nature geoscience

Analysis

https://doi.org/10.1038/s41561-026-01928-z

Widespread presence of anthropogenic compounds in marine dissolved organic matter

Received: 29 March 2025

Accepted: 21 January 2026

Published online: 16 March 2026

Check for updates

A list of authors and their affiliations appears at the end of the paper

Chemical pollution has profound impacts on marine ecosystem health and services. Most investigations of the distributions of anthropogenic organic chemicals (xenobiotics) have been regionally focused, which limits our understanding of the extent of chemical pollution in the world's largest biome. To address this gap, we mapped the presence of xenobiotics across marine ecosystems. Here we present a meta-analysis of 21 public non-targeted tandem mass spectrometry datasets, which are not restricted to a predefined set of compounds but rather capture thousands of chemicals. These datasets comprise 2,315 seawater samples, spanning coastal to open ocean environments across three ocean basins. Our analysis revealed that common pollutants such as pesticides and pharmaceuticals were predominantly detected in estuaries and coastal areas but declined with distance from shore, whereas industrial chemicals and additives, including polyalkylene glycols, phthalates and organophosphates, were widely distributed across marine ecosystems. A total of 248 annotated xenobiotic features contributed a median of 2% to the total detected peak area per sample. We observed highest median levels of xenobiotic contribution in coastal datasets (up to 20%) and lowest levels (0.5%) in open ocean datasets, which indicates that anthropogenic organic substances contribute substantially to the dissolved organic matter pool in the surface ocean.

Population growth, urbanization, industrialization and consumerism have caused an unprecedented release of pollutants into marine environments, posing severe risks to ecosystem health, including disruptions in marine food webs, loss of biodiversity and the degradation of critical ecosystem services[1–5]. The global influx of anthropogenic organic molecules (xenobiotics) into some ecosystems has been suggested to reach similar levels as natural organic matter (NOM), which may directly influence global carbon cycling[6]. Despite these substantial impacts, we still have little knowledge about the distribution, persistence and ecological impact of xenobiotics in marine systems.

Liquid chromatography high-resolution tandem mass spectrometry (LC-HR-MS/MS) has substantially advanced our ability to detect and characterize xenobiotics within the dissolved organic matter (DOM) pool with high sensitivity and structural resolution[7–11]. However, the analysis via LC-HR-MS/MS is usually limited to small spatial scales[12,13], which prohibits the inference of ecosystem-wide patterns of xenobiotics within and between marine environments. The meta-analysis of LC-HR-MS/MS data from different studies and laboratories offers new opportunities to assess global DOM composition and planetary-scale patterns of anthropogenic compounds[14].

To characterize the distribution of xenobiotics in marine systems, we employed an analysis strategy for comparing non-targeted LC-HR-MS/MS studies across publicly available marine DOM datasets. Twenty-one public datasets comprising 2,315 samples that were acquired by three laboratories, using Quadrupole-Orbitrap instruments with standardized chromatographic methods and MS/MS settings, were included in our analysis.

These data enabled us to map out chemical pollutants in environments across the central and eastern Pacific, Indian and North Atlantic oceans, including the Baltic and Caribbean seas (Fig. 1). Datasets were

e-mail: dpetras@ucr.edu

https://doi.org/10.1038/s41561-026-01928-z

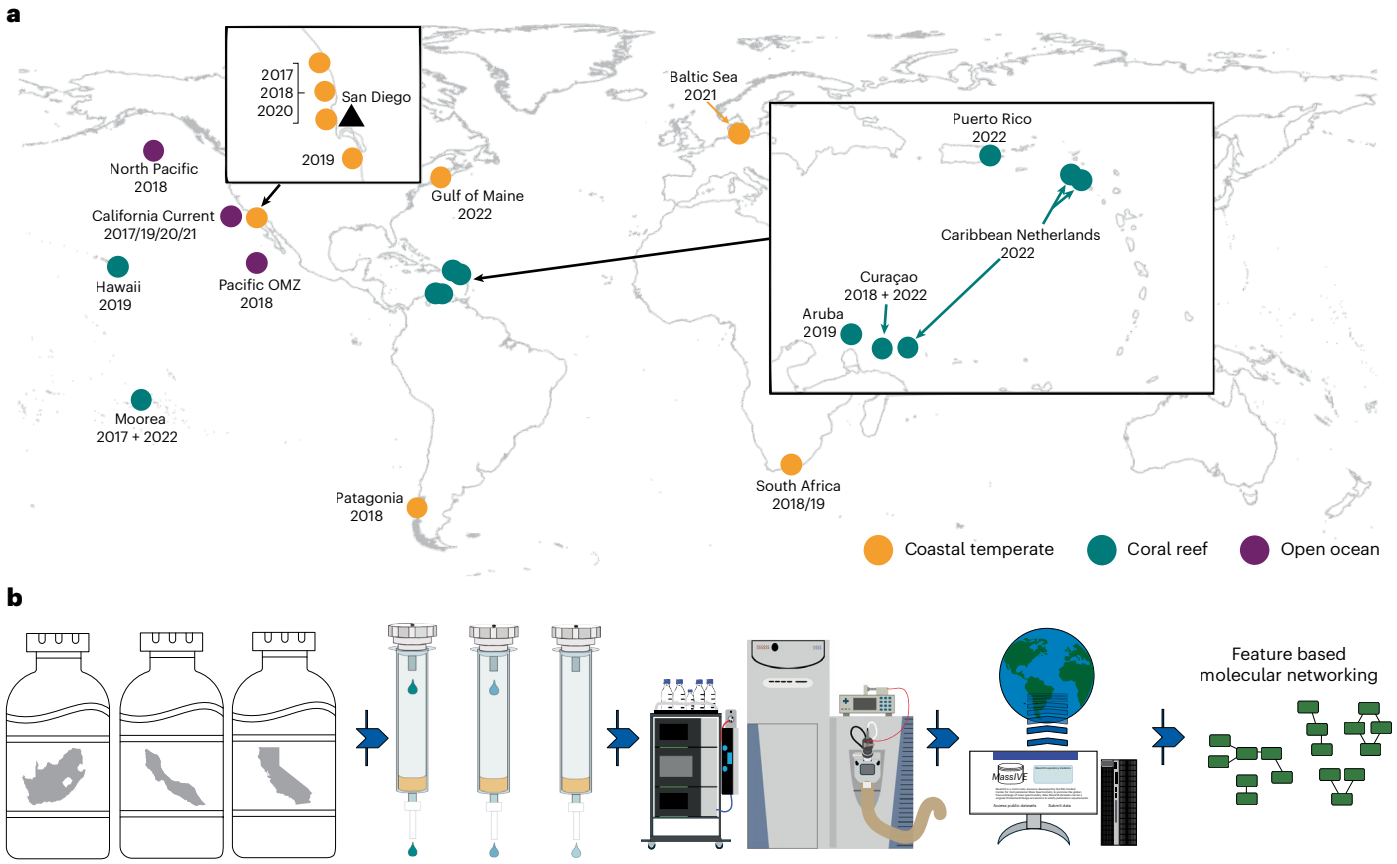


Fig. 1 | Sampling and analytical workflow. a, Map of sampling sites of the 21 public tandem mass spectrometry datasets included in this study. **b**, Schematic workflow showing how all datasets were prepared from marine water extracted with solid-phase extraction, followed by positive mode analysis on the Q Exactive platform, data deposition on MassIVE and subsequent analysis through FBMN. Colours in the map denote marine ecosystem type. OMZ, oxygen minimum zone. The map was generated with QGIS with Natural Earth shapefiles (https://www.naturalearthdata.com).

allocated to 'coastal temperate', 'coral reef' and 'open ocean' categories based on whether the majority of samples were collected in estuaries or near-shore on the coastal shelf in temperate regions ('coastal temperate'), from tropical or subtropical coastal regions with extensive benthic coral reef coverage ('coral reef') or from areas mostly off the coastal shelf ('open ocean'). Our findings reveal that distinct classes of xenobiotics are widespread throughout marine ecosystems and contribute substantially to the marine DOM pool.

Mapping xenobiotic diversity across marine habitats

To generate an overview of xenobiotics presence and abundance within the selected public LC-HR-MS/MS data, we processed all files with the feature-based molecular networking (FBMN) workflow, using the software tool mzmine for consensus feature detection[15,16] and the GNPS2 (gnps2.org) platform for spectral library matching and molecular networking[17]. FBMN is based on feature alignment between samples using retention time and MS1 m/z tolerances, followed by annotation and networking using MS/MS spectra. Xenobiotic features were annotated via MS/MS library matching (Level 1 and 2 identification[18]) to compounds with known industrial, pharmaceutical or agricultural uses and no reasonable natural source explaining their presence in marine DOM samples. After data processing, we first confirmed that overall feature peak areas were stable and comparable among the different datasets (Supplementary Fig. 1). Inspecting a set of commonly observed xenobiotics (DEET, diethyl phthalate and cocamidopropyl betaine), we observed excellent mass accuracy (< 2.5 ppm) and retention time stability (< 0.1 min) within the datasets. Retention time variance between datasets was moderate but within our FBMN feature finding tolerance (< 1.7 min). Stability of feature intensity was comparable between ecosystem types (Supplementary Figs. 2–5).

