch	48.00523	-122.22237		√	√	√	√
14	Salmon Bay, Commodore Park	47.66630	-122.40180		√	√	√	√
15	Smith Cove, Terminal 91	47.63237	-122.37868		√	√	√	√
16	West Point South	47.65930	-122.43348		√	√	√	√
17	Joe Block Park	47.58491	-122.36708			√	√	√
18	Jack Perry Memorial Park	47.58890	-122.34869			√	√	√

April 17th: accumulated precipitation 0.5-0.7 inch

Oct 30th: accumulated precipitation 0.9-1.2 inch (data source: weather.gov)

Table S2. Isotope-labelled internal standard mixture for qTOF analysis.

Compound	Formula	RT[a]	ESI polarity	Conc. in injection vial (ng/mL)	Range of Matrix Effects (%)
Carbamazepine-d10	$C_{15}H_2D_{10}N_2O$	6.32	+	25	77-120
Caffeine-13C3	$^{13}C_3C_5H_{10}N_4O_2$	3.98	+	50	93-157
Cotinine-d3	$C_{10}H_9D_3N_2O$	3.39	+	100	67-86
Sulfamethoxazole-d4	$C_{10}H_7D_4N_3O_3S$	4.13	+, -	100	69-133
Vanillin-d3	$C_8H_5D_3O_3$	4.29	-	100	87-149
Sulfadimethoxine-d6	$C_{12}H_8D_6N_4O_4S$	4.86	+, -	100	52-127
5-methyl-1H-benzotriazole-d6	$C_7HD_6N_3$	4.93	+, -	100	67-91
Ethyl paraben-d4	$C_9H_6D_4O_3$	5.92	-	100	109-170
DEET-d7	$C_{12}H_{10}D_7NO$	7.15	+	100	62-89
Atrazine-d5	$C_8H_9D_5ClN_5$	6.98	+, -	100	51-77
Prometon-d3	$C_{10}H_{16}D_3N_5O$	7.81	+	100	51-74
Ibuprofen-d3	$C_{13}H_{15}D_3O_2$	8.93	-	100	66-109
Bis(2-ethylhexyl)phthalate-d4	$C_{24}H_{34}D_4O_4$	16.73	+	100	58-92
Theobromine-d3	$C_7H_2D_6N_4O_2$	2.99	+	200	54-87
Nicotine-d3	$C_{10}H_{11}D_3N_2$	1.88	+	500	76-131
Propyl paraben-d4	$C_{10}H_8D_4O_3$	7.07	-	500	97-161

a: RT= retention time, unit: min.

Table S3. Parameters for XCMS Online. Positive and negative mode shared same parameters except for adduct.

Parameter Category	Parameter	Setting
Feature detection	method	centWave
	mzdiff	0.01
	S/N threshold	6
	noise filter	1000
Retention time correction	method	obiWarp
	profStep (*m/z*)	0.1
Alignment	bin width (s)	5
	minfrac	0.65
	mzwid	0.015
Statistics	test	Welch T test
	p value threshold	0.05
	fold change threshold	5
Annotation	ppm	10
	m/z absolute error	0.015
	adduct	[M+H], [M+Na], [M+NH4], [M-H]

Table S6. Predicted "dye" concentration from the Salish Sea Model, describing the distribution, resulting from transport of wastewater effluent from 99 unique WWTP outfalls throughout Puget Sound (mean ± standard deviation). The concentration of all WWTP effluent was 100.

Site No.	Site Name	May	June	August
4	HoodCanal	0.06515 ± 0.00173	0.06650 ± 0.00251	0.06261 ± 0.00212
5	EvergreenRPk	0.22197 ± 0.03161	0.21370 ± 0.03522	0.28917 ± 0.06196
8	BuddInlet	0.53766 ± 0.13805	0.63766 ± 0.12424	0.73726 ± 0.13873
9	SaltarsPoint	0.25351 ± 0.04964	0.24373 ± 0.04720	0.20957 ± 0.03695
11	CommBayTheaFoss	0.20374 ± 0.01218	0.18944 ± 0.00430	0.20060 ± 0.01115
12	EdmondsFerry	0.15112 ± 0.00728	0.14909 ± 0.00579	0.13920 ± 0.00289
13	EverettBoatLaunch	0.15342 ± 0.01429	0.16812 ± 0.01210	0.14782 ± 0.00685
14	SalmonBayCPk	0.04041 ± 0.01176	0.05731 ± 0.01635	0.09332 ± 0.00703
15	SmithCove	0.16226 ± 0.01017	0.15680 ± 0.00631	0.14526 ± 0.00393
16	WestPointSouth	0.16500 ± 0.00851	0.16211 ± 0.00691	0.14453 ± 0.00309
18	JackPerryMPk	0.07627 ± 0.01148	0.08622 ± 0.01082	0.12631 ± 0.00604

Table S7. Numbers of non-target features at each site/date (passed through replicate filters and blank subtraction).

Site No.	Apr pos	Apr neg	May pos	May neg	Jun pos	Jun neg	Aug pos	Aug neg	Oct pos	Oct neg	Sum Pos	Sum Neg	Sum	Average No. of features
1			335	252	625	308	604	332	606	121	2170	1013	3183	796
2	231	141	388	226	442	319	473	289	472	101	2006	1076	3082	616
3	315	210	337	199	449	228	733	415	597	136	2431	1188	3619	724
4	268	135	265	175	350	192	526	373	251	98	1660	973	2633	527
5	258	111	327	196	503	290	697	412	445	117	2230	1126	3356	671
6			405	213	420	209	533	368	480	131	1838	921	2759	690
7	343	264	468	312	635	291	611	349	562	151	2619	1367	3986	797
8	271	139	133	60	544	269	586	357	490	131	2024	956	2980	596
9	335	204	334	170	349	153	669	381	436	129	2123	1037	3160	632
10	357	285	353	158	381	192	342	271	385	114	1818	1020	2838	568
11			417	257	619	308	628	366	535	134	2199	1065	3264	816
12			377	251	509	244	706	412	210	55	1802	962	2764	691
13			68	61	318	147	248	136	351	69	985	413	1398	350
14			255	151	379	191	452	270	291	57	1377	669	2046	512
15			437	316	552	241	642	284	326	74	1957	915	2872	718
16			496	335	470	339	628	368	303	97	1897	1139	3036	759
17					347	218	417	266	148	47	912	531	1443	481
18					289	120	311	192	261	60	861	372	1233	411

Table S8. Summary of linear regression between the most prevalent CECs and the predicted dye concentration from the Salish Sea Model. The dye represents the combined contributions and transport of WWTP effluent from 99 outfalls throughout Puget Sound.

	May		**June**		**August**	
	R^2	p^a	R	*p*	R^2	*p*
Lamotrigine	0.5	0.085	0.73	**0.024**[b]	0.65	**0.018**
Sucralose	0.25	0.25	0.82	**0.047**	0.7	**0.0051**
2-Hydroxyquinoline	0.078	0.77	0.0035	0.99	0.13	0.36
4-Nitrophenol	0.071	0.74	0.41	0.18	0.0079	0.91

a. *p* values were calculated from 95% quantile of a 10,000-replicate bootstrap
b. bold figures indicate statistically significant results ($\alpha = 0.05$)

