

## Supporting Information – Part 1 (SI-1)

### Apex marine predators and ocean health: proactive screening of halogenated organic contaminants reveals ecosystem indicator species

Jennifer M. Cossaboon<sup>a</sup>, Eunha Hoh<sup>a</sup>, Susan J. Chivers<sup>b</sup>, David W. Weller<sup>b</sup>, Kerri Danil<sup>b</sup>, Keith A. Maruya<sup>c</sup>, and Nathan G. Dodder<sup>a1</sup>

<sup>a</sup>School of Public Health, San Diego State University, San Diego, CA 92182; <sup>b</sup>Southwest Fisheries Science Center, National Oceanographic and Atmospheric Administration, La Jolla, CA 92037; and <sup>c</sup>Southern California Coastal Water Research Project Authority, Costa Mesa, CA 92626, USA

### Automated Data Reduction Strategy

#### Halogenation Filter Optimization

Data reduction software was developed based on the algorithm described by Pena-Abaurrea et al. (2014)<sup>1</sup>, which classified isotopic clusters in mass spectra as either halogenated (containing bromine and/or chlorine) or non-halogenated using eleven rules that examined the ion intensity ratios<sup>1</sup>. The original algorithm was re-written in the R programming language and optimized (as described below) by comparison against manual analysis. One sample from each species was selected (5 samples total), and each deconvoluted chromatographic peak and corresponding mass spectrum (average of 9210 features per sample) was manually inspected for the presence of chlorine and/or bromine.

The original algorithm was modified by the addition of three rules that increased the overall efficiency by minimizing the false positive rate. First, isotopic distributions identified by the original algorithm as halogenated were compared against theoretical bromine (n=1 to 10), chlorine (n=1 to 10), and bromine plus chlorine (all combinations of 1 to 5 bromines and 1 to 5 chlorines) using a dot product based spectral similarity score. The closest matching halogen

content was selected. If the experimental isotopic cluster had a similarity score  $< 0.85$  against the closest matching theoretical distribution, the experimental isotopic cluster was specified as non-halogenated. Second, any experimental isotopic clusters matching  $Cl_1$  were specified as non-halogenated. This simple cluster frequently matches non-halogenated ion clusters and lead to an unacceptably high false positive rate. Third, for a compound to be specified as halogenated, it was required that at least two isotopic clusters in the mass spectrum met the above criteria, since prior work demonstrated that bioaccumulative HOCs typically contain more than one halogen<sup>2</sup>.

The original algorithm identified an average of 2093 halogenated spectra/sample. Application of the three additional rules reduced the average number of spectra to 479 per sample. This reduction in false positives and the resulting average number of mass spectra per sample enabled non-targeted analysis of all 25 samples within a reasonable timeframe. An additional crosschecking procedure was developed that required mass spectra (with the same retention times) present in  $> 2$  samples to be manually searched in the remaining 23 samples. Crosschecking resulted in a final detection rate of 80-85% compared to a fully manual process (Table S3). It is recognized that the automated method had a higher false negative rate compared to full manual review, however compromise was required in order to meet the efficiency requirements of analyzing 25 samples. Core components of the R script are provided below (Code 1 and 2), and the full script is provided at <https://github.com/OrgMassSpec/IdentifyHalogenatedSpectra>.

### Compound Identification

GC×GC/TOF-MS data was processed within the LECO ChromaTOF mass spectrometer data system using automatic peak finding and electron impact fragmentation mass spectral deconvolution with a signal-to-noise threshold (S/N) of 50 or higher. The optimized halogenation filtering script was applied to all 25 samples for halogenation screening. The chromatographic peaks and corresponding mass spectra selected by the filter were manually reviewed and halogenated compounds were structurally identified by searches against existing mass spectral libraries<sup>2,3</sup> including the NIST 2014 Electron Impact Mass Spectral Library, matches to authentic reference standards, or manual interpretation. The crosschecking procedure was then performed on true positives appearing in more than two samples (Figure S4). Compound relative abundances were calculated using methods established by Shaul et al. (2014). Differences in contaminant loads between marine mammal species were analyzed via statistical comparisons in R<sup>4,5</sup> using hierarchical clustering displayed as heat maps<sup>6</sup>, as well as box and whisker diagrams<sup>7</sup> to display the distribution of compound structural class abundances.