To identify the most prominent xenobiotics in our datasets, we manually inspected all spectrum-library matches (mirror matches in Supplemental Fig. 6) and assessed the degree to which they were shared among the 21 datasets (Fig. 2b). The presence of most xenobiotic features was confined to small intersection sizes between datasets, meaning only a few xenobiotic features shared the exact same distribution pattern across all datasets. To control for background contaminants from the laboratories and sampling equipment (for example, plasticizers from laboratory consumables), we removed features found in the process blanks of each dataset with a sample/blank ratio (mean peak area) equal to or less than 10, in addition to subtraction of the maximum blank peak area from all sample peak areas.

Chemicals commonly detected were plastic- and personal care product-related chemicals or surfactants (all grouped as industrial pollutants for subsequent analyses). Pharmaceuticals and agricultural pesticides were mostly found to be dataset specific, and several of these were shared between the two datasets containing the most xenobiotic features namely 'South Africa 2018/19' and 'Southern San Diego Coast 2019'. The insect repellent DEET[19] and the fungicide metalaxyl[20] were detected across various datasets, whereas the most commonly detected pharmaceutical-related feature across all datasets was atenolol acid (metabolite of the beta-blocker atenolol)[21].

Industrial pollutants dominate marine contamination

The most widespread annotated xenobiotic features (Level 1 and 2[18]) per class (industrial pollutants, pharmaceuticals and pesticides) are shown

https://doi.org/10.1038/s41561-026-01928-z

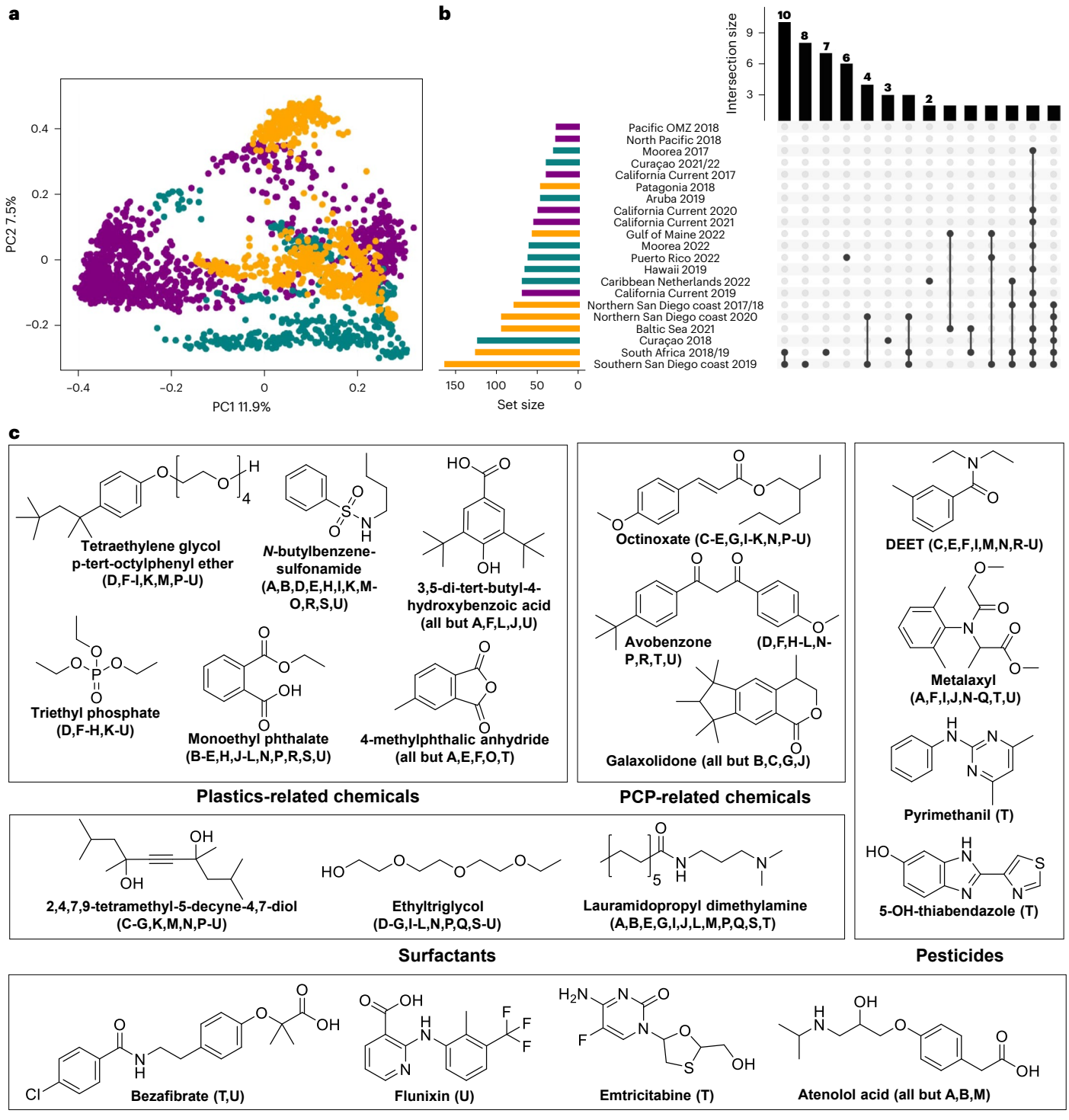


Fig. 2 | Xenobiotics compositional profiles differ between marine ecosystems. **a**, Principal coordinate analysis (PCoA) results (sample-centric relative abundance feature normalization, Bray–Curtis distance metric, PERMANOVA for ecosystem types $p < 0.001$ $R^2 = 0.1072$). The plot was generated in Python using scikit-bio (for distance matrices, PERMDISP and PERMANOVA), SciPy (for distance calculations). **b**, UpSet plot, generated in Intervene (https://asntech.shinyapps.io/intervene/)[48], showing the overlap of features annotated as xenobiotics between datasets. Intersection size describes the number of features that share the same presence/absence profile across datasets, defined as peak areas exceeding a minimum threshold. Individual datasets are assigned capital letters as identifiers. Colour denotes marine ecosystem type (orange: coastal temperate, turquoise: coral reef, purple: open ocean). **c**, Molecular structures of annotated xenobiotics detected throughout the sampled ecosystems. Capital letters in brackets indicate in which datasets the various molecular structures were detected.