Supporting Information Figures

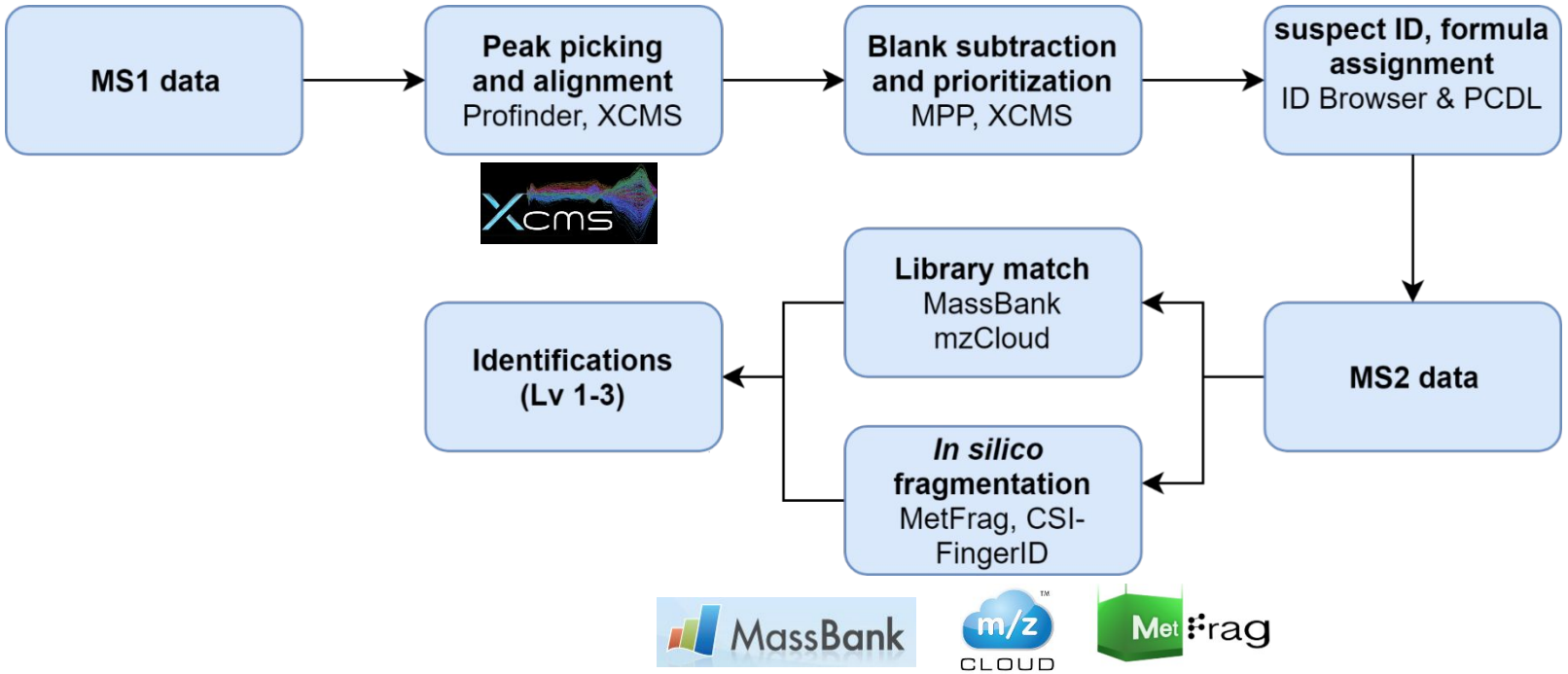


Figure S1. The workflow for data prioritization and identification. MPP = Mass Profiler Professional, PCDL = Personal Compound Database and Library.

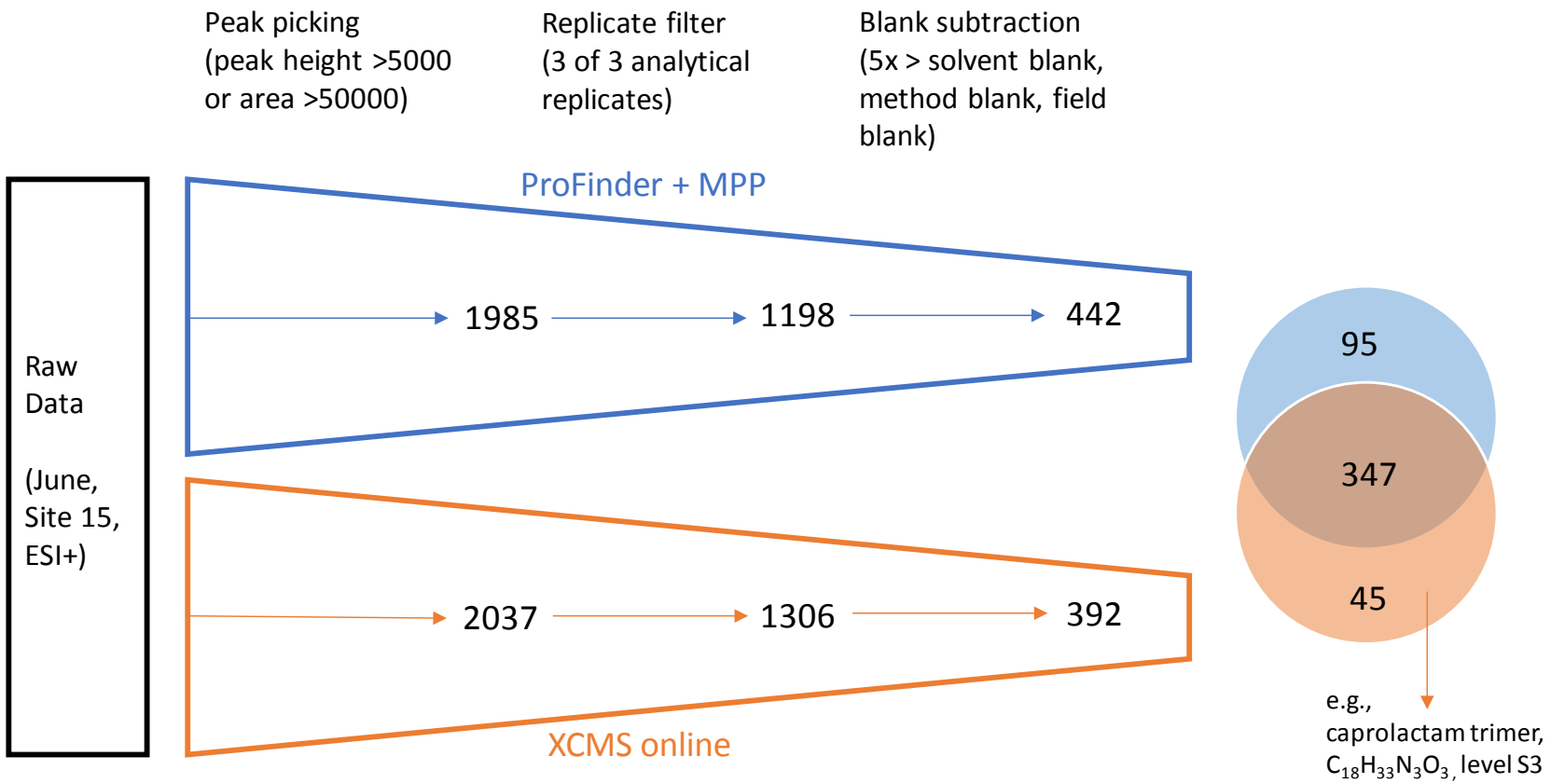


Figure S2. Comparison of the two workflows (Profinder + MPP and XCMS online) for feature detection and prioritization. The example data was from Site 15, June 2018, in ESI+ mode. The numbers represent the amounts of features.

