Development of suspect and non-target screening methods for detection of organic contaminants in highway runoff and fish tissue with high-resolution time-of-flight mass spectrometry†‡

Bowen Du, Jonathan M. Lofton, Katherine T. Peter, Alexander D. Gipe, C. Andrew James, Jenifer K. McIntyre, Nathaniel L. Scholz, Joel E. Baker and Edward P. Kolodziej

Untreated urban stormwater runoff contributes to poor water quality in receiving waters. The ability to identify toxicants and other bioactive molecules responsible for observed adverse effects in a complex mixture of contaminants is critical to effective protection of ecosystem and human health, yet this is a challenging analytical task. The objective of this study was to develop analytical methods using liquid chromatography coupled to high-resolution quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) to detect organic contaminants in highway runoff and in runoff-exposed fish (adult coho salmon, Oncorhynchus kisutch). Processing of paired water and tissue samples facilitated contaminant prioritization and aided investigation of chemical bioavailability and uptake processes. Simple, minimal processing effort solid phase extraction (SPE) and elution procedures were optimized for water samples, and selective pressurized liquid extraction (SPE) procedures were optimized for fish tissues. Extraction methods were compared by detection of non-target features and target compounds (e.g., quantity and peak area), while minimizing matrix interferences. Suspect screening techniques utilized in-house and commercial databases to prioritize high-risk detections for subsequent MS/MS characterization and identification efforts. Presumptive annotations were also screened with an in-house linear regression (log K_{ow} vs. retention time) to exclude isobaric compounds. Examples of confirmed identifications (via reference standard comparison) in highway runoff include ethoprophos, prometon, DEET, caffeine, cotinine, 4(5)-methyl-1H-methylbenzotriazole, and acetanilide. Acetanilide was also detected in runoff-exposed fish gill and liver samples. Further characterization of highway runoff and fish tissues (14 and 19 compounds, respectively with tentative identification by MS/MS data) suggests that many novel or poorly characterized organic contaminants exist in urban stormwater runoff and exposed biota.

Environmental significance

This work specifically focuses on developing methods for the chemical characterization of environmental systems with as yet unexplained acute mortality events linked to water quality. Simple procedures requiring minimal sample processing were developed for extraction of water and tissue samples. In combination with the use of suspect screening databases, processing of paired water and tissue samples facilitated contaminant prioritization and aided investigation of chemical bioavailability and uptake processes. These efforts help guide analytical methodology and workflow development, and provide biological relevance to identification of novel contaminants in highway runoff and runoff-exposed fish tissues. Novel detection of acetanilide in stormwater runoff presents an example where a focus on a specific mode of action was successfully used to prioritize our HRMS detection efforts, and can now guide development of toxicity-related hypotheses that are testable with bioassays. The combination of HRMS analytical chemistry and aquatic toxicology is a promising tool in identifying potential pollutants in complex environmental mixtures and guiding future source control efforts.
Introduction

Stormwater runoff is one of the most important forms of nonpoint source pollution that impairs aquatic ecosystems globally.\textsuperscript{1} Urban runoff is typically contaminated with synthetic organic compounds, metals, and inorganics arising from a wide variety of sources.\textsuperscript{2} Without appropriate treatment, these chemicals and related transformation products are transported to surface water or groundwater by intentional discharge, infiltration, or overland flow where they degrade receiving water quality.\textsuperscript{3} Many stormwater contaminants are potentially toxic at elevated concentrations,\textsuperscript{4} including various pesticides,\textsuperscript{5-7} petroleum hydrocarbons,\textsuperscript{8} heavy metals,\textsuperscript{9} and other chemicals originating from urban sources,\textsuperscript{10,11} including uncharacterized compounds with potential ecological hazards.\textsuperscript{12,13} Understanding the occurrence, sources, concentrations, and transport pathways of stormwater-derived chemical contaminants, including identification of compounds most hazardous to aquatic life, is critical for effective water quality protection in urbanized watersheds.\textsuperscript{14}

As one motivating example of stormwater impacts on a commercially, recreationally, and ecologically important species, urban runoff is acutely lethal to adult coho salmon (Oncorhynchus kisutch) in the Pacific Northwest. Notably, this syndrome is not correlated with conventional water chemistry parameters (e.g., temperature, dissolved oxygen, and suspended solids), disease, spawner conditions, or exposure to pesticides, metals or polycyclic aromatic hydrocarbons (PAHs).\textsuperscript{1,15} As an example of why additional characterization of contaminant flows in urban stormwaters is needed, the available evidence suggests that as-yet unknown toxicant(s) are causing this mortality event. Typically, chemical profiling relies upon use of targeted analyses built from reference standards (e.g., PAHs, metals, and pesticides), often in combination with controlled exposures and insights into mechanisms of action to correlate detections with biological relevance.\textsuperscript{7,11,16} While accurate and sensitive, these approaches remain narrowly defined options for water quality assessment because such targeted analyses only detect a small subset of contaminants present and can be constrained by the lack of reference standards for emerging contaminants and bioactive transformation products.