https://doi.org/10.1038/s41561-026-01928-z

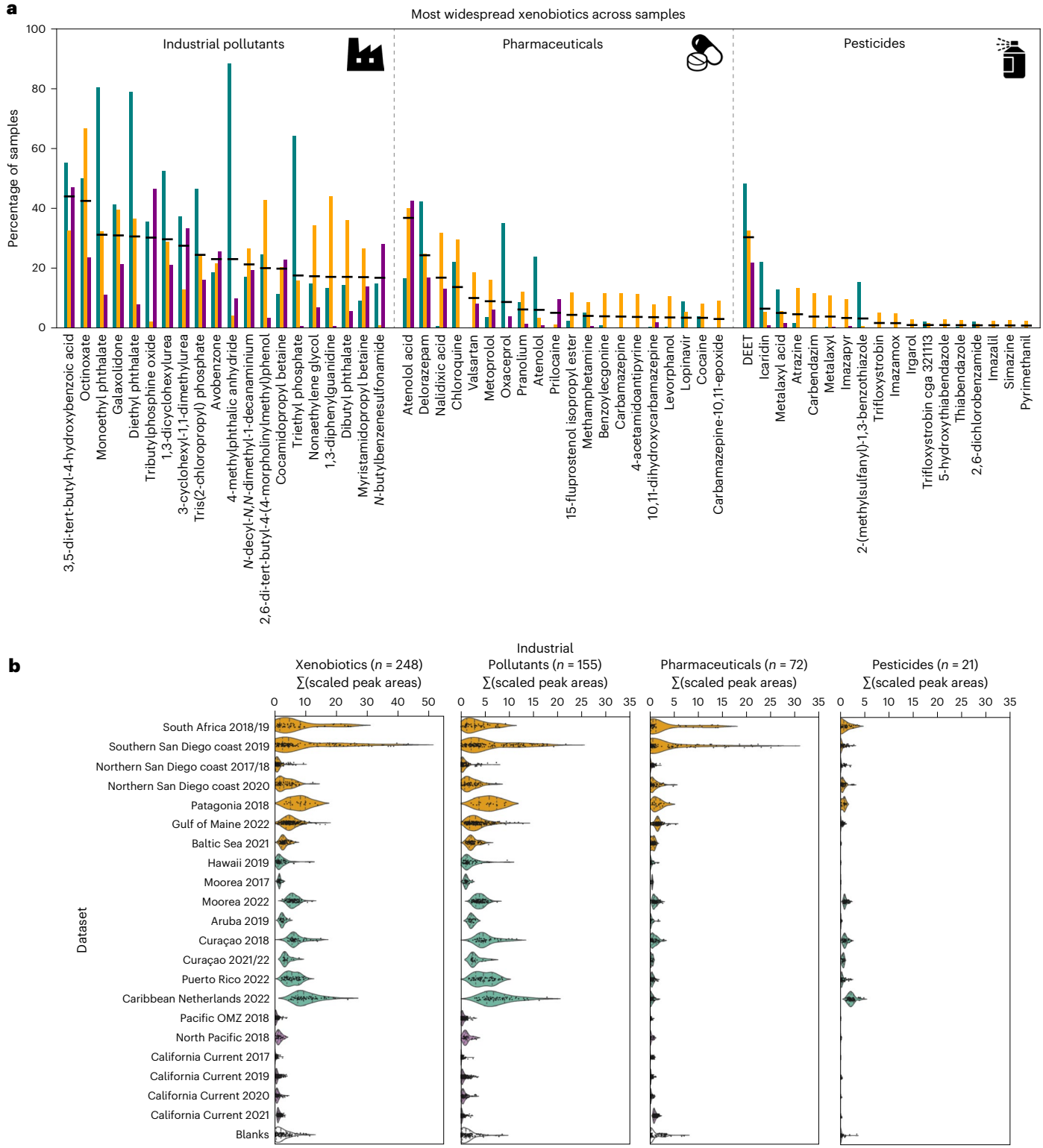


Fig. 3 | Industrial pollutants dominate xenobiotic profiles. a, Bar graphs showing the percentage of samples in which the most widespread annotated xenobiotics exceeded our minimum peak area threshold. Horizontal bars indicate the average frequency across all ecosystem types. **b**, Violin plots of the summed scaled peak areas (feature-centric scaled, arbitrary units) for annotated xenobiotics in each sample (points). Summed scaled peak areas are an estimate of the total amount of each xenobiotic type in each sample accounting for variation in feature intensities of different compounds. Raw peak areas from each ion feature are min-max scaled between 0 and 1 across all samples to account for ionization differences among features, then these scaled peak areas are summed by xenobiotic category (industrial pollutants, pharmaceuticals, pesticides). '*n*' represents the number of features contributing to each panel. Colours denote marine ecosystem type (orange: coastal temperate, turquoise: coral reef, purple: open ocean).

in Fig. 3a. Features annotated as industrial pollutants were found to be much more widely spread throughout the samples than pharmaceuticals and pesticides. Among industrial pollutants, the top-five features were detected in more than 30% of all samples. These features exemplify compound classes often associated with water pollution, including 3,5-di-*tert*-butyl-4-hydroxybenzoic acid[22], octinoxate[23], mono- and diethyl phthalate[24] and galaxolidone[25].

Most of these annotated compounds are derived from petrochemicals and associated with plastic materials, a known source of adverse health effects observed globally[26]. This finding provides further evidence that plastic-derived carbon, including micro- and nano-plastics, contributes a substantial portion to the marine carbon pool[6,27]. The five most widespread pharmaceutical features present in 10% to 37% of all samples were annotated as atenolol acid (metabolite of the beta-blocker atenolol)[21], delorazepam (antidepressant known to be persistent in the environment)[28] along with nalidixic acid[29] chloroquine[30] and valsartan[31]. For pesticides, the most widespread features were DEET and icaridin (both insect repellents used in, for example, mosquito spray)[32], that were detected in 30% and 6% of all samples, respectively.

Considering the sum of scaled peak areas of xenobiotics, 'coastal temperate' and 'coral reef' datasets showed more pronounced and variable pollution fingerprints than 'open ocean' datasets (Fig. 3b and Supplementary Fig. 7). We interpret the higher peak areas as evidence of increased relative concentrations and richness of individual xenobiotics and the imprint of pollution across datasets and ecosystems. The predominant annotated xenobiotics were industrial pollutants, which exhibited similar scaled peak area profiles to the grouped xenobiotics as a whole. Features annotated as pharmaceuticals were most pronounced in samples collected from or near estuarine locations (South Africa 2018/19 and Southern San Diego Coast 2019). Pesticides showed strongest imprints in the South Africa 2018/19[13] and Caribbean Netherlands 2022 dataset, while contributing almost nothing to 'open ocean' datasets. These results indicate that industrial pollutants are more widely distributed in Pacific 'open ocean' environments compared to pharmaceuticals and pesticides.

Molecular networking reveals xenobiotic diversity

Limited annotation of xenobiotic derivatives, degradation products, and different ion adducts that are not annotated through spectral matching is a central bottle-neck in the analysis of non-targeted LC-HR-MS/MS data. Molecular networking (MN) allows for structural annotation through propagation from spectrum-library matches through network connections[17]. This allows for further insight into the magnitude and structural complexity of xenobiotics by including unknown derivatives and degradation products.

Molecular networks of xenobiotics showed that the compound family of polyalkylene glycols (PAGs), which are widely used as lubricants, surfactants and plasticizers, were both highly abundant and structurally diverse (Fig. 4a). The peak areas of PAGs were highest in the Southern San Diego Coast 2019 dataset, followed by several 'coral reef' and other 'coastal temperate' datasets. In contrast, the 'open ocean' datasets exhibited notably lower scaled PAG peak area sums. While specific PAG features were matched in all ecosystem types, most were predominantly detected in 'coastal temperate' or 'coral reef' datasets and the polyethylene glycols (PEGs), polypropylene glycols (PPGs) and *tert*-octylphenyl ethers of PEGs (OPPEGs) clusters exhibited distinct shifts in relative contributions from these ecosystem types. Interestingly, many of these polymers fall in the lower operational range (5–10 nm) of nanoplastics[33,34], which have recently gained attention as major pollution sources in the ocean[27].

Scaled peak areas of organophosphate flame retardants showed strong differences between 'coastal temperate' and 'coral reef' datasets as compared to 'open ocean' datasets (Fig. 4b). For instance, tris(chloropropyl) phosphate and triethyl phosphate were observed as prominent features in several 'coral reef' datasets, whereas diphenylmethylphosphine oxide was only detected in 'coastal temperate' datasets and triphenyl phosphate was also detected in some 'open ocean' samples. Another diverse set of features was annotated as phthalates, comprising 22 features in total (Fig. 4c). The four features with the most intense signal intensity were detected in all marine ecosystem types: three of these features, annotated as diethyl phthalate, monoethyl phthalate and 4-methyl-phthalic anhydride, were detected with highest average abundance in 'coral reef' datasets, whereas dibutyl phthalate showed comparably higher contributions from 'coastal temperate' datasets. Overall, Curaçao 2018 and Caribbean Netherlands 2022 appear to exhibit elevated levels of phthalates, whereas 'open ocean' datasets generally showed lower levels than the other two types of ecosystem. Noteworthy, several smaller chain PAGs, organophosphate compounds and phthalates are particularly prominent in 'coral reef' datasets, and we suspect this could be a result of localized point sources and temperature and photoageing promoting increased leaching from immersed plastic debris[35–39].