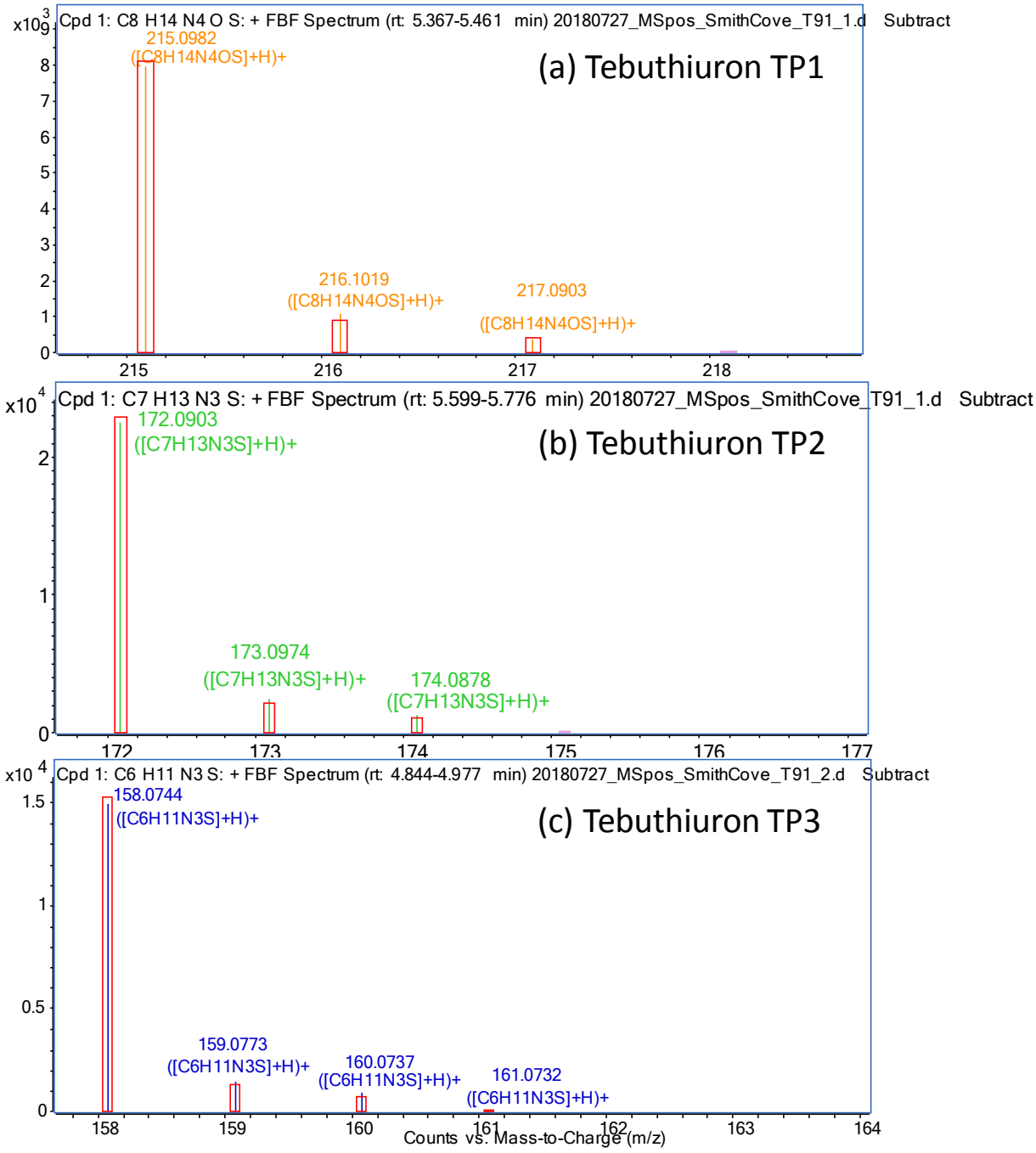


Figure S3. Isotope patterns of tebuthiuron TPs and chlorothalonil TPs: (**a**) tebuthiuron TP1; (**b**) tebuthiuron TP2; (**c**) tebuthiuron TP3; (**d**) 4-hydroxy-chlorothalonil; (**e**) 1-amide-4-hydroxy-chlorothalonil. The red boxes represent theoretical isotope patterns calculated by Agilent Qualitative Analysis, and the bars are measured mass spectra.

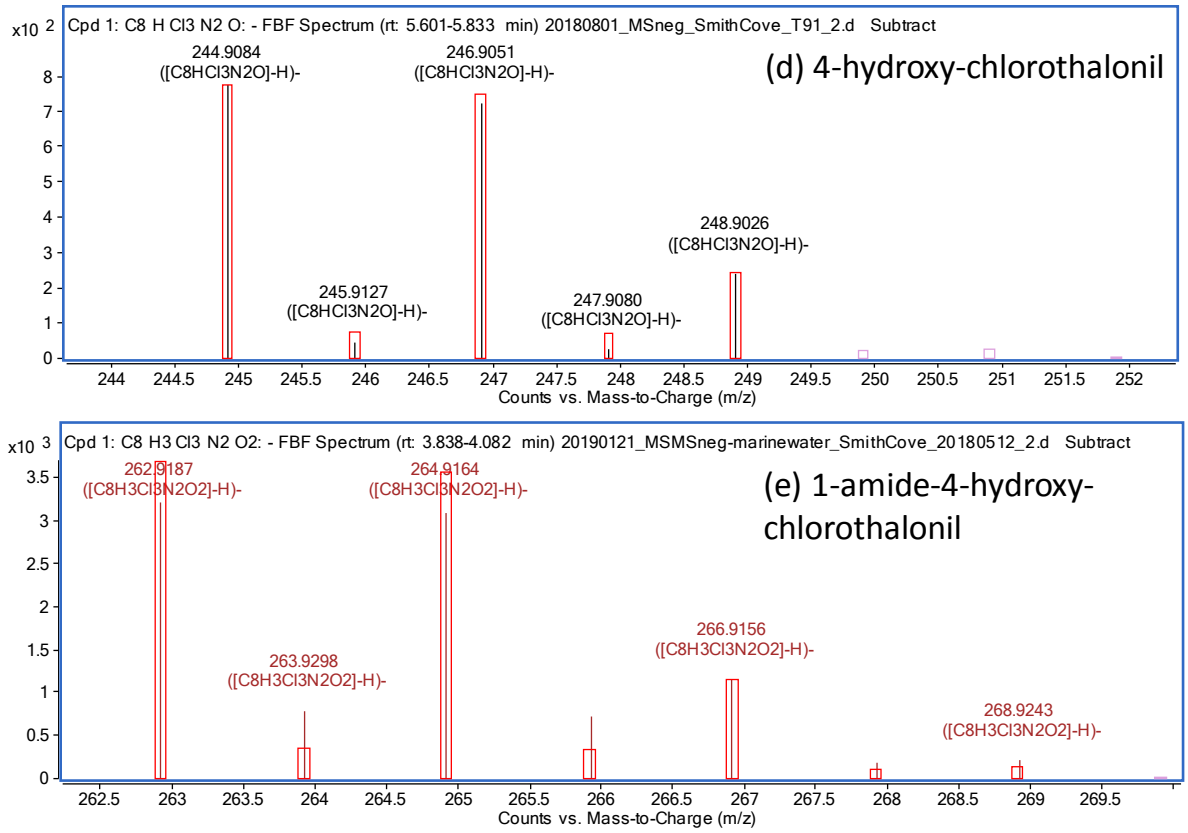
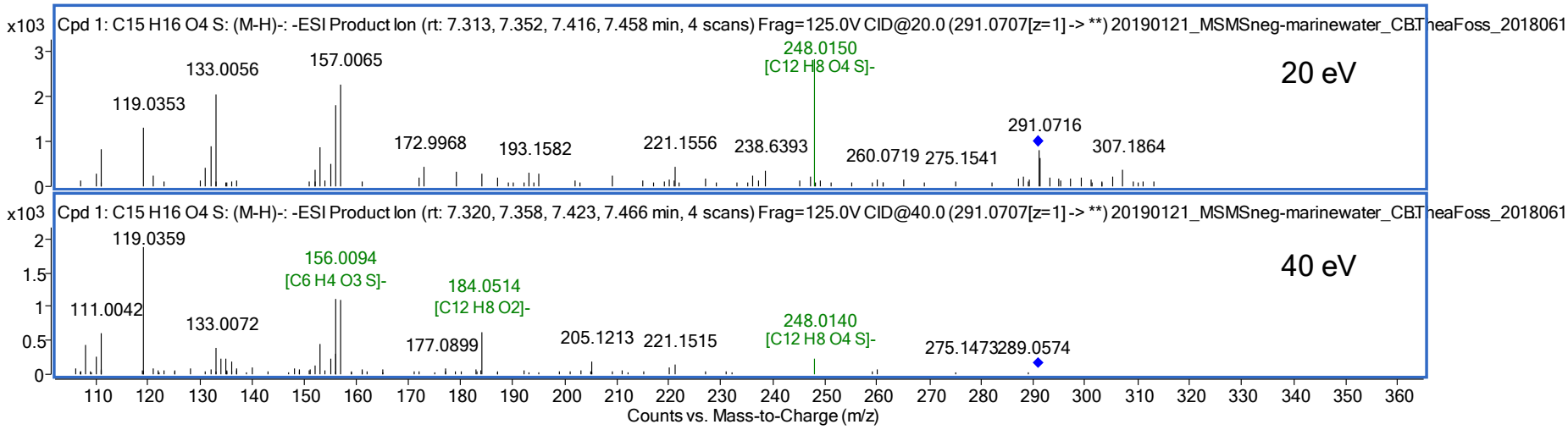


Figure S3, continued.