To broaden detection capabilities, non-target and suspect screening approaches typically detect higher numbers of distinct contaminants relative to targeted approaches,\textsuperscript{17} although often at a cost to method sensitivity and selectivity.\textsuperscript{18} High-resolution mass spectrometry (HRMS) is a broad spectrum screening method based upon accurate-mass detection to aid chemical characterization of complex mixtures.\textsuperscript{19} Extraction of chromatographic features can demonstrate thousands of contaminant detections per sample. While non-target screening (i.e., HRMS data acquisition and assignment of tentative identities using accurate mass and isotopic information) does not start with reference standards, MS/MS fragmentation patterns are typically matched to authentic standards to enable conclusive confirmation of chemical identity.\textsuperscript{20-22} HRMS analyses often employ orbitrap or time-of-flight detectors, typically coupled with gas or liquid chromatography for separation.\textsuperscript{9,23} Screening with some prior information (i.e., a given structure; suspect screening) and identification starting from exact mass, isotope, adduct, and fragmentation information (non-target screening) are approaches for identifying contaminants in environmental samples.\textsuperscript{8} However, HRMS characterization of water, tissues, and sediments from aquatic environments is time-consuming, and still needs significant method development and optimization for more efficient performance.\textsuperscript{24} Confident identification of small molecules using HRMS-based suspect and non-target analysis depends upon consistent data acquisition, effective data reduction, and the ability to prioritize identification efforts within large datasets, as well as the availability of screening databases that include environmental contaminants and reference standards to confirm detections.\textsuperscript{25}

Notably, urban receiving waters subject to non-point source pollution remain poorly characterized relative to point sources like municipal wastewater (Web of Science search for “stormwater contaminant” yields 470 references vs. 5442 references for “wastewater contaminant”). Contaminant flows in urban systems are still poorly understood, and biological responses (i.e., bioaccumulation, ecotoxicology) are often not integrated with chemical characterization efforts. Previous studies have optimized extraction methods for non-target analysis of contaminants in biological matrices\textsuperscript{26} and examined contaminant co-occurrence in surface water and exposed fish.\textsuperscript{27,28} To develop methods for characterization of organic contaminants in urban stormwater runoff, especially biologically relevant compounds with potential adverse effects on aquatic species, we analyzed paired highway runoff and runoff-exposed coho salmon tissues from controlled stormwater exposure experiments. We focused on optimizing sample extraction and data reduction methods to enhance detection capabilities using HRMS techniques, particularly to identify contaminants that might be implicated in pre-spawn mortality. We initially developed a liquid chromatography quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) analytical method, optimized sample extraction methods to minimize false negatives (i.e., maximize non-target detections), and then developed a data reduction and analysis workflow for paired highway runoff and adult coho tissue samples. These efforts ultimately resulted in the detection of a limited suite of chemical contaminants in highway runoff and coho salmon tissues, providing a proof of concept for the approach. More generally, this work provides a potential framework and approach for the prioritization and identification of pollutants that are bioavailable and potentially toxic to fish and other aquatic organisms.

Materials and methods

Chemicals

Standards and labeled analogs were obtained from commercial vendors. Most reference standards were obtained from Sigma-Aldrich (St Louis, MO, USA) and Toronto Research Chemicals (Toronto, ON, Canada). A mixture of 240 pesticides (LC/MS Pesticide Comprehensive Test Mix; p/n 5190-0551) was


This journal is © The Royal Society of Chemistry 2017
purchased from Agilent Technologies (Santa Clara, CA, USA). A complete list of reference standards, adsorbents, and the compounds in the pesticide mix is provided in the ESL‡. OPTIMA® grade methanol (MeOH) was obtained from Fisher Scientific (Fair Lawn, NJ, USA). Acetic acid (>99.7%) and ammonium acetate (HPLC grade, 97.8%) were purchased from VWR Scientific (Radnor, PA, USA). A Thermo Barnstead Nanopure Diamond UV water purification system (Dubuque, IA, USA) was used to provide 18 MΩ water. Dichloromethane (DCM; >99.9%) was purchased from Tedia (Fairfield, OH, USA).