Xenobiotics contribute substantially to the DOM pool

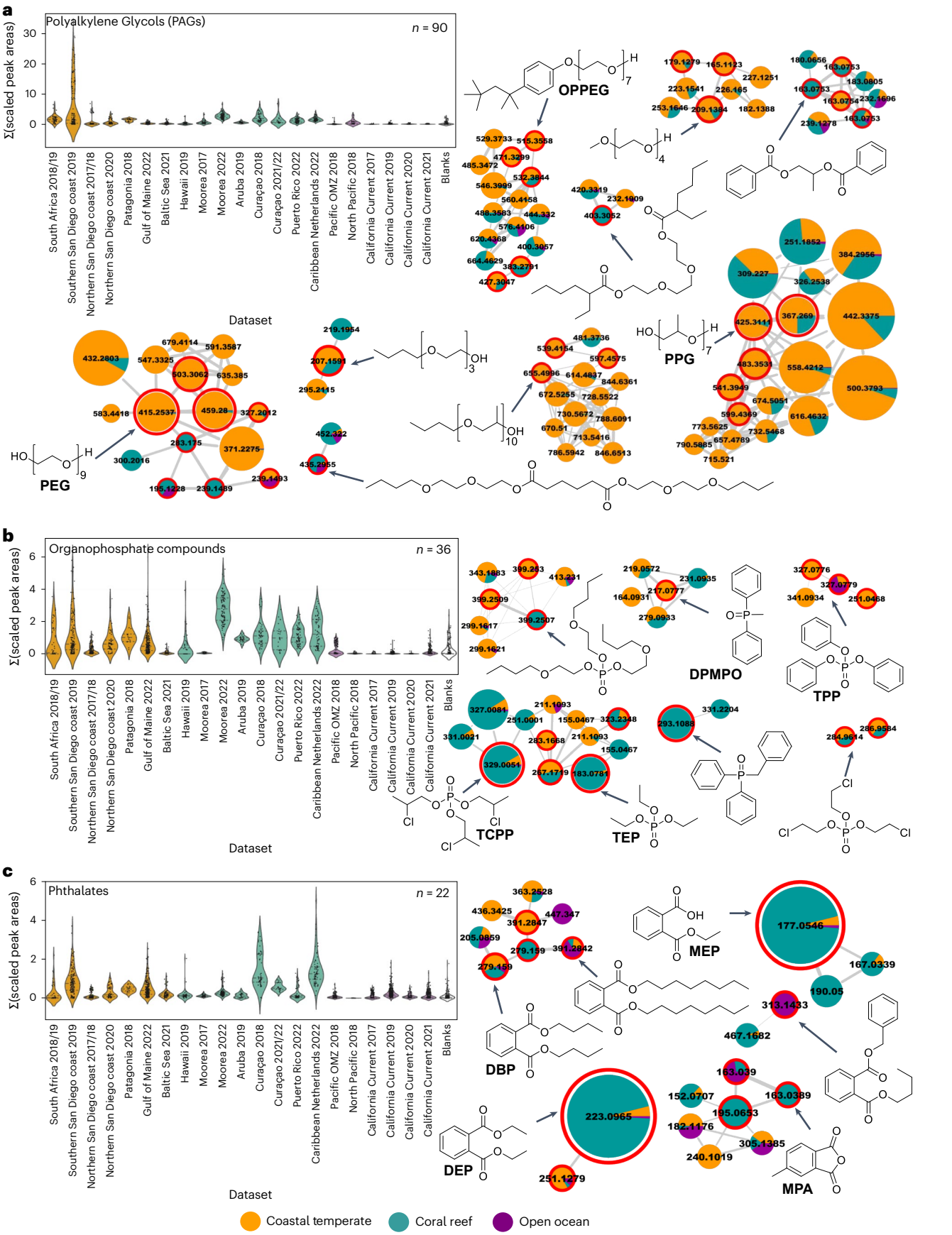
When considering all features that are direct (first) MN neighbours of xenobiotics, approximately 4% of the total number of detected ion features are linked to putative anthropogenic pollutants (1.6% have a direct spectral library match to a xenobiotic). Comparing the peak area from all annotated xenobiotics to the total peak area from all features per sample, showed that up to 63% of peak area could be attributable to xenobiotic features (up to 76% for estuarine samples; Fig. 5a).

For 'coastal temperate' datasets, the median peak area percentages of xenobiotics range between 0.5% and 9%, with the highest values for South Africa 2018/19 and Southern San Diego Coast 2019 (8%). Several 'coral reef' datasets exhibited lower median percentages of 2–8% (Moorea 2017 and 2022, Curaçao 2018 and 2021/22, Aruba 2019), whereas three datasets showed the highest median percentages, namely, Puerto Rico 2022, Caribbean Netherlands 2022 and Hawaii 2019 (20%, 11% and 10%, respectively). For 'open ocean' datasets, the median percentages ranged from 0.5 to 4%.

When including the first MN neighbours with the annotated xenobiotics, the contribution of xenobiotics to total signal naturally increases (median range 2–34%; Fig. 5b). The high proportion of signal intensity attributed to xenobiotic features, while representing only a small fraction of total number of features, suggests that xenobiotics are less diverse but observed at higher abundance compared to natural DOM components.

Fig. 4 | Molecular networking propagates xenobiotic molecular diversity beyond spectral library-based annotations. Examples of selected xenobiotic classes and clusters of individual xenobiotics with their first molecular network neighbours. **a**–**c**, Each panel comprises a violin plot of summed scaled peak areas (feature-centric, arbitrary units; Fig. 3b) and molecular network subclusters of first neighbours for PAGs (PPG=polypropylene glycol, PEG=polyethylene glycol, OPPEG=*p-tert*-octylphenol polyethylene glycols) (**a**), organophosphate compounds (TCPP=tris(2-chloropropyl) phosphate, DPMPO=diphenylmethylphosphine oxide, TEP=triethyl phosphate, TPP=triphenyl phosphate) (**b**) and phthalates (DEP=diethyl phthalate, MEP=monoethyl phthalate, DBP=dibutyl phthalate, MPA = 4-methylphthalic anhydride) (**c**). Nodes representing compounds that were annotated through spectral library matching are shown with red borders. Molecular network node pie chart slices correspond to average peak area across dataset types, for example, a node representing a feature found most in the 'coastal temperate' dataset will be mostly orange, as with PEGs in **a**. Node sizes increase with total average peak area. Colours denote marine ecosystem type (orange: coastal temperate, turquoise: coral reef, purple: open ocean).

https://doi.org/10.1038/s41561-026-01928-z



https://doi.org/10.1038/s41561-026-01928-z

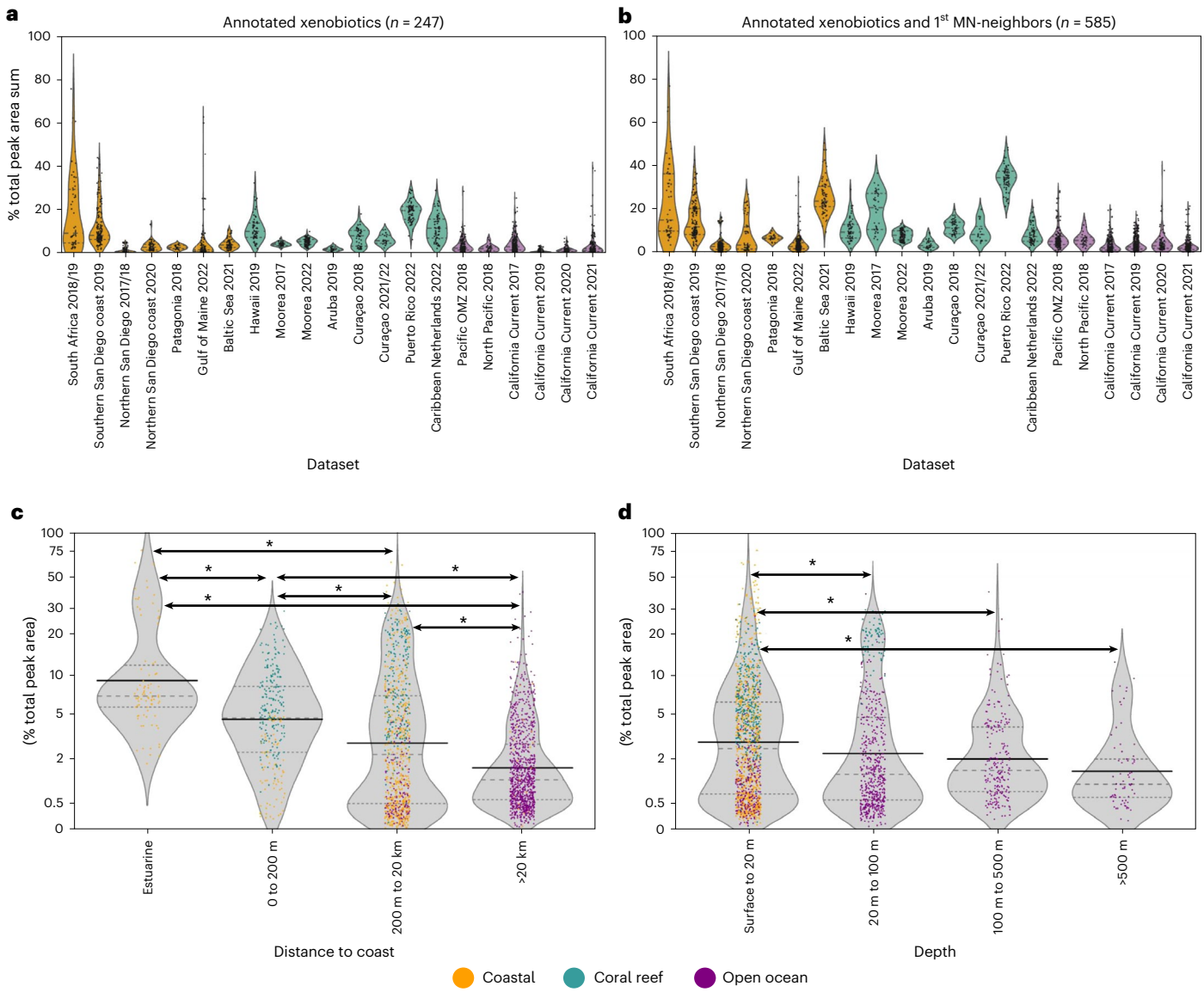


Fig. 5 | Xenobiotics comprise a significant proportion of the dissolved organic matter pool in coastal areas and decline with distance from land. **a**, Percentage of total peak area sum contributed by annotated xenobiotics ('*n*' indicates the number of contributing ion features). **b**, Percentage of total peak area sum contributed by annotated xenobiotics and their first molecular network neighbours. **c**, Percentage of total peak area sum contributed by annotated xenobiotics plotted according to binned geodesic distances to coast ($p < 0.01$; Tukey HSD, two-sided, adjusted for multiple comparisons). Linear regression of log transformed percentage values also showed a significant relationship ship to the geodesic distances to coast ($p < 0.01$; Pearson). **d**, Percentage of total peak area sum contributed by annotated xenobiotics plotted according to binned depths. Black horizontal bars indicate the mean, dashed lines indicate the median (50th percentile) and the 25th and 75th percentiles (quartiles), arrows with asterisks (*) indicate bin pairs with statistically significant differences ($p < 0.01$; Tukey HSD, two-sided, adjusted for multiple comparisons), with exact *p*-values reported in Supplementary Table 1. Corresponding plots resolved by xenobiotic class are shown in Supplementary Fig 8. The DEET feature was removed due to excessive peak area contribution in several 'coral reef' datasets (Supplementary Fig. 9). Colours denote marine ecosystem type (orange: coastal temperate, turquoise: coral reef, purple: open ocean).