(a) 4-((4-Isopropoxyphenyl)sulfonyl)phenol ($C_{15}H_{16}O_4S$)



(b) 1-Amide-4-hydroxy-chlorothalonil ($C_8H_3Cl_3N_2O_2$)

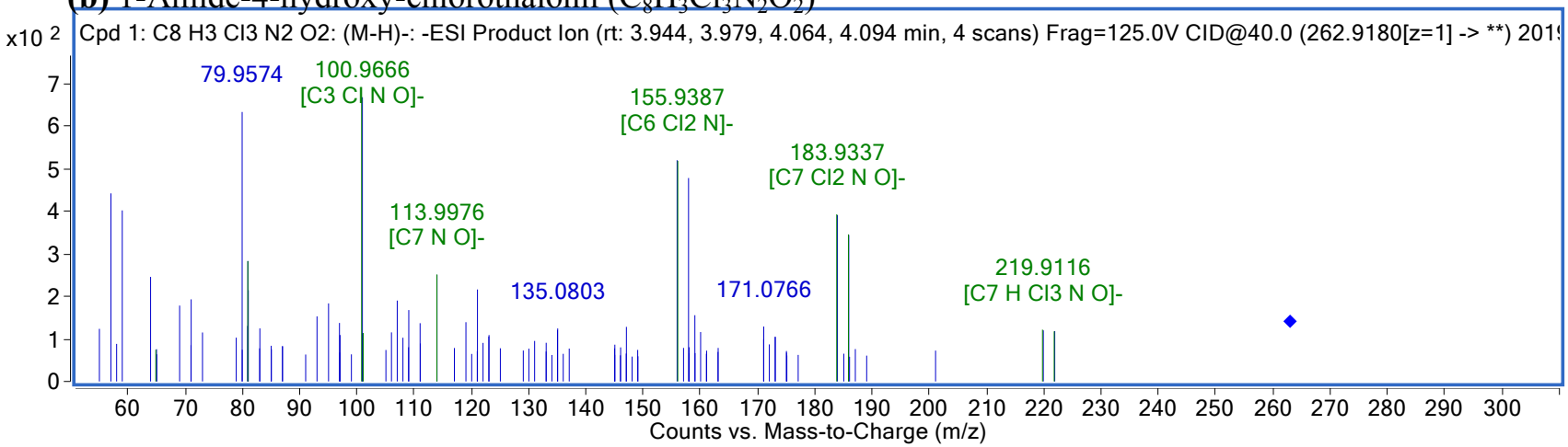
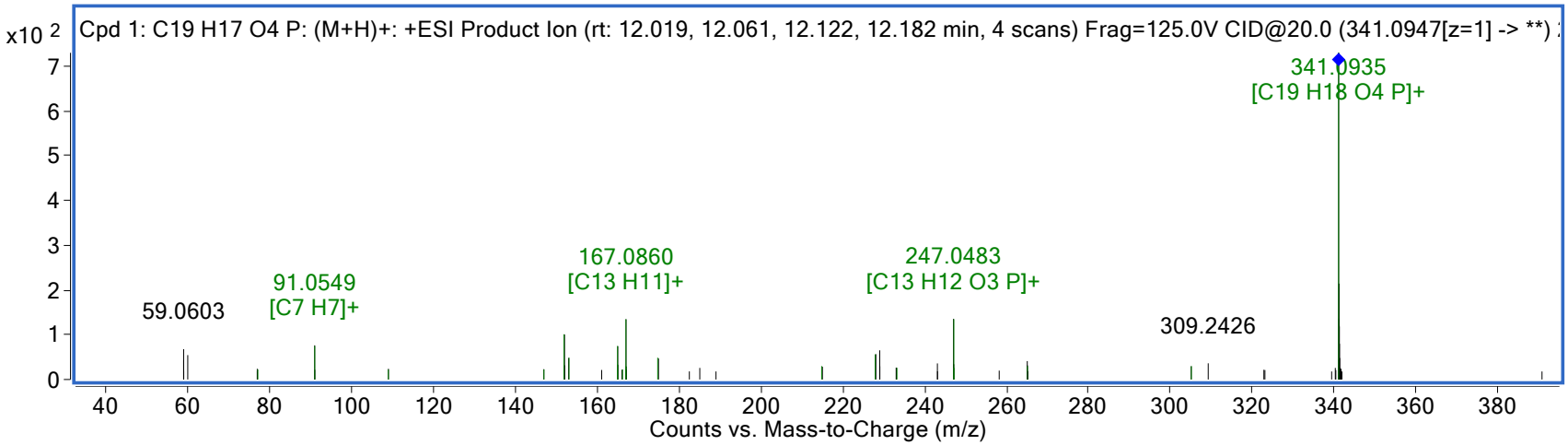


Figure S4. MS/MS spectra for level S3 tentatively identified contaminants. The blue diamonds in the spectra indicate the parent ions. **(a)**: 4-((4-Isopropoxyphenyl)sulfonyl)phenol featured by a fragment related to bisphenol S ($C_{12}H_8O_4S$). **(b)**:1-Amide-4-hydroxy-chlorothalonil lost a $CONH_2$ (amide group) and a Cl atom.

(c) Tolyl diphenyl phosphate ($C_{19}H_{17}O_4P$)



(d) Phenyl di(p-tolyl) phosphate ($C_{20}H_{19}O_4P$)

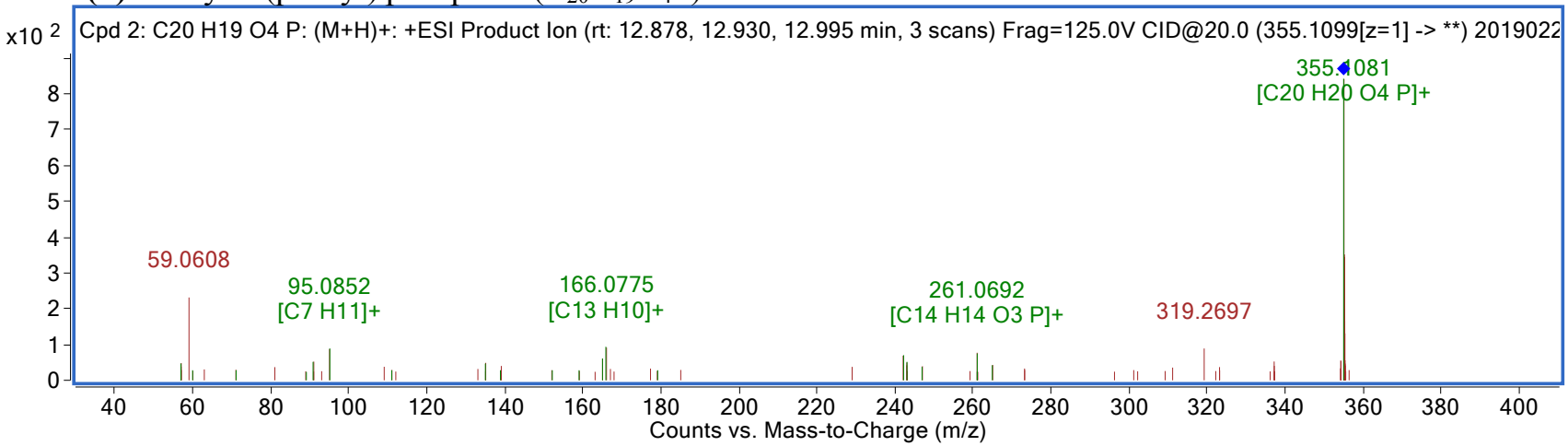
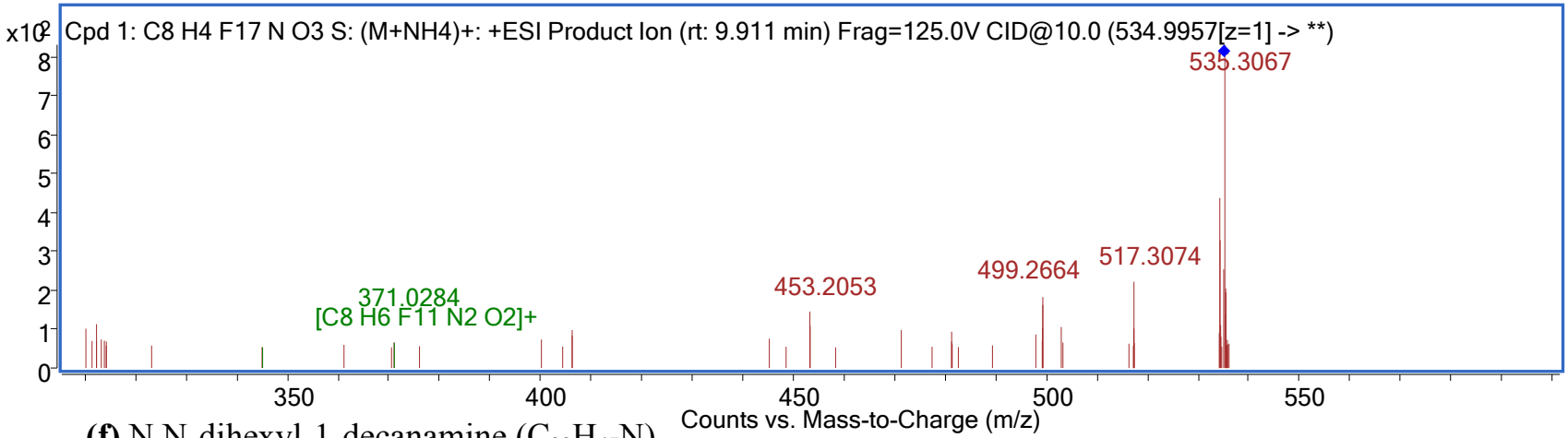