**Sample collection**

Consistent with previous studies on the coho pre-spawn mortality syndrome, high runoff samples were collected from four storms in a stainless steel tote (Custom Metalcraft; Springfield, MO, USA) from an elevated urban highway during October and November 2015 (SR 520 in Seattle, WA, USA; GPS: 47°38′38″, −122°18′25″). The highway runoff was transported to Grovers Creek Hatchery (Poulsbo, WA; GPS: 47°46′26″, −122°33′23″), where returning adult coho salmon spawners were collected and exposed to the runoff following previously reported protocols. Briefly, for each storm event, a total of 4 ventilated PVC tubes, each containing an adult coho salmon, were placed in a polyethylene tank with either 440 L of highway runoff or clean well water (control exposure, 8 salmon total per event). Waters were aerated to ensure sufficient oxygenation, which was monitored along with temperature and pH. Symptomatic fish or time-matched controls were euthanized by blunt-force trauma. Tissues, including gill, liver, brain, and heart, were collected immediately after mortality from runoff-associated exposure, transported to the lab on ice, and stored at −20 °C. The corresponding runoff and control water samples were collected in 4 L pre-rinsed amber glass bottles without headspace and transported to the laboratory on ice. We note that these exposure studies were ongoing and not specific to this study; we opportunistically leveraged existing ecotoxicology efforts for these samples.

**Extraction methods for water samples**

Method optimization for off-line solid phase extraction (SPE) sought to maximize two criteria: the total number of non-target features and their peak area. SPE performance was compared for 3 mL, 100 mg Infinity SPE cartridges (ABS Materials, Wooster, OH, USA) and Oasis 6 mL, 200 mg hydrophilic–lipophilic-balanced (HLB) SPE cartridges (Waters, Milford, MA, USA). Both cartridges were preconditioned with 3 mL 50% (v/v) methanol in deionized water, followed by deionized water (25 mL). The runoff samples were pre-filtered (0.45 μm, polyethersulfone, hydrophilic) prior to loading (5–10 mL min⁻¹) on HLB (but not Infinity) cartridges; sample extraction was otherwise identical. After extraction, cartridges were rinsed with deionized water (25 mL), nitrogen-dried (15 min), and eluted with methanol or DCM (2×, 1 mL). Additionally, sequential elution with DCM (after elution with MeOH) to extract more/ complementary features was evaluated separately. Extracts were concentrated with nitrogen to 1 mL, sonicated (1 min) and filtered (13 mm, 0.2 μm Pall Acrodisc PTFE syringe filters) to prevent clogging of the analytical column. All water samples were extracted and analyzed within 24 hours of collection.

**Extraction methods for tissue samples**

To optimize tissue extraction methods with selective pressurized liquid extraction (SPLE) techniques, including in-cell cleanup, we used archived coho embryo and liver tissues (stored at −20 °C at the NOAA Fisheries Science Center, Seattle, WA) collected during October–November 2014. Extractions used an accelerated solvent extractor (ASE; ASE 300, Dionex, Salt Lake City, UT, USA) with 33 or 66 mL ASE extraction cells. The adsorbent combination and in-cell cleanup method was optimized toward removal of bulk interferences, improved chromatography, and numbers of target compounds and non-target features detected (Table S1†). Methanol was selected as the elution solvent to extract a similar spectrum of compounds from water and tissue samples, although acetonitrile was also assessed (Table S1†). Evaluated adsorbents included neutral alumina, basic alumina, acidic alumina, silica gel, and Florisil®, based upon anticipated suitability for lipid-rich matrices. Adsorbent performance was examined individually (n = 3; 10 g adsorbent to 1 g tissue) and in combination (adsorbent to sample mass ratios ranged from 13.5 : 1 to 30 : 1). For each extraction, an aliquot of 1 g fish tissue was homogenized with diatomaceous earth (1 : 1 m : m ratio) using mortar and pestle, then loaded above pre-cleaned adsorbents, ordered sequentially from top to bottom: 7.5 g basic alumina, 5 g silica gel, and 1 g Florisil® above a cellulose filter. Ottawa sand filled any remaining headspace in the ASE cell. The homogenates were spiked with the pesticide suite (100 ppb; for target characterization) and equilibrated (20 min) prior to extraction. Instrumental sensitivity for the pesticide standard was evaluated at both high (100 ng mL⁻¹) and low (10 ng mL⁻¹) concentrations. We selected 103 pesticides that were detected with confidence (match score >70) at 10 ng mL⁻¹ for targeted analyte evaluation. The match score is based on mass accuracy, isotopic abundance, and isotopic scoring (weighted 100%, 60%, and 50%, respectively). The rate of false negatives was calculated as the percentage of the 103 pesticides that were not detected with confidence. Samples were methanol-extracted under the following conditions: 40 °C, 1500 psi, 2 cycles (5 min each), 50% flush, and 100 s purge. Extracts were then concentrated to 1 mL in a heated bath (40 °C) under nitrogen gas (12–15 psi). Colored extracts containing particulate matter were observed after reconstitution, regardless of the adsorbent type. Thus, extracts were sonicated for one minute and filtered (13 mm, 0.2 μm Pall Acrodisc PTFE syringe filter) prior to analysis to remove particulates.