A decrease in total xenobiotic signal intensity was observed from estuarine to near coastal (0 to 200 m) and intermediate (200 m to 20 km) to open ocean (>20 km) ecosystems (Fig. 5c, geodesic distance). These decreases were statistically significant ($p < 0.01$, Tukey's HSD) for all comparisons. This gradient suggests that the primary source of xenobiotics is terrestrial and may result from dilution and attenuation processes of river-transported anthropogenic contaminants[40]. Estimates of fossil contribution to riverine dissolved organic carbon (DOC) in 'completely human-disturbed' river catchments using radiocarbon ages have shown that fossil carbon can contribute between on average 22% (6–34%) of river DOC[41], which falls in similar range as some of our observations in the estuarine datasets.

When considering samples retrieved from different depths in the marine water column (Fig. 5d), a less pronounced but significant trend ($p < 0.01$, Tukey's HSD) was observed for xenobiotic levels elevated in samples collected at the surface compared to deeper water bodies.

Limitations of our analysis

It is important to note that this analysis bears several limitations, such as the extraction efficiency during solid phase extraction (SPE) procedures and ionization efficiency during electrospray ionization (ESI), which is structure dependent and can span wide ranges[42]. Predictions for logP and pK_a values for all annotated compounds and subsets of the NORMAN susDat[43] and COCONUT[44] libraries suggested no pronounced

differences of extraction and ionization bias of our methodology between xenobiotics and natural products (Supplementary Fig. 10). Due to limited publicly available datasets, we only considered positive mode ESI data, which predominantly capture metabolites amenable to cationic ionization. This reduces the scope of detectable compounds and common organic anthropogenic chemicals, such as per- and polyfluoroalkyl substances and polychlorinated biphenyls. In addition, our alignment approach is imperfect and suffers from inherent batch effects between datasets. This is a general drawback of current data-dependent acquisition-based LC-MS/MS data analysis strategies that could be addressed through DIA and future hybrid alignment workflows. Spectral annotations are mostly of confidence level 2[18], which results in a typically 1% false discovery rate at 0.7 cosine scores[45]. To further estimate potential false positive annotations, we compared the top-five spectral matches for each xenobiotic feature, which showed that 96% of features were unambiguously matched (when enantiomers and regional isomers were ignored). The high confidence of our spectral matching approach is also reflected in the MN, where first neighbours of xenobiotics features were also annotated as xenobiotics (90 to 96%), whereas xenobiotic to natural product edges resulted in significantly fewer network connections (4% to 10%; Supplementary Fig. 11).

Conversely, the incompleteness of current spectral libraries probably results in many xenobiotics remaining unannotated. To further expand the structural annotation, in silico structure predictions (SIRIUS6)[46] were compared to the NORMAN susDat and the COCONUT libraries. This indicated approximately 10% of the unannotated features may represent xenobiotics (Supplementary Fig. 12). Some xenobiotics labelled as industrial pollutants in this study, can also be observed as contaminants in lab plasticware. To exclude potential background contaminants from our analysis, we applied a strict blank removal approach for all features that were observed at ≥10% average abundance in the process blanks, in addition to subtracting the maximum blank peak area for all features. To verify this procedure, we manually inspected the sample and blank signals for all xenobiotic features, indicating the effectiveness of our strategy (Supplementary Fig. 13). The semi-quantitative comparison of our analysis relies on positive mode ESI peak area of annotated xenobiotic features to the total peak area across all deconvoluted ions. The reported percentages cannot be directly interpreted as concentration percentages without the use of analytical standards. Nevertheless, we assume that global differences in extraction and ionization efficiencies between detected chemical species will be largely uniform across datasets and compound origin types, as indicated by our comparison of logP, pK_a and pK_b values between entries of the NORMAN susDat (xenobiotics) and COCONUT (natural products) databases (Supplementary Fig. 10).

To estimate whether the reported total signal percentages are within realistic concentration percentages, we accumulated a list of reported xenobiotic concentrations from previous targeted studies (Supplementary table 2). By considering the maximum values of 232 reported substances, we calculated a scenario for contaminated seawater in which the cumulative xenobiotic DOC concentration would amount up to 42 μg l^{-1}, corresponding to around 3.5% of total DOC commonly found in seawater (1,200 μg l^{-1}) (ref. 47). Extrapolating these concentrations to the 585 putatively assigned xenobiotics in our study, a range of 10% total DOC thus appears plausible. Due to the lack of data, our analysis only covers a portion of global ocean basins. However, we believe that our multi-regional analysis provides a robust insight into the distribution of xenobiotics in marine systems. Our results highlight the substantial percentage of total marine DOM features attributed to anthropogenic input, which should be further expanded and validated in future studies. In addition to absolute quantification, future studies may also be guided by estimates based on radiocarbon age, which have been applied to riverine systems to estimate fossil carbon contributions[41].

Anthropogenic compounds in marine DOM need monitoring

Our analysis provides new insights into the distribution of anthropogenic organic compounds in marine ecosystems. By analysing over 2,300 samples from diverse marine environments, our results indicate that the investigated ecosystems ('coastal temperate', 'coral reef' and 'open ocean') have unique xenobiotic profiles, corresponding to the presence of industrial pollutants, pesticides and pharmaceuticals (Fig. 3). Within these data, we uncovered widespread xenobiotics found within large portions of samples and ecosystems, highlighting the diverse and pervasive nature of anthropogenic pollution in our oceans (Fig. 2). The detection of industrial pollutants traditionally associated with coastal environments (for example, PAGs and phthalates) within open ocean datasets suggests a high degree of environmental persistence.

Strikingly, in the most extreme cases, up to 63% of the total feature peak areas (76% when considering estuarine samples) could be attributed to xenobiotics, indicating that anthropogenic substances can substantially contribute to the ocean's organic matter pool (Figs. 3 and 5a,b). This raises further questions about the impact of anthropogenic micropollutants on the ocean DOM reservoir and its role in global carbon cycling, particularly considering that several of the detected compound classes are known to affect phytoplankton metabolism. To determine the extent of chemical pollution in global marine systems, large-scale studies with unified analytical approaches, including non-targeted analysis and targeted absolute quantification are necessary. We recommend the development of long-term sampling efforts with standardized protocols for sample collection and data acquisition within coordinated laboratories along with standardized data storage and analysis following FAIR (findable, accessible, interoperable and reusable) data principles.

In summary, our analysis offers a snapshot of the widespread distribution of xenobiotics in marine environments and their contribution to the marine dissolved organic matter pool. These results emphasize the need for continued, standardized monitoring efforts to track human chemical impacts on marine ecosystems.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41561-026-01928-z.

References

1. Halpern, B. S. et al. A global map of human impact on marine ecosystems. *Science* **319**, 948–952 (2008).
2. Zhou, W., Li, M. & Achal, V. A comprehensive review on environmental and human health impacts of chemical pesticide usage. *Emerg. Contam.* **11**, 100410 (2025).
3. Sigmund, G. et al. Addressing chemical pollution in biodiversity research. *Glob. Change Biol.* **29**, 3240–3255 (2023).
4. Nguyen, M.-K. et al. A comprehensive review on ecological effects of microplastic pollution: an interaction with pollutants in the ecosystems and future perspectives. *TrAC, Trends Anal. Chem.* **168**, 117294 (2023).
5. Yang, Y. et al. Which micropollutants in water environments deserve more attention globally?. *Environ. Sci. Technol.* **56**, 13–29 (2022).
6. Stubbins, A., Law, K. L., Munoz, S. E., Bianchi, T. S. & Zhu, L. Plastics in the Earth system. *Science* **373**, 51–55 (2021).
7. Tian, Z. et al. Suspect and nontarget screening for contaminants of emerging concern in an urban estuary. *Environ. Sci. Technol.* **54**, 889–901 (2020).