Figure S4 (continued). MS/MS spectra for level S3 tentatively identified contaminants. Both **(c)** tolyl diphenyl phosphate and **(d)** phenyl di(p-tolyl) phosphate lost C6H6O (phenol group).

(e) Ammonium perfluorooctanesulfonate ($C_8H_4F_{17}NO_3S$)



(f) N,N-dihexyl-1-decanamine ($C_{22}H_{47}N$)

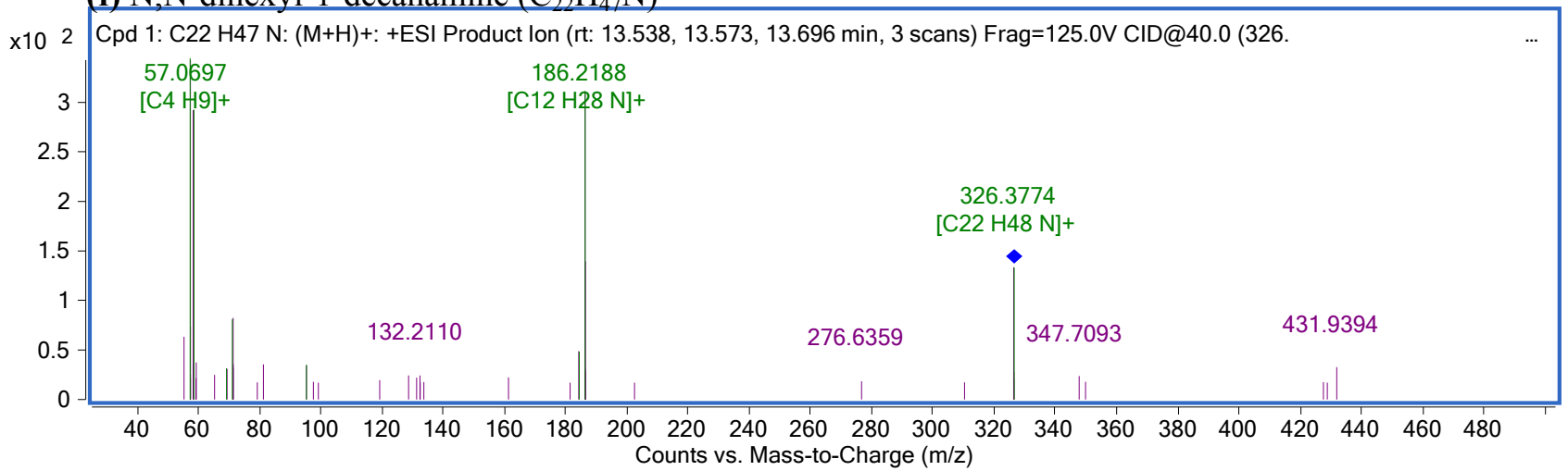
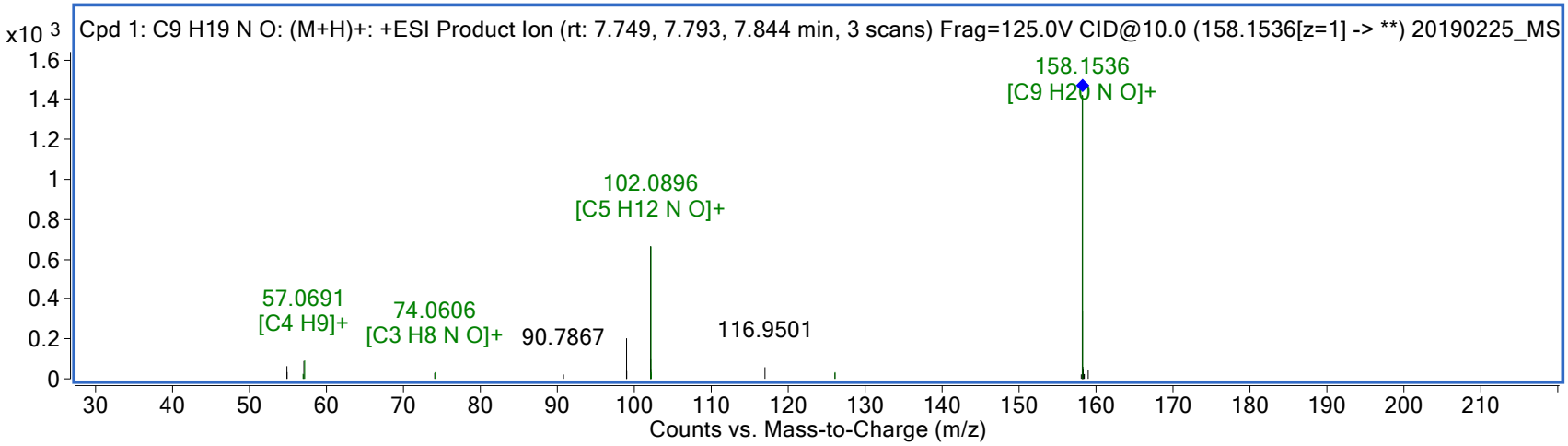


Figure S4 (continued). MS/MS spectra for level S3 tentatively identified contaminants. **(e)**: the fragment represents PFOS (*m/z* 499) and PFOS ammonium salt (*m/z* 517) can be seen. **(f)**: the loss of $C_{10}H_{20}$ represent the decyl group (calculated by CSI:FingerID).