**Table S1.** Structural class descriptions for HOCs detected in the five SCB marine mammal species.

Structural Class	Constituents	Source and Description
B/CDE	Brominated/chlorinated diethers	Limited data; previously detected in wastewater treatment plant drying lagoons <sup>8</sup> and e-waste recycling plants <sup>9</sup>
Brominated anisole		Mixed natural and anthropogenic <sup>2</sup>
Chlordane-related	Chlordane-related compounds	Legacy organochlorine insecticide formulation
Chlorinated benzene	Chlorinated benzenes	Solvent for pesticide formulations and chemical manufacturing
Chlorinated styrene	Chlorinated styrenes	Derived from styrene, an industrial chemical used in manufacturing (still in use)
DDT-related	Dichlorodiphenyltrichloroethane-related compounds	Legacy pesticide and associated degradation products
Dichlorobenzophenone	4,4'-dichlorobenzophenone	Degradation product of dicofol and DDT pesticides <sup>10,11</sup>
Dieldrin		Legacy insecticide
DMBP	Dimethyl bipyroles	Nitrogen-containing heterocyclic compounds with no known anthropogenic sources, suspected to be natural products <sup>12</sup>
HCH-related	Hexachlorocyclohexane-related compounds	Legacy insecticide
Heptachlor-related	Heptachlor-related compounds	Legacy insecticide
MBP	Methyl bipyroles	Nitrogen-containing heterocyclic compounds suspected to be natural products <sup>12</sup>
MeO-B/CDE	Methoxy brominated/chlorinated diphenyl ethers	Suspected natural product <sup>13</sup>
MeO-BDE	Methoxy brominated diphenyl ethers	Only known natural sources; isolated from marine algae, sponges, and bacteria <sup>14</sup>
MeO-CDE	Methoxy chlorinated diphenyl ethers	Unknown, hydroxyl analogues identified as impurities in technical pentachlorophenol <sup>15</sup>
MeO-PBB	Methoxy polybrominated biphenyls	Natural sources <sup>2</sup>
Methylenebistrichloroanisole	Methylene bis(trichloro)anisole	Methylated version of hexachlorophene disinfectant (banned in 1972)
Methylsulfonyl-PCB	Methylsulfonyl polychlorinated biphenyls	PCB metabolites <sup>16</sup>
Mirex-related	Mirex-related compounds	Legacy pesticide and flame retardant
PBB	Polybrominated biphenyls	Flame retardants
PBDE	Polybrominated diphenyl ethers	Flame retardants
PBHD	Polybrominated hexahydroxanthenes	Only known natural sources <sup>17</sup>
PCT	Polychlorinated terphenyls	Flame retardants <sup>18</sup>
Pyrrolidinecarbonyl chloride		Synthetic precursor to a carbamate fungicide <sup>19</sup>

TCPM	Tris(4-chlorophenyl)methanes	Possible industrial sources, suspected minor constituent in technical DDT and dicofol <sup>20,21</sup>
TCPMOH	Tris(4-chlorophenyl)methanol	Presumed primary breakdown product of TCPM <sup>22</sup>
Toxaphene	Toxaphenes	Legacy pesticide
Unknown	Miscellaneous unknown compounds	Unknown
Unknown-3	Unknown compounds with similar fragmentation ion clusters	Previously detected in Atlantic dolphin blubber <sup>2</sup> , fish oil supplements <sup>23</sup> , a Mediterranean sponge <sup>17</sup> , and passive samplers deployed along the Great Barrier Reef <sup>24</sup>
Unknown-4	Unknown compounds with similar fragmentation ion clusters	Suspected to be polychlorinated diphenyl ethers or hydroxyl PCBs <sup>2</sup>
Unknown-5 to Unknown-8	Unknown compounds with similar fragmentation ion clusters	Unknown

**Table S2.** Sample collection details. All samples were acquired from the Marine Mammal Life History Archive of the NOAA National Marine Fishery Service, Southwest Fisheries Science Center (SWFSC), in La Jolla, California.