8. Carve, M., Singh, N., Askeland, M., Allinson, G. & Shimeta, J. Salting-out assisted liquid–liquid extraction combined with LC–MS/MS for the simultaneous determination of seven organic UV filters in environmental water samples: method development and application. *Environ. Sci. Pollut. Res.* **30**, 104870–104885 (2023).
9. Papadopoulos Lambidis, S. et al. Two-dimensional liquid chromatography tandem mass spectrometry untangles the deep metabolome of marine dissolved organic matter. *Environ. Sci. Technol.* **58**, 19289–19304 (2024).
10. Patrone, J. et al. Enhancing molecular characterization of dissolved organic matter by integrative direct infusion and liquid chromatography nontargeted workflows. *Environ. Sci. Technol.* **58**, 12454–12466 (2024).
11. Schramm, T., Kalinski, J.-C. J., Arini, G. S., Da Silva, R. R. & Petras, D. Uncovering the structural space of marine dissolved organic matter. *Annu. Rev. Mar. Sci.* **18**, 165–194 (2026).
12. Petras, D. et al. Non-targeted tandem mass spectrometry enables the visualization of organic matter chemotype shifts in coastal seawater. *Chemosphere* **271**, 129450 (2021).
13. Kalinski, J.-C. J. et al. Urban and agricultural influences on the coastal dissolved organic matter pool in the Algoa Bay estuaries. *Chemosphere* **355**, 141782 (2024).
14. Kalinski, J.-C. et al. Comparability of liquid chromatography tandem mass spectrometry analysis of dissolved organic matter across laboratories. *Environ. Sci. Technol.* **60**, 4814–4829 (2026).
15. Schmid, R. et al. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* **41**, 447–449 (2023).
16. Heuckeroth, S. et al. Reproducible mass spectrometry data processing and compound annotation in MZmine 3. *Nat. Protoc.* **19**, 2597–2641 (2024).
17. Wang, M. et al. Sharing and community curation of mass spectrometry data with global natural products social molecular Networking. *Nat. Biotechnol.* **34**, 828–837 (2016).
18. Sumner, L. W. et al. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **3**, 211–221 (2007).
19. Costanzo, S. D., Watkinson, A. J., Murby, E. J., Kolpin, D. W. & Sandstrom, M. W. Is there a risk associated with the insect repellent DEET (*N*,*N*-diethyl-m-toluamide) commonly found in aquatic environments?. *Sci. Total Environ.* **384**, 214–220 (2007).
20. Celis, R., Gámiz, B., Adelino, M. A., Hermosín, M. C. & Cornejo, J. Environmental behavior of the enantiomers of the chiral fungicide metalaxyl in Mediterranean agricultural soils. *Sci. Total Environ.* **444**, 288–297 (2013).
21. Küster, A. et al. Environmental risk assessment of human pharmaceuticals in the European Union: a case study with the β-blocker atenolol. *Integr. Environ. Assess. Manage.* **6**, 514–523 (2010).
22. Han, B. et al. Prevalence of synthetic phenolic antioxidants in food contact materials from China and their implications for human dietary exposure through take-away food. *J. Hazard. Mater.* **473**, 134599 (2024).
23. Kwon, B. & Choi, K. Occurrence of major organic UV filters in aquatic environments and their endocrine disruption potentials: a mini-review. *Integr. Environ. Assess. Manage.* **17**, 940–950 (2021).
24. Weaver, J. A. et al. Hazards of diethyl phthalate (DEP) exposure: a systematic review of animal toxicology studies. *Environ. Int.* **145**, 105848 (2020).
25. Su, Y. et al. Ecological risk of galaxolide and its transformation product galaxolidone: evidence from the literature and a case study in Guangzhou waterways. *Environ. Sci. Processes Impacts* **25**, 1337–1346 (2023).
26. Cropper, M., Dunlop, S., Hinshaw, H. & Symeonides, C. The benefits of removing toxic chemicals from plastics. *Proc. Natl Acad. Sci. USA* **121**, e2412714121 (2024).
27. ten Hietbrink, S., Materić, D., Holzinger, R., Groeskamp, S. & Niemann, H. Nanoplastic concentrations across the North Atlantic. *Nature* **643**, 412–416 (2025).
28. Fogliano, C., Motta, C. M., Acloque, H., Avallone, B. & Carotenuto, R. Water contamination by delorazepam induces epigenetic defects in the embryos of the clawed frog *Xenopus laevis*. *Sci. Total Environ.* **896**, 165300 (2023).
29. Hashemi, S. H., Kaykhaii, M., Jamali Keikha, A. & Naruie, N. Application of molecularly imprinted polymer pipette tip micro-solid phase extraction of nalidixic acid and acetaminophen from pills and seawater samples and their determination by spectrophotometry. *Chem. Pap.* **74**, 4009–4023 (2020).
30. Gupta, P. & Rani, V. A comprehensive assessment of environmental implications, ecotoxicity and bioaccumulation potential of repurposed drug hydroxychloroquine: from challenges to sustainability. *Rev. Environ. Contam. Toxicol.* **262**, 10 (2024).
31. Leite, M. et al. Pharmaceuticals in coastal sediments: validation of an ultra-high-performance liquid chromatography coupled with time-of-flight mass spectrometry (UHPLC-TOF-MS) methodology. *Environ. Pollut.* **380**, 126571 (2025).
32. Goodyer, L. & Schofield, S. Mosquito repellents for the traveller: does picaridin provide longer protection than DEET?. *J. Travel Med.* **25**, S10–S15 (2018).
33. Gigault, J. et al. Current opinion: what is a nanoplastic?. *Environ. Pollut.* **235**, 1030–1034 (2018).
34. Rani-Borges, B. & Ando, R. A. How small a nanoplastic can be? a discussion on the size of this ubiquitous pollutant. *Cambridge Prisms: Plast.* **2**, e23 (2024).
35. Li, C. & Tang, K. H. D. Effects of pH and Temperature on the Leaching of Di(2-Ethylhexyl) Phthalate and Di-n-butyl Phthalate from Microplastics in Simulated Marine Environment. *Biointerface Res. Appl. Chem.* **13**, 269 (2022).
36. Henkel, C. et al. Photoaging enhances the leaching of di(2-ethylhexyl) phthalate and transformation products from polyvinyl chloride microplastics into aquatic environments. *Commun. Chem.* **7**, 218 (2024).
37. Pinheiro, H. T. et al. Plastic pollution on the world's coral reefs. *Nature* **619**, 311–316 (2023).
38. Wang, W. et al. Bioaccumulation and potential risk of organophosphate flame retardants in coral reef fish from the Nansha Islands, South China Sea. *Chemosphere* **287**, 132125 (2022).
39. Schupp, T. et al. in *Reviews of Environmental Contamination and Toxicology* vol. 244 (ed. De Voogt, P.) 53–111 (Springer International Publishing, 2017).
40. Massicotte, P., Asmala, E., Stedmon, C. & Markager, S. Global distribution of dissolved organic matter along the aquatic continuum: across rivers, lakes and oceans. *Sci. Total Environ.* **609**, 180–191 (2017).
41. Butman, D. E., Wilson, H. F., Barnes, R. T., Xenopoulos, M. A. & Raymond, P. A. Increased mobilization of aged carbon to rivers by human disturbance. *Nat. Geosci.* **8**, 112–116 (2015).
42. Cancelada, L. et al. Assessment of styrene-divinylbenzene polymer (PPL) solid-phase extraction and non-targeted tandem mass spectrometry for the analysis of xenobiotics in seawater. *Limnol. Oceanogr. Methods* **20**, 89–101 (2022).
43. Mohammed Taha, H. et al. The NORMAN Suspect List Exchange (NORMAN-SLE): facilitating European and worldwide collaboration on suspect screening in high resolution mass spectrometry. *Environ. Sci. Eur.* **34**, 104 (2022).
44. Sorokina, M., Merseburger, P., Rajan, K., Yirik, M. A. & Steinbeck, C. COCONUT online: collection of open natural products database. *J. Cheminformatics* **13**, 2 (2021).