(g) N-butyl-3-pentanamide ($C_9H_{19}NO$)



(h) Octyldimethylamine oxide ($C_{10}H_{23}NO$)

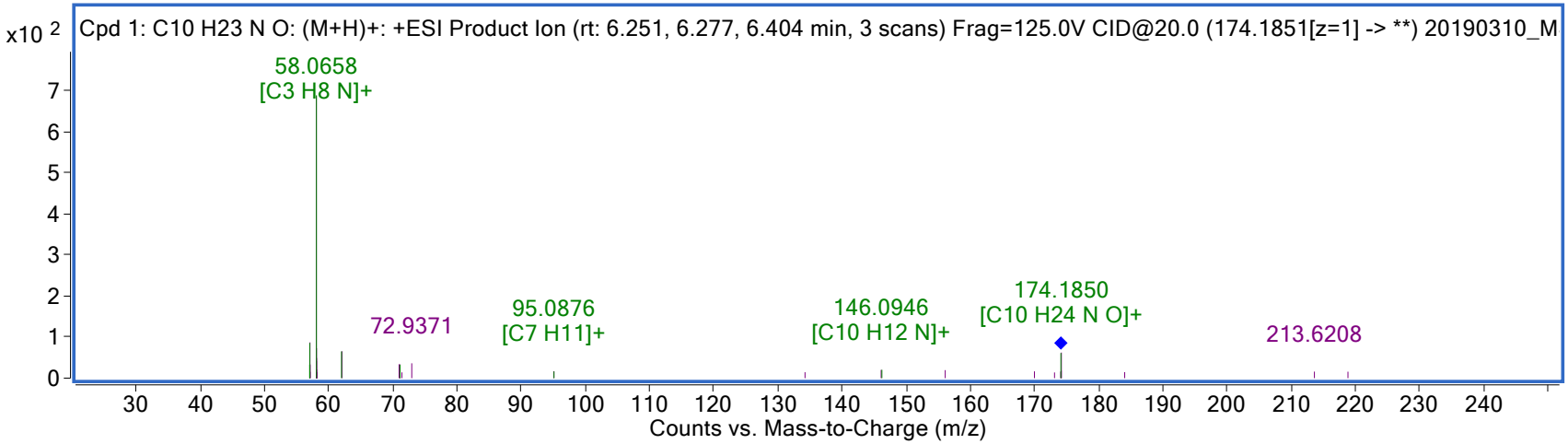
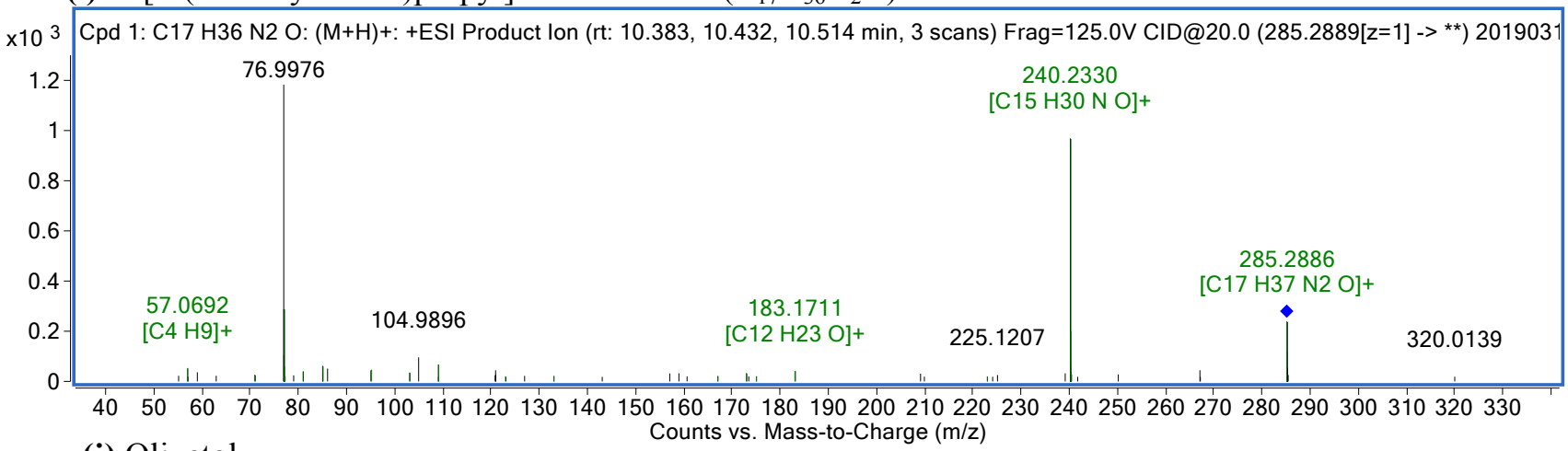


Figure S4 (continued). MS/MS spectra for level S3 tentatively identified contaminants. **(g)**: the lost C4H8 represented the butyl group, and $C_5H_{12}NO$ is the left pentanamide. **(h)**: octyldimethylamine oxide was proposed by *in silico* fragmentation.

(i) N-[3-(dimethylamino)propyl]dodecanamide ($C_{17}H_{36}N_2O$)



(j) Olivetol

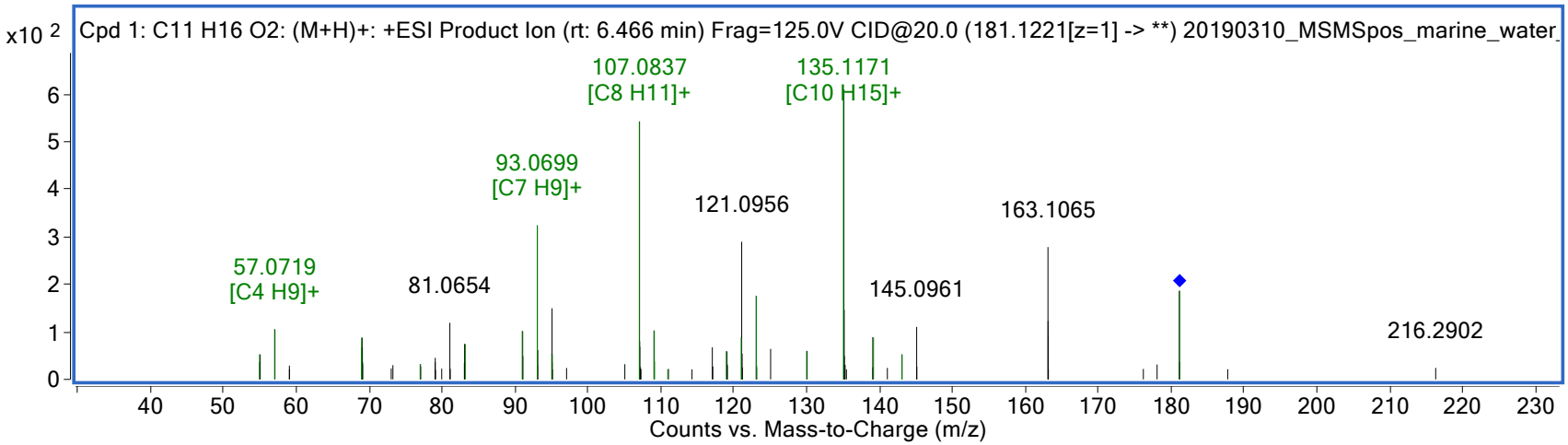
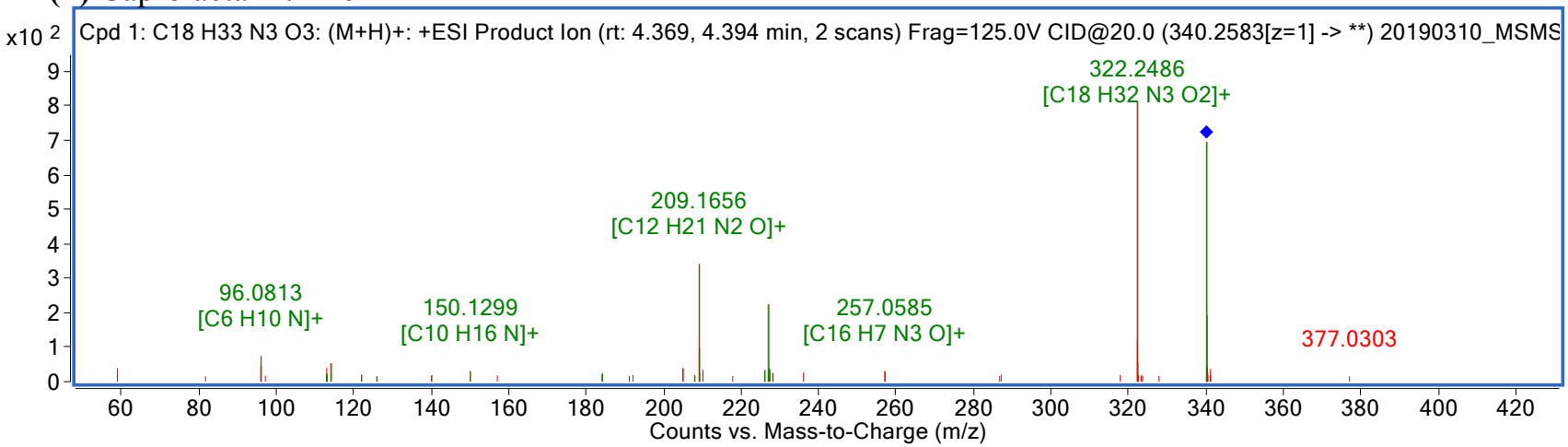


Figure S4 (continued). MS/MS spectra for level S3 tentatively identified contaminants. **(i)**: the loss of C2H7N represented dimethylamino group. **(j)**: olivetol was proposed by *in-silico* fragmentation.