**Instrumental analysis**

Extracts were analyzed on an Agilent 1290 Infinity UHPLC system (Santa Clara, CA, USA) for separation and an Agilent 6530 Accurate-Mass Quadrupole Time-of-Flight LC-MS system with electrospray Jet Stream Technology for detection. Chromatographic separation utilized a reversed-phase C18 analytical
column (Agilent ZORBAX Eclipse Plus 2.1 x 100 mm, 1.8 μm particle size) connected to a similar C18 guard column, at a temperature of 45 °C. Chromatography employed an injection volume of 5 μL, a flow rate of 0.4 mL min⁻¹ and a binary gradient of 5 mM ammonium acetate plus 0.1% acetic acid in water (A) and 5 mM ammonium acetate plus 0.1% acetic acid in methanol (B) [5% B at 0–1 min, 50% B at 4 min, 100% B at 17–20 min, 5% B at 20.1 min; stop time 22.5 min; post-time 2 min]. Over a 22 minute run, HRMS spectra were acquired across the range 100–1700 m/z (MS only) and 50–1700 m/z (MS/MS) in 2 GHz Extended Dynamic Range resolution mode (collision-induced dissociation; data-dependent acquisition). The resolving power of the detector was 6000–12 400 within the acquisition range, and the mass accuracy is ~1 ppm. Other instrument parameters are listed in Table S2.†

Quality assurance and quality control (QA/QC)

Currently, there are no consensus QA/QC guidelines to validate analytical performance for non-target HRMS analysis. We used a variety of laboratory control samples and background detections to track instrument performance. Detector performance was monitored by checking mass accuracy and re-tuned or re-calibrated if mass error exceeded 2 ppm. A mixture of reference standards, each at 100 ng mL⁻¹, containing cotinine d₅ (RT 3.42 min), carbamazepine ¹³C₁⁵N (RT 6.47 min), and prometryn (RT 9.43 min) was analyzed every 8–10 samples to check chromatography and sensitivity during data acquisition. Area counts were monitored and expected to be within 20% of initial sensitivity and mass accuracy was limited to <5 ppm. If the instrument failed these criteria, performance was corrected by tuning or detector maintenance. Background signals were identified by analysis of lab control samples (deionized (DI) water, methanol) and method blanks (DI water through SPE, methanol through ASE), then exempted from MS/MS analysis regardless of peak intensity. We also validated consistent chromatography by monitoring triethyl citrate (RT 6.13 min), oleamide (RT 15.76 min), stearamide (RT 16.36 min), and an unidentified background ion (300.2019 Da@3.68 min). These ions were observed regardless of instrumental usage and sample type, they elute at different retention times, and therefore were used to monitor chromatographic stability. Oleamide and stearamide were identified as HRMS background ions elsewhere.‡ Throughout our analyses, the chromatography was consistent, without frequent or large shifts in retention time (<5 s) of reference ions and standards. Analytical runs also included solvent blanks (every 3 h) to monitor column carryover, which was not detected.

Data reduction and analysis

Initial data screening relied upon manufacturer software packages to identify peaks and relationships between urban runoff and different coho tissues. Agilent MassHunter Profinder software (B.06.00) was used to isolate peaks (referred to as features, or unique exact mass-retention time pairs) with mass height counts above 300 (noise level) as positive adducts ([M + H]+, [M + Na]+, and [M + NH₄]⁺) or negative adducts ([M – H]⁻) due to its fast batch-processing capacity. Profinder uses a recursive feature-finding algorithm to address missing features and simplifies results to interface with Agilent MassHunter Mass Profiler Professional (MPP, B.13.00) for subsequent data analysis. MPP is a statistical package to align, filter, and understand relationships across conditions by matching retention time (RT) (±0.01 min) and mass accuracy (±0.01 ppm) of any features extracted by MassHunter Profinder.