45. Scheubert, K. et al. Significance estimation for large scale metabolomics annotations by spectral matching. *Nat. Commun.* **8**, 1494 (2017).
46. Dührkop, K. et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* **16**, 299–302 (2019).
47. Lønborg, C. et al. A global database of dissolved organic matter (DOM) concentration measurements in coastal waters (CoastDOM v1). *Earth Syst. Sci. Data* **16**, 1107–1119 (2024).
48. Khan, A. & Mathelier, A. Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. *BMC Bioinf.* **18**, 287 (2017).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2026

Jarmo-Charles J. Kalinski[1,2], Abzer K. Pakkir Mohamed Shah[1,3], Bruno Ruiz Brandão da Costa[1,4], Shane P. Farrell[5], Lisa Schellenberg[6,7], Lana G. Graves[1,8], Tilman Schramm[1], Paolo Stincone[1,3], Irina Koester[9], Brandon M. Stephens[10], Ralph R. Torres[9], Lucia Cancelada[9,11], Caroline Utermann-Thüsing[12], Zachary A. Quinlan[9,13], Linda Wegley Kelly[9], Craig A. Carlson[14], Cristóbal Castillo-Ilabaca[15], Silvio Pantoja-Gutiérrez[16], J. Michael Beman[17], Aaron Hartmann[18,19], Allegra Aron[20], Xavier Siwe Noundou[21], Rosemary A. Dorrington[2], Deniz Tasdemir[12,22], Andreas F. Haas[6], Pieter C. Dorrestein[23], Craig E. Nelson[24], Lihini I. Aluwihare[9], Mingxun Wang[25] & Daniel Petras[1,3] ✉

[1]Department of Biochemistry, University of California Riverside, Riverside, CA, USA. [2]Department of Biochemistry, Microbiology and Bioinformatics, Rhodes University, Makhanda, South Africa. [3]Interfaculty Institute of Microbiology and Infection Medicine, University of Tuebingen, Tuebingen, Germany. [4]Department of Botany, Institute of Biosciences, University of São Paulo, São Paulo, Brazil. [5]Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, USA. [6]Department of Marine Microbiology and Biogeochemistry, Royal Netherlands Institute for Sea Research (NIOZ), Texel, Netherlands. [7]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, the Netherlands. [8]Department of Geosciences, Eberhard Karls University of Tuebingen, Tuebingen, Germany. [9]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. [10]Institute of Oceanography, National Taiwan University, Taipei, Taiwan. [11]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA, USA. [12]GEOMAR Centre for Marine Biotechnology, Research Unit Marine Natural Product Chemistry, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany. [13]Hawai'i Institute of Marine Biology, University of Hawai'i at Mānoa, Honolulu, HI, USA. [14]The Marine Science Institute, Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, Santa Barbara, CA, USA. [15]Graduate Program in Oceanography, Department of Oceanography, Faculty of Natural Sciences and Oceanography, University of Concepción, Concepción, Chile. [16]Department of Oceanography and Center for Oceanographic Research COPAS Coastal, University of Concepción, Concepción, Chile. [17]Department of Life and Environmental Sciences and Sierra Nevada Research Institute, University of California, Merced, Merced, CA, USA. [18]Perry Institute for Marine Science, Waitsfield, VT, USA. [19]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. [20]Department of Chemistry and Biochemistry, University of Denver, Denver, CO, USA. [21]Department of Pharmaceutical Sciences, School of Pharmacy, Sefako Makgatho Health Sciences University, Pretoria, South Africa. [22]Faculty of Mathematics and Natural Sciences, Kiel University, Kiel, Germany. [23]Skaggs School of Pharmacy, University of California San Diego, La Jolla, CA, USA. [24]Daniel K. Inouye Center for Microbial Oceanography: Research and Education, Department of Oceanography and Sea Grant College Program, University of Hawai'i at Mānoa, Honolulu, HI, USA. [25]Department of Computer Science, University of California Riverside, Riverside, CA, USA. ✉e-mail: dpetras@ucr.edu

Methods

Dataset selection

For our global analysis, we selected publicly available non-targeted LC-HR-MS/MS data from styrene-divinylbenzene-based solid phase extraction (Priority PolLutant (PPL-SPE)) extracts from seawater samples. An overview of all datasets is provided as Supplementary Table 3 in the Supplementary Information. All datasets were acquired on quadrupole-orbitrap platforms in our laboratories with identical LC methods and harmonized MS/MS settings based on our initial method development[12].

Extraction and LC-HR-MS/MS instrument methods

All samples (100–1,000 ml seawater) were acidified with HCl to pH 2 and extracted with PPL-SPE cartridges (200 or 1,000 mg) after optional prefiltration with sterivex (0.2 µm), supore (0.2 µm) or GF/F (0.7 µm) filters. Extracts were evaporated using vacuum centrifuges and resuspended in 100–200 µl of 80–100% MeOH. Injection volumes for LC-MS/MS analysis were either 5 or 10 µl. Ultra-high-performance liquid chromatography (UHPLC) separation was performed with Phenomenex Kinetex C18 or Evo C18 columns (2.1 ×150 mm, 1.7 µm particle size) using a standardized gradient elution with H20 + 0.1% formic acid (Solvent A) and Acetonitrile + 0.1% formic acid (Solvent B). Gradients were from 5%–50% B from 0–8 min and then increased from 50%–99% B from 8–10 min. After a 3 min washout phase at 99% B, the column was equilibrated at 5% for either 2 or 3 min.

MS and MS/MS spectra were acquired in heated electrospray ionization (HESI) positive mode, on Q Exactive and Q Exactive HF platforms. Unified ESI settings were used among all datasets with 52 AU sheath gas flow, 14 AU auxiliary gas flow, 0 AU sweep gas flow and 400 °C auxiliary gas temperature. The spray voltage was set to 3.5 kV and the inlet capillary to 320 °C. Fifty V S-lens voltage (Q Exactive) was applied. MS scan range was set to 150–1,500 m/z with a resolution at m/z 200 ($R_{m/z\,200}$) of 140,000 with one micro-scan in positive mode. The maximum ion injection time was set to 100 ms with automated gain control of 1.0E6. MS/MS spectra were recorded in data dependent acquisition (DDA) mode. Both MS1 survey scans (150–1,500 m/z) and up to 5 DDA MS/MS scans of the most abundant ions per duty cycle were measured with $R_{m/z\,200}$ of 17,500 with one micro-scan. The maximum ion injection time was set to 100 ms with automated gain control targets set to 1.0E6 for survey scans and 3.0E5 for MS/MS with minimum 10% C-trap filling. MS/MS precursor selection windows were set to m/z 1. Normalized collision energy was set to a stepwise increase from 20 to 30 to 40% with $z = 1$ as default charge state. MS/MS acquisition was triggered at the apex of peaks within 2–15 s from their first occurrence. Dynamic exclusion was set to 5 s. Unassigned charge states as well as isotope peaks were excluded from MS/MS acquisition.

Feature-based molecular networking

Centroided .mzML or .mzXML files were downloaded from massIVE and processed in mzmine (ver.4.2.0)[15]. Mass detection was carried out with a noise level of 1E5 at MS1 level and 1E3 at MS/MS level. Chromatograms were assembled with the Chromatogram builder requiring at least five consecutive scans, a minimum intensity for consecutive scans of 3E5 and a minimum absolute height of 2E6 with a m/z tolerance of 5 ppm. Deconvolution was accomplished with the Local minimum search module using a chromatographic threshold of 50%, a minimum search range of 0.1 min, a minimum absolute height of 2E6, a minimum ratio of peak top to edge of 1.4, a peak duration range of 0.05 to 3 minutes and a minimum of five consecutive scans. MS/MS scans were paired with a m/z tolerance of 0.002 or 10 ppm using the feature edges and a minimum relative height of 25% and at least one signal. Next, the feature lists were filtered to retain only features with an MS/MS spectrum (to reduce computational load). ^{13}C isotope signals were removed with the ^{13}C isotope filter with an m/z tolerance of 5 ppm, a retention time tolerance of 0.1 min and a maximum charge of 2 with the most intense isotope selected as representative. Next isotope patterns were allocated with the isotopic peaks finder considering the elements C, N, Cl, H, O, S and Br, an m/z tolerance of 5 ppm and a maximum charge of 2 with searching in the single most intense scan. Feature lists were merged with the Join Aligner with an m/z tolerance of 5 ppm and a retention time tolerance of 1.7 min and weights of m/z to retention time of 3 to 1. The feature list was then filtered to retain only features that occurred in at least ten samples and had an assigned MS/MS spectrum, minimum features in an isotope pattern were set to 2 and validation of ^{13}C isotope patterns was activated. The duplicate peak filter was employed with the mode 'NEW AVERAGE', an m/z tolerance of 5 ppm and a retention time tolerance of 1.7 min. The final aligned feature list was exported for GNPS and SIRIUS analysis with merging of MS/MS spectra using the 'weighted average' mode and summing of intensities, expected mass deviations of 5 ppm, cosine thresholds of 70%, signal count thresholds of 20%, isolation window offsets of 0 m/z and isolation window widths of 3 m/z. The files were uploaded to GNPS2 (gnps2.org) and processed in the feature-based molecular networking workflow[49], with the default library (https://gnps2.org/status?task=454547be77d94ae1ad00d798046ee7a0) and with added libraries, respectively (https://gnps2.org/status?task=5e1d73844a1e48068643bf57246a4149). For both jobs precursor and fragment ion tolerances were set to 0.01 m/z, window and precursor window filters were deactivated, and only library matches with a mass error <10 ppm were considered. For networking a minimum cosine of 0.7, four matched peaks and maximum shift of 200 m/z were set with the topology parameters of 10 for Top K and 100 for maximum component size. Library search parameters were a minimum cosine of 0.7 and 4 matched peaks.