Sample ID	Specimen ID	Common Name	Source*	Length (cm)	Standard Length?	Collection Year	Latitude	Longitude
MMCP 1	DWJ0022	Harbor seal	Fishery	121	Y	1991	33.096666	-117.343333
MMCP 2	MAI0004	Risso's dolphin	Fishery	252	Y	1993	32.418333	-118.146666
MMCP 3	DFP0099	Risso's dolphin	Fishery	254	N	1993	32.728333	-118.981666
MMCP 4	SCB0076	Harbor seal	Fishery	97	Y	1992	33.6	-117.945
MMCP 5	RAC0011	California sea lion	Fishery	204	Y	1997	32.656666	-118.028333
MMCP 6	SMC0023	Long-beaked common dolphin	Fishery	218	Y	1997	33.321666	-117.886666
MMCP 7	AXR0005	Risso's dolphin	Fishery	203	Y	1997	32.861666	-117.501666
MMCP 8	CMN0011	Long-beaked common dolphin	Fishery	235	N	1997	34.968333	-120.918333
MMCP 9	KSS0010	Short-beaked common dolphin	Fishery	194	Y	2008	32.723333	-117.43
MMCP 10	LSE0064	Short-beaked common dolphin	Fishery	188	Y	1996	32.483333	-118.725
MMCP 11	MSK0045	Short-beaked common dolphin	Fishery	187	Y	1991	32.323333	-117.795
MMCP 12	JCP0025	California sea lion	Fishery	224	n.d.	2000	32.57	-117.99
MMCP 13	KMP0047	Harbor seal	Fishery	110	Y	1994	34.116666	-119.268333
MMCP 14	MJK0048	Short-beaked common dolphin	Fishery	188	Y	1999	32.715	-117.85
MMCP 15	BXD0026	California sea lion	Fishery	227	N	2004	32.645	-117.635
MMCP 16	MSK0056	Risso's dolphin	Fishery	219	n.d.	1992	34.435	-121.061666
MMCP 17	KXD0197	Long-beaked common dolphin	Stranding	226.7	y	2011	32.62575	-117.139166
MMCP 18	JBG0017	California sea lion	Fishery	209	Y	1997	33.518333	-117.88
MMCP 19	JNB0038	California sea lion	Fishery	207	Y	1999	32.598333	-117.778333
MMCP 20	SWC-Dc-1401B	Long-beaked common dolphin	Stranding	211.8	y	2014	32.865565	-117.25538
MMCP 21	KXD0066	Long-beaked common dolphin	Stranding	200	N	2004	33.216666	-117.383333
MMCP 22	FHM0001	Harbor seal	Fishery	99	Y	1990	32.551666	-117.171666
MMCP 23	RGH0129	Harbor seal	Fishery	92	Y	1994	34.173333	-119.373333
MMCP 24	DKG0002	Short-beaked common dolphin	Fishery	196	Y	2000	32.331666	-117.725
MMCP 25	EAF0001	Risso's dolphin	Fishery	243	Y	1993	34.121666	-121.946666