(k) Caprolactam trimer



(l) 1-[5-(1-Aminoethyl)-2-methylsulfanyl-4-propylsulfanylphenyl]ethanamine

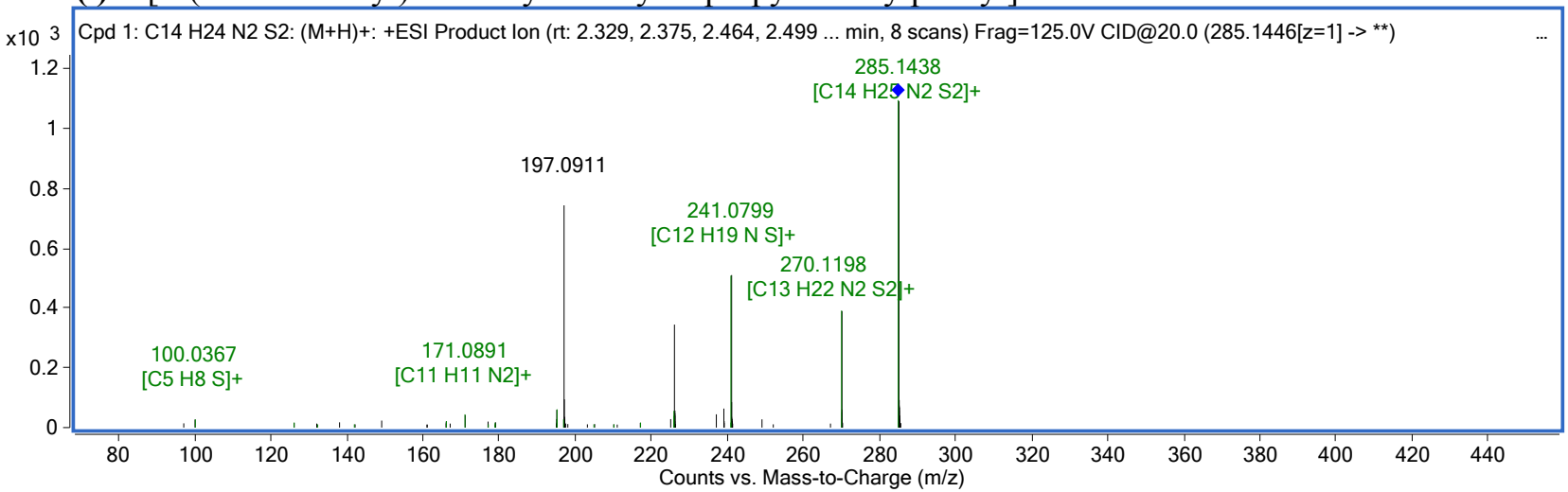


Figure S4 (continued). MS/MS spectra for level S3 tentatively identified contaminants. **(k)**: a regular pattern of losing $C_6H_{11}NO$ (caprolactam). **(l)**: consecutive loss of CH_3 and C_2H_6N represented methyl and aminoethyl groups.

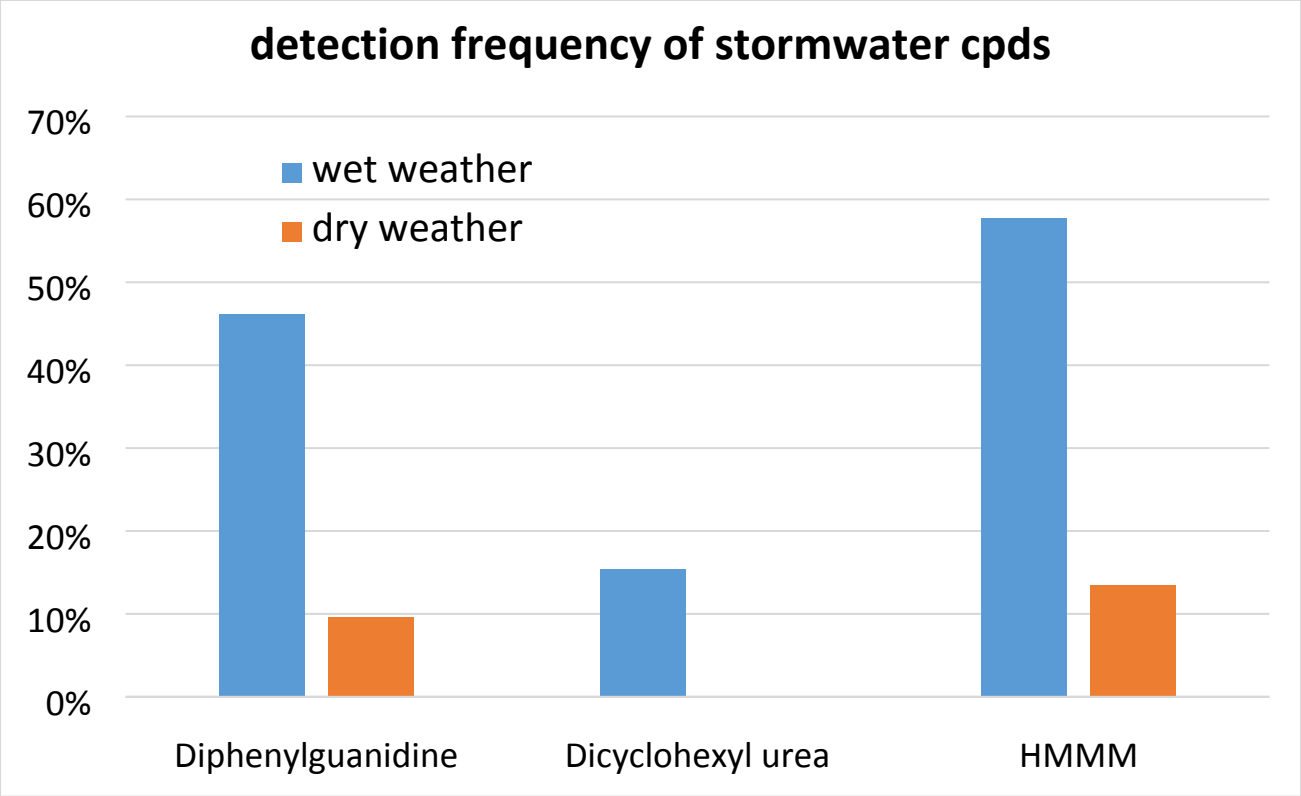


Figure S5. Detection frequencies of contaminants associated with stormwater runoff in wet weather sampling (April, October) and dry weather sampling (May, June, August). The detection frequency of stormwater-associated CECs was much higher in the wet weather sampling compared to dry weather, illustrating the potential range of impacts of stormwater in the marine nearshore. HMMM: Hexa(methoxymethyl)melamine.