Data clean-up, normalization, statistical analysis and visualizations

UpSet plots were created using Intervene (https://asntech.shinyapps.io/intervene/). Molecular networks were downloaded from GNPS2 and visualized in Cytoscape (ver.3.10.1). The geographic maps and geodesic distance to coast measurements were created using QGIS (https://www.qgis.org/) and the PyQGIS API (EPSG:4326, WGS84) with the Natural Earth shapefile (1:10 m; https://www.naturalearthdata.com/). All other analyses and visualization downstream of molecular networking were created using customized python scripts including elements of our standardized FBMN statistical pipeline[50]. Blank subtraction was carried out on the processed feature table from mzmine, where features in a dataset were set to zero if the mean peak area within the blanks of the dataset was greater or equal to 0.1 times the mean peak area within the samples of the dataset. Furthermore, the maximum blank peak area was subtracted from all peak areas per feature. After this, all blank samples were removed and all features that contained only zero entries were removed. We then aligned this table with the GNPS library hits output which we prefiltered to only contain hits of interest using a 'taglist' (Supplementary Table 2). This taglist also included information on which annotated compounds were considered xenobiotics and whether they were classified as 'industrial pollutants', 'pharmaceuticals' or 'pesticides'. The python packages utilized for building our jupyter notebook included pandas, numpy, plotly, scikit-learn, matplotlib and seaborn. To explore compositional shifts among samples, we normalized peak areas of each feature to the total peak area within each sample, converting peak areas to relative abundances of features. This approach allows us to visualize compositional shifts after accounting for variation in the total amount of DOM among samples. To explore variation in the number of xenobiotics across all samples and datasets from various ecosystems, we scaled raw peak areas of each feature using minimum–maximum scaling (function MinMaxScaler from scikit-learn), which converted absolute raw peak areas for each feature individually to a value within the range of 0 and 1 across all samples. We did this for xenobiotic features to achieve a feature-specific measure of xenobiotic abundance (Fig. 3b and Fig. 4)

that is independent from well-known ionization efficiency differences among ion features. For the violin plots with percentage *y* axes in Fig. 5, the raw peak areas for annotated xenobiotic features were summed per sample and then divided by the total peak area sum for all features of the respective sample, followed by multiplication with 100 to calculate the percentage of total peak area attributable to annotated xenobiotic features. All code is available as jupyter notebooks via Github (see the code availability statement).

Data availability

All datasets are publicly available through the MassIVE repository (massive.ucsd.edu) under the following identifiers: Aruba 2019 (UCSD, Q Exactive): MSV000085891; Baltic Sea 2021 (Tuebingen, Q Exactive HF): MSV000089481; California Current 2017 (UCSD, Q Exactive): MSV000082082; California Current 2019 (UCSD, Q Exactive): MSV000085852; California Current 2020 (UCSD, Q Exactive): MSV000086236; California Current 2021 (UCSD, Q Exactive): MSV000090123; Caribbean Netherlands 2022 (NIOZ, Q Exactive): MSV000094015; Curaçao 2018 (UCSD, Q Exactive): MSV000083522; Curaçao 2021/22 (NIOZ, Q Exactive): MSV000094018; Gulf of Maine 2022 (Tuebingen, Q Exactive HF): MSV000090696; Hawaii 2019 (UCSD, Q Exactive): MSV000082085; Moorea 2017 (UCSD, Q Exactive): MSV000082083; Moorea 2022 (UCSD, Q Exactive): MSV000090798; Northern Pacific 2018 (UCSD, Q Exactive): MSV000083365; Northern San Diego Coast 2017/18 (UCSD, Q Exactive): MSV000082312; Northern San Diego Coast 2020 (UCSD, Q Exactive): MSV000085555 and MSV000085779; Pacific OMZ 2018 (UCSD, Q Exactive): MSV000085480; Puerto Rico 2022 (Tuebingen, Q Exactive HF): MSV000092677; South Africa 2018/19 (UCSD, Q Exactive): MSV000085843; Southern San Diego Coast 2019 (UCSD, Q Exactive): MSV000083889; Patagonia 2018 (UCSD, Q Exactive): MSV000083521. All processed data and mzmine batch files used in this study are available via Zenodo at https://doi.org/10.5281/zenodo.16687158 (ref. 51). Source data are provided with this paper.

Code availability

All code, including python scripts, is available via Github at https://github.com/Functional-Metabolomics-Lab/Marine-Xenobiotics and via Zenodo at https://doi.org/10.5281/zenodo.16687158 (ref. 51).

References

49. Nothias, L.-F. et al. Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **17**, 905–908 (2020).
50. Pakkir Shah, A. K. et al. Statistical analysis of feature-based molecular networking results from non-targeted metabolomics data. *Nat. Protoc.* **20**, 92–162 (2025).
51. Kalinski, J.-C. Widespread detection of xenobiotics in marine dissolved organic matter (manuscript data upload). *Zenodo* https://doi.org/10.5281/zenodo.16687158 (2025).

Acknowledgements

We thank D. Rasher and D. Yiu for facilitating sample collection in the Gulf of Maine. D.P. was supported by the Simons Foundation International through a Simons Early Career Investigator in Aquatic Microbial Ecology and Evolution Award (SFI-LS-ECIAMEE-00013858) and the German Research Foundation through the Cluster of Excellence Controlling Microbes to Fight Infection (EXC 2124). S.P.-G. and C.C.-I. were supported by the Center for Oceanographic Research (COPAS) Coastal ANID FB210021 (Chile). R.R.T. and Z.A.Q. were supported by the National Science Foundation (NSF) through Graduate Research Fellowship and a Postdoctoral Fellowship (DGE-2038238, OCE-2308400 and award #2019252845). J.M.B. was supported through an NSF CAREER grant (OCE-1555375). We would like to thank the California Current Ecosystem Long Term Ecological Research (CCE-LTER) Program and the Captains and Crew of the RV *Atlantis*, RV *Oceanus*, RV *Revelle* and RV *Sally Ride*. We would also like to thank the team at University of California Gump Marine Station on Moorea for logistical support. Datasets presented here were supported in part by CCE-LTER funding (NSF OCE-1637632) and Moorea funding (OCE-2023509 to L.I.A. and P.C.D., OCE-2118618 to L.W.K., OCE-2023298 to C.E.N.), and NASA award 80NSSC18K0437 to C.A.C. A.H. was supported by the United States Department of Defense Environmental Security Technology Certification Program (no. CR20-5175). B.R.B.d.C. was supported by São Paulo Research Foundation (FAPESP, grant no. 2023/15215-4) through a Research Internship Abroad Fellowship. J.-C.J.K. was supported by a Rhodes University postdoctoral fellowship and by the South African Medical Research Council as well as the UK Medical Research Council, through the UK Government's Newton Fund (grant no. 96185). We thank the Shallow Marine and Coastal Research Infrastructure platform for research infrastructure support as well as the crew of the RV *Ukwabelana* and *Koos Smith* of the African Coelacanth Ecosystem Programme.

Author contributions

J.-C.J.K., A.K.P.M.S. and D.P. conceptualized the study. L.W.K., C.E.N., C.A.C., J.M.B., R.A.D., D.T., P.C.D., L.I.A. and D.P. supervised the study. S.P.F., R.R.T., L.C., B.M.S., C.U.-T., I.K., Z.A.Q., L.W.K., C.E.N., C.A.C., C.C.-I., S.P.-G., J.M.B., A.A., A.H., X.S.N., R.A.D., D.T., A.F.H., P.C.D., L.I.A. and D.P. performed sampling. S.P.F., P.S., R.R.T., L.C., B.M.S., C.U.-T., I.K., Z.A.Q., L.W.K., C.E.N., S.P.-G., J.M.B., A.H., X.S.N., D.T., A.F.H. and D.P. acquired LC-MS/MS data. J.-C.J.K., B.R.B.d.C., L.G.G., C.E.N., M.W. and D.P. analysed the data. J.-C.J.K., A.K.P.M.S., B.R.B.d.C., S.P.F., L.G.G., T.S., Z.A.Q., C.E.N., M.W. and D.P. wrote the manuscript. All authors edited and approved the final manuscript.

Competing interests

P.C.D. is an advisor and holds equity in Cybele, BileOmix and Sirenas and is a scientific co-founder, advisor, holds equity in and/or received income from Ometa, Enveda and Arome with prior approval by UC-San Diego. P.C.D. also consulted for DSM animal health in 2023. M.W. is a co-founder of Ometa Labs LLC. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41561-026-01928-z.

Correspondence and requests for materials should be addressed to Daniel Petras.

Peer review information *Nature Geoscience* thanks Eric Carmona, Mengyang Liu and Aron Stubbins for their contribution to the peer review of this work. Primary Handling Editor: James Super, in collaboration with the *Nature Geoscience* team.

Reprints and permissions information is available at www.nature.com/reprints.