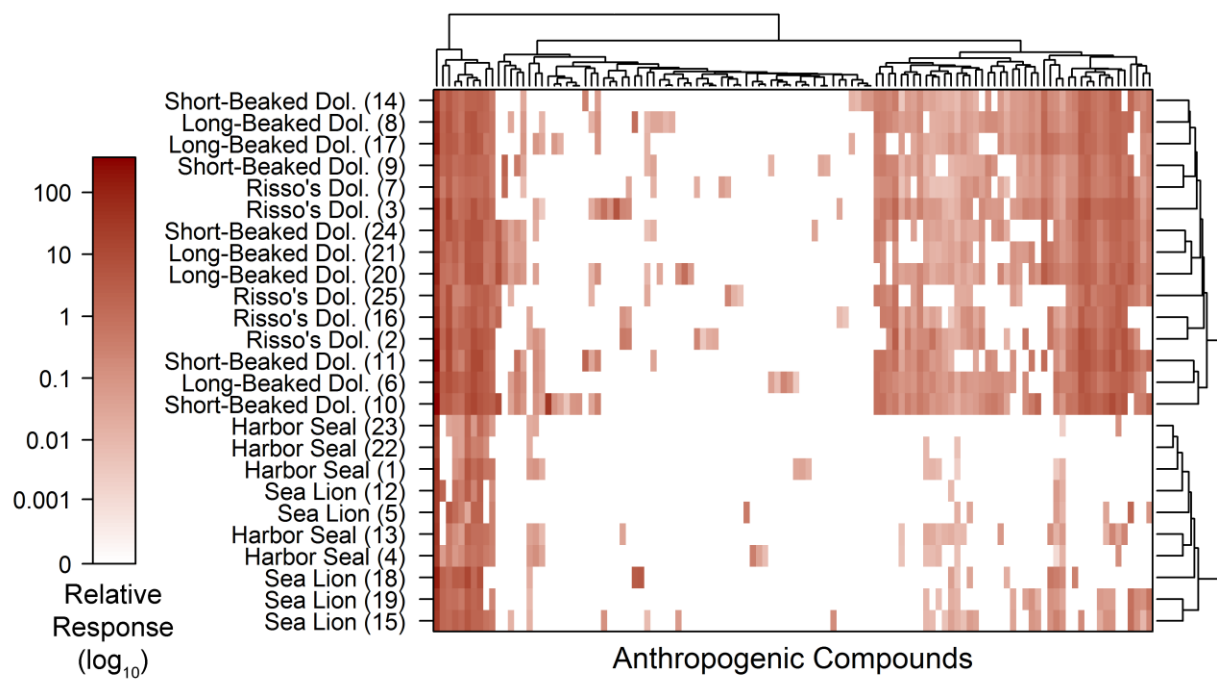
\*Fishery = fishery by-catch

n.d. = no data

**Table S3.** Comparison of total number of HOCs (including PCBs) identified by manual analysis, the optimized filtering script, and cross-checking of filtering script results in a subset of five samples.

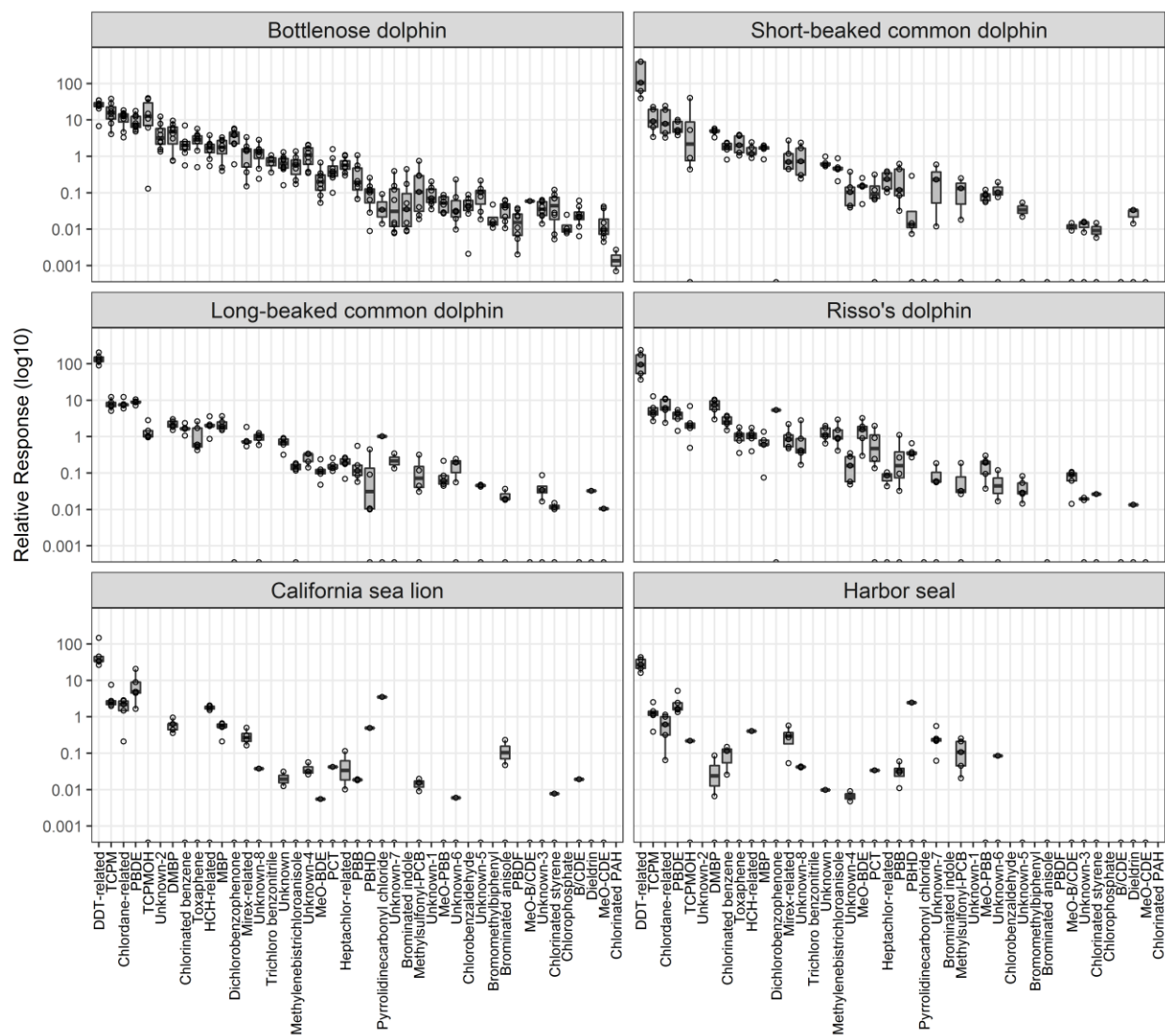
Species	Sample	No. HOCs Identified Manually	No. HOCs Identified by Filtering Script	No. HOCs Identified after Cross-checking
Long-beaked common dolphin	MMCP 8	159	108 (68%)	134 (84%)
Short-beaked common dolphin	MMCP 10	170	121 (71%)	144 (85%)
Risso's dolphin	MMCP 2	149	95 (64%)	119 (80%)
California sea lion	MMCP 15	117	76 (65%)	94 (80%)
Harbor seal	MMCP 13	67	41 (61%)	57 (85%)
Pacific common bottlenose dolphin	n=8	*206		