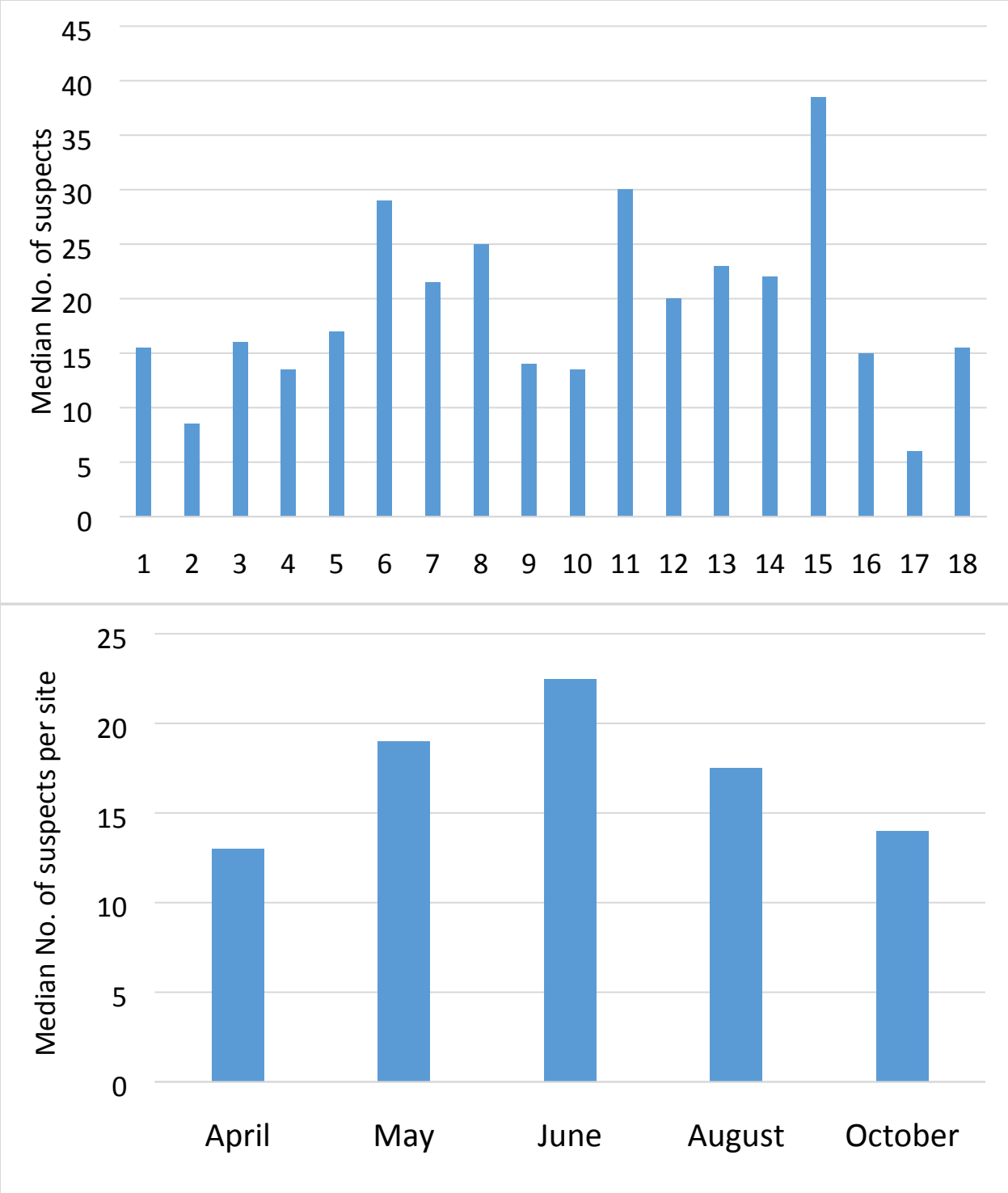


Figure S6. Temporal and spatial trends of contaminants in Puget Sound, measured by: **(a)** median numbers of matched suspects by sites. Site number, names, and locations are shown in Table S1 and Figure 1. **(b)** median numbers of matched suspects per site by dates. The suspects were from a standard compound database (CUW standards database) including wastewater-derived contaminants, stormwater contaminants, and chemicals from EPA's ToxCast library, which implies anthropogenic impacts. The database includes a total of ~1100 compounds.

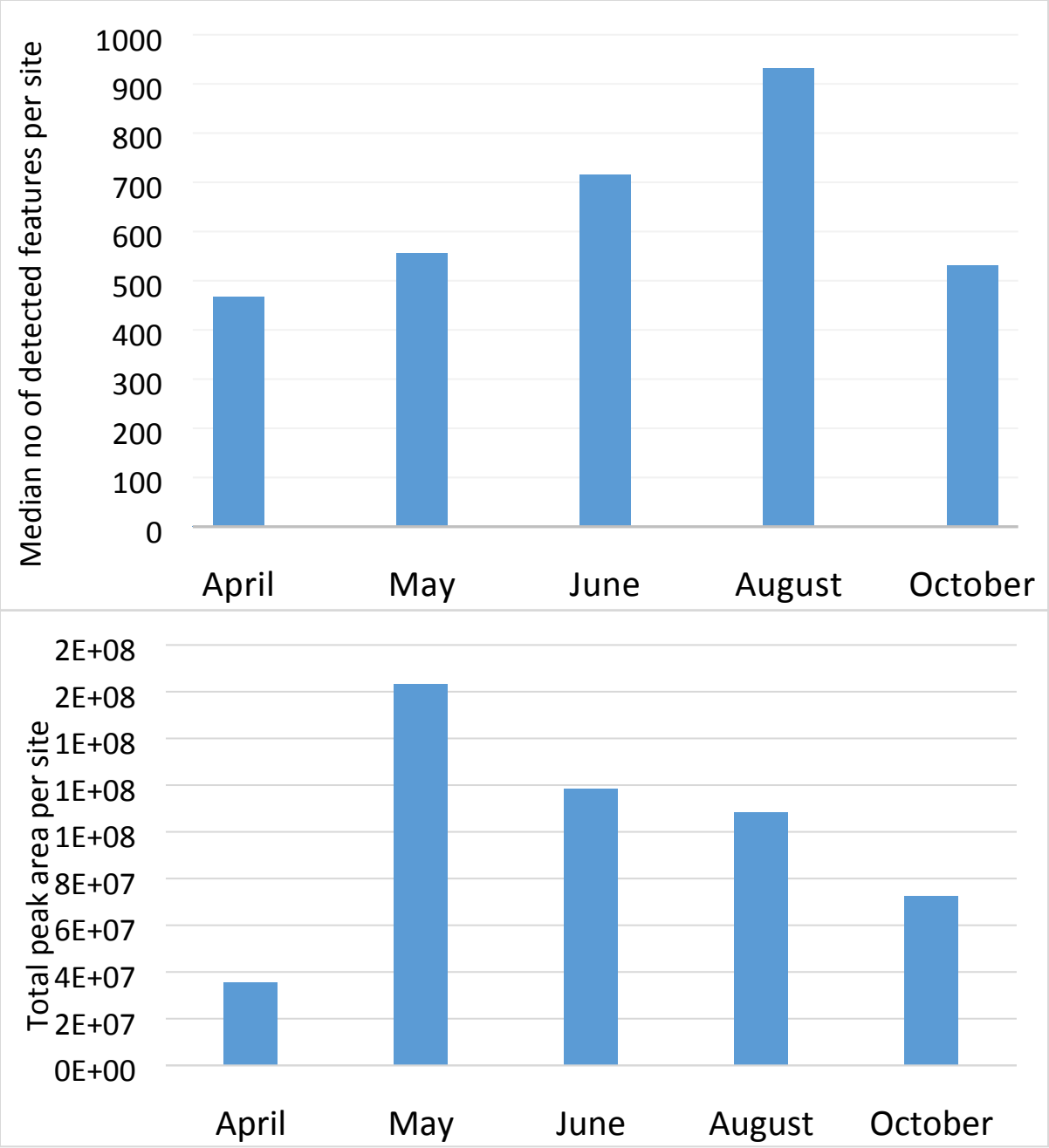


Figure S6, continued. **(c)** median number of features detected across all sites for each sampling event; **(d)** median total peak areas across all sites for each sampling event.