\*This number represents the average from eight samples analyzed by Shaul et al., 2014 and excludes PCBs.

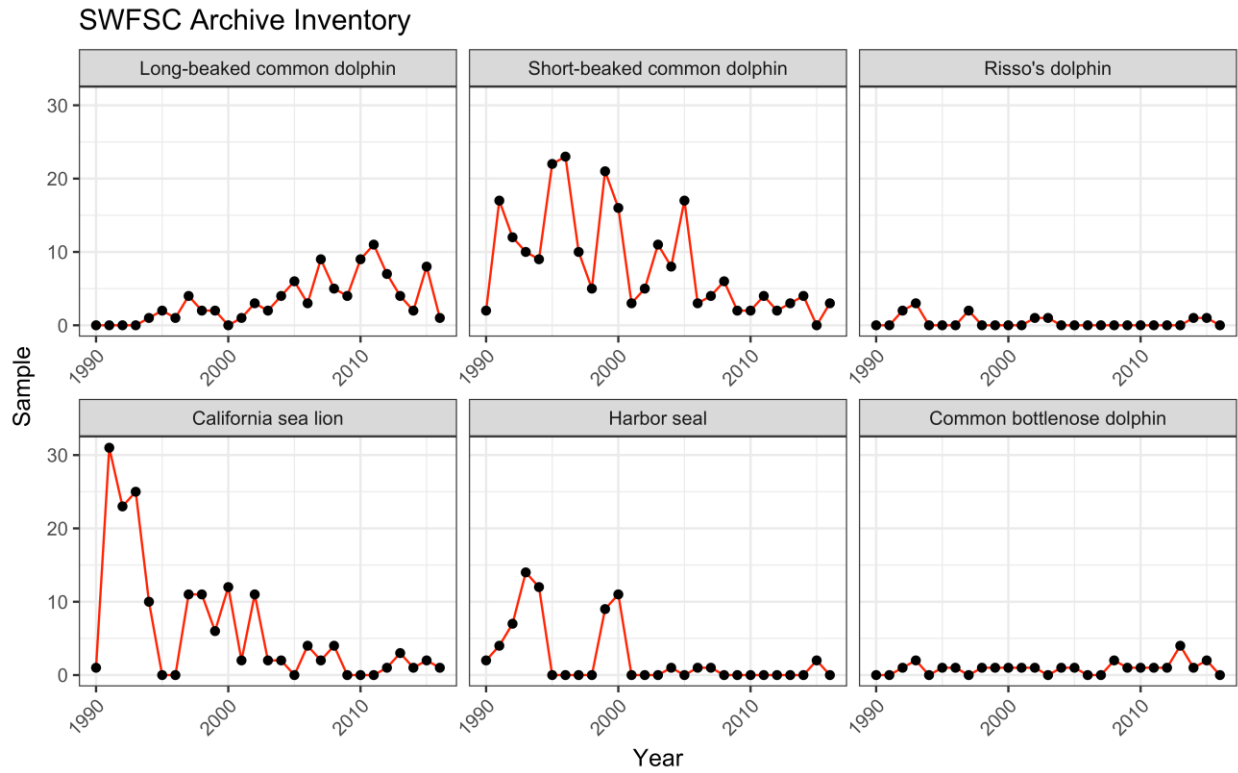


**Figure S1.** Heat map illustrating hierarchical clustering results for all 25 samples based upon detection and relative abundance of HOCs from anthropogenic sources.

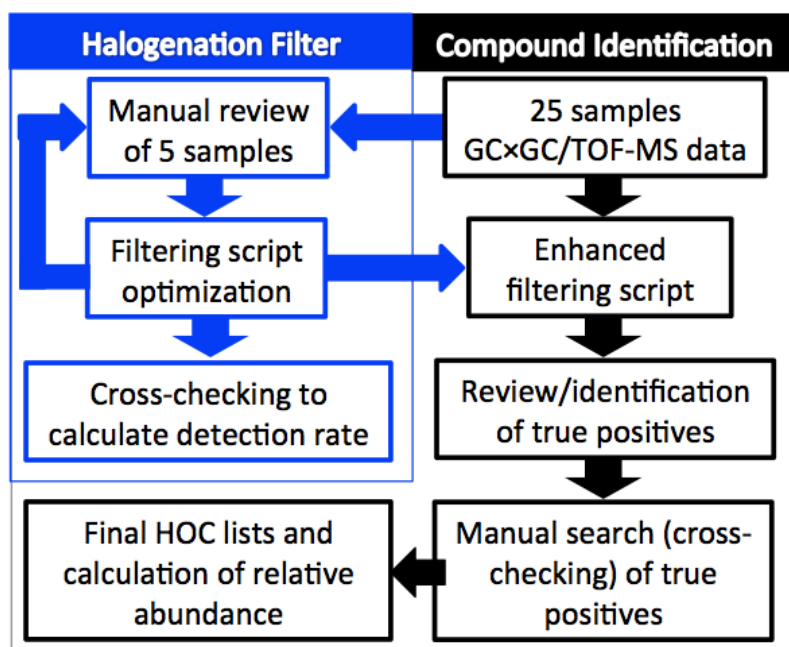




**Figure S2.** Summed relative abundances of HOCs by structural class for each species, including Pacific common bottlenose dolphin profile from Shaul et al., 2014.



**Figure S3.** Inventory of full-depth blubber specimens from male individuals (all age groups, stranding and fisheries sources) of the six proposed sentinel species collected from the Southern California Bight (1990–2016).



**Figure S4.** Flowchart of automated data handling and compound identification procedure.

**Code 1. Halogenation Search Filter. Full description of a working script is available at <http://orgmassspec.github.io/IdentifyHalogenationScript.html>.**

```
ProcessSpectrum <- function(x) {  
  
  # Normalize peak intensity to percentage of max peak.  
  x$percentIntensity <- with(x, intensity / max(intensity) * 100)  
  
  x$Filter <- FALSE  
  
  x <- x[x$mz > 100, ]  
  
  for(i in (nrow(x) - 6):5) {  
  
    if(x$percentIntensity[i] > 3) {  
  
      # Calculate peak ratios.  
  
      peakRatio1 <- x$percentIntensity[i + 1] / x$percentIntensity[i]  
      peakRatio2 <- x$percentIntensity[i + 2] / x$percentIntensity[i]  
      peakRatio3 <- x$percentIntensity[i + 3] / x$percentIntensity[i]  
      peakRatio4 <- x$percentIntensity[i + 4] / x$percentIntensity[i]  
      peakRatio5 <- x$percentIntensity[i + 5] / x$percentIntensity[i]  
      peakRatio6 <- x$percentIntensity[i + 6] / x$percentIntensity[i]  
      peakRatioa <- x$percentIntensity[i - 1] / x$percentIntensity[i]  
      peakRatiob <- x$percentIntensity[i - 2] / x$percentIntensity[i]  
      peakRatioc <- x$percentIntensity[i - 4] / x$percentIntensity[i]  
  
      # Determine if peak ratios pass rules.  
  
      if(peakRatio1 < 0.5 & peakRatio1 > 0.003 & # Rule 1  
        peakRatio2 > 0.225 & peakRatio2 <= 1 & # Rule 2  
        peakRatio1 < peakRatio2 & # Rule 3  
        peakRatio3 < 0.5 & peakRatio3 > 0.001 & # Rule 4  
        peakRatio3 < peakRatio1 & # Rule 5  
        peakRatio4 < peakRatio2 & # Rule 6  
        peakRatio5 < peakRatio1 & # Rule 7  
        peakRatio6 < peakRatio2 & # Rule 8  
        peakRatioa < 0.5 & # Rule 9  
        peakRatiob < 1 & # Rule 10  
        peakRatioc < 1) { # Rule 11  
  
        x$Filter[i] <- TRUE  
  
      }  
  
    } # encloses filter  
  
  } # encloses for loop that iterates through each m/z  
  
  # Returns TRUE if compound is halogenated.  
}
```

## Code 2. Additional Rules.

```
x <- x[x$Theory_PercentIntensity != 0, ] #Error check

# Apply additional filters.

# Rule 1. Exclude C11 isotopic clusters.
x <- x[x$DistributionLabel != 'C11', ]

# Rule 2. Clusters must be at least 85% similar to theoretical distribution.
x <- x[x$SimilarityScore > 0.85, ]

# Rule 3. Select only the spectra with > 1 halogenated fragment per spectrum.
y <- setNames(aggregate(Expt_mz ~ Spectrum, data = x, FUN = length),
c('Spectrum', 'Count'))
countFilter <- y$Spectrum[y$Count > 1]
x <- x[x$Spectrum %in% countFilter, ]
```

## References

- (1) Pena-Abaurrea, M.; Jobst, K. J.; Ruffolo, R.; Shen, L.; McCrindle, R.; Helm, P. A.; Reiner, E. J. Identification of Potential Novel Bioaccumulative and Persistent Chemicals in Sediments from Ontario (Canada) Using Scripting Approaches with GC×GC-TOF MS Analysis. *Environ. Sci. Technol.* **2014**, *48* (16), 9591–9599.
- (2) Shaul, N. J.; Dodder, N. G.; Aluwihare, L. I.; Mackintosh, S. A.; Maruya, K. A.; Chivers, S. J.; Danil, K.; Weller, D. W.; Hoh, E. Nontargeted Biomonitoring of Halogenated Organic Compounds in Two Ecotypes of Bottlenose Dolphins (*Tursiops truncatus*) from the Southern California Bight. *Environ. Sci. Technol.* **2015**, *49* (3), 1328–1338.
- (3) Mackintosh, S. A.; Dodder, N. G.; Shaul, N. J.; Aluwihare, L. I.; Maruya, K. A.; Chivers, S. J.; Danil, K.; Weller, D. W.; Hoh, E. Newly Identified DDT-Related Compounds Accumulating in Southern California Bottlenose Dolphins. *Environ. Sci. Technol.* **2016**, *50* (22), 12129–12137.
- (4) R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017.
- (5) Wickham, H.; Francois, R.; Henry, L.; Müller, K. *Dplyr: A Grammar of Data Manipulation*; 2017.
- (6) Sarkar, D.; Andrews, F. *LatticeExtra: Extra Graphical Utilities Based on Lattice*; 2016.
- (7) Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag: New York, 2009.
- (8) La Guardia, M. J.; Hale, R. C.; Harvey, E.; Chen, D. Flame-Retardants and Other Organohalogenes Detected in Sewage Sludge by Electron Capture Negative Ion Mass Spectrometry. *Environ. Sci. Technol.* **2010**, *44* (12), 4658–4664.
- (9) Yu, Z.; Zheng, K.; Ren, G.; Wang, D.; Ma, S.; Peng, P.; Wu, M.; Sheng, G.; Fu, J. Identification of Monochloro-Nonabromodiphenyl Ethers in the Air and Soil Samples from South China. *Environ. Sci. Technol.* **2011**, *45* (7), 2619–2625.
- (10) Bumpus, J. A.; Aust, S. D. Biodegradation of DDT [1,1,1-Trichloro-2,2-Bis(4-Chlorophenyl)Ethane] by the White Rot Fungus *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* **1987**, *53* (9), 2001–2008.
- (11) Thiel, A.; Guth, S.; Böhm, S.; Eisenbrand, G. Dicofol Degradation to p,P'-Dichlorobenzophenone – A Potential Antiandrogen. *Toxicology* **2011**, *282* (3), 88–93.
- (12) Teuten, E. L.; Pedler, B. E.; Hangsterfer, A. N.; Reddy, C. M. Identification of Highly Brominated Analogues of Q1 in Marine Mammals. *Environ. Pollut.* **2006**, *144* (1), 336–344.
- (13) Malmvärn, A.; Marsh, G.; Kautsky, L.; Athanasiadou, M.; Bergman, Å.; Asplund, L. Hydroxylated and Methoxylated Brominated Diphenyl Ethers in the Red Algae *Ceramium tenuicorne* and Blue Mussels from the Baltic Sea. *Environ. Sci. Technol.* **2005**, *39* (9), 2990–2997.
- (14) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. Biosynthesis of Polybrominated Aromatic Organic Compounds by Marine Bacteria. *Nat. Chem. Biol.* **2014**, *10* (8), 640–647.

- (15) Deinzer, M.; Lamberton, J.; Griffin, D.; Miller, T. Isolation of Hydroxynonachlorodiphenyl Ethers from Pentachlorophenol. Mass Spectrometry of Hydroxynonachlorodiphenyl Ethers and Their Methylated Derivatives. *Biol. Mass Spectrom.* **1978**, *5* (10), 566–571.
- (16) Fängström, B.; Athanasiadou, M.; Grandjean, P.; Weihe, P.; Bergman, A. Hydroxylated PCB Metabolites and PCBs in Serum from Pregnant Faroese Women. *Environ. Health Perspect.* **2002**, *110* (9), 895.
- (17) Melcher, J.; Janussen, D.; Garson, M. J.; Hiebl, J.; Vetter, W. Polybrominated Hexahydroxanthene Derivatives (PBHDs) and Other Halogenated Natural Products from the Mediterranean Sponge *Scalariispongia scalaris* in Marine Biota. *Arch. Environ. Contam. Toxicol.* **2007**, *52* (4), 512–518.
- (18) Filyk, G. PCTs\_UECE.Pdf. *U. N. Econ. Comm. Eur.* **2011**, 1–29.
- (19) Katz, L. E. Selected 2-Trichloromethyl-4-Pyrimidinyl Carbamates and Their Use as Fungicides. EP0100795, 1984.
- (20) Buser, H. R. DDT, a Potential Source of Environmental Tris(4-Chlorophenyl)Methane and Tris(4-Chlorophenyl)Methanol. *Environ. Sci. Technol.* **1995**, *29* (8), 2133–2139.
- (21) de Boer, J. Environmental Distribution and Toxicity of Tris(4-Chlorophenyl)Methanol and Tris(4-Chlorophenyl)Methane. In *Reviews of Environmental Contamination and Toxicology*; 1997; Vol. 150, pp 95–106.
- (22) Zook, D. R.; Buser, H. R.; Bergqvist, P.-A.; Rappe, C.; Olsson, M. Detection of Tris(Chlorophenyl)Methane and Tris(4-Chlorophenyl)Methanol in Ringed Seal (*Phoca hispida*) from the Baltic Sea. *Ambio* **1992**, *21* (8), 557–560.
- (23) Hoh, E.; Lehotay, S. J.; Mastovska, K.; Ngo, H. L.; Vetter, W.; Pangallo, K. C.; Reddy, C. M. Capabilities of Direct Sample Introduction–Comprehensive Two-Dimensional Gas Chromatography–Time-of-Flight Mass Spectrometry to Analyze Organic Chemicals of Interest in Fish Oils. *Environ. Sci. Technol.* **2009**, *43* (9), 3240–3247.
- (24) Rosenfelder, N.; Van Zee, N. J.; Mueller, J. F.; Gaus, C.; Vetter, W. Gas Chromatography/Electron Ionization-Mass Spectrometry-Selected Ion Monitoring Screening Method for a Thorough Investigation of Polyhalogenated Compounds in Passive Sampler Extracts with Quadrupole Systems. *Anal. Chem.* **2010**, *82* (23), 9835–9842.