Alignment of features across sample groups in MassHunter Profinder was based upon matching retention time and mass within spans of 0.3 min and 30 ppm, respectively. To screen ions, only features with mass height above 5000 (S/N ~ 17) and appearing in 50% of replicates from at least one condition (2 of 4 replicates) were extracted. Characteristic fragmentation patterns become less reliable for ions with lower mass heights. The recursive feature extraction rescanned samples and extracted ions with heights above 3000 (S/N ~ 10) and match score (based on mass accuracy, isotopic abundance, and isotopic spacing) >50 to capture any features missed during the first feature extraction. Based on previous strategies to reduce false positives,§ MPP was used to align compounds with mass height above 5000 (S/N ~ 17) by retention time and mass across the different samples, limiting the identification of unaligned ions as different features.&& We applied a replication filter, excluding features not occurring in 50% of replicates, from further analysis, which typically excluded 15% and 30% of non-target features observed in stormwater and tissue samples, respectively. All features from water and tissue controls were also excluded from further analysis. Remaining features were exported to ID Browser, and screened (matched using accurate mass, isotope-pattern, and isotope abundance) against an in-house custom database that includes 377 compounds (and their isomers) previously reported in stormwater, as well as potential metabolic toxicants. The remaining features were also screened against both the Agilent Metlin (~80 000 compounds) and Forensic/Toxicology (9000 compounds) databases. For further prioritization of features matched with databases at high confidence (match score >70), we evaluated the relationship between chemical properties and observed retention time, as well as the compounds’ potential to be ionized. Lower scoring (<70) detections were selectively examined. MS/MS characterization (10, 20, and 40 eV) was performed for prioritized features by re-injecting and re-analyzing extracts, and detection was further confirmed by subsequently matching MS/MS fragmentation patterns with reference standards, when available (Fig. 1). Detections were scored against criteria proposed by Schymanski et al. for HRMS-based identification to communicate detection confidence. The highest confidence level (S1) is achieved when retention time and fragmentation patterns match reference standards, while the lowest confidence level (S5) includes measured exact mass only. In parallel to compound identification by reference standards (and for exclusion of false positives), we developed an in-house linear regression (n = 260, Fig. 2) between octanol–water partition coefficient (log Kow) and retention time to sort HRMS results. Although successful prediction of retention time to exclude interfering compounds is highly dependent on the accuracy of
log $K_{ow}$, chromatographic consistency, and the number of compounds used to develop the regression, this approach helped to confirm identification, particularly when reference standards are not available.

Results and discussion

Comparison of SPE cartridges

Optimization criteria for method development included the use of minimal sample preparation steps, minimal inspection of instrumental performance, and maximal detection of feature numbers and peak area. Non-target screening methods should value minimal sample preparation to maintain detection capability for the broadest spectrum of compounds while trying to avoid false negatives. In stormwater, some contaminants are particle-associated. We tried to avoid pre-filtration, and compared Infinity SPE cartridges, specifically designed for analysis of high-solid samples such as wastewater and turbid runoff, to Oasis HLB cartridges for extraction of highway runoff. HLB SPEs completely clogged after loading ~100 mL of unfiltered highway runoff. In contrast, >1 L of unfiltered high-solids runoff (~200 mg L$^{-1}$ of total suspended solids) was loaded on Infinity SPEs without noticeable reduction in flow rate, thus simplifying and accelerating sample processing. Thus, highway runoff required pre-filtration (0.45 μm paper filters) prior to extraction with HLB SPEs (“HLB MeOH”), while unfiltered highway runoff was loaded directly onto Infinity SPEs (“Infinity MeOH”).

Beyond comparison of cartridge extraction performance, we tested DCM as an alternative elution solvent for the Infinity SPE
cartridges (“Infinity DCM”) to yield additional, especially more hydrophobic, non-target features (Fig. 3). Using positive electrospray ionization (ESI⁺), the highest number (4802, n = 3) of non-target features was extracted with Infinity MeOH (Fig. 3A). For ESI⁻, the highest number (1182, n = 3) of non-target features was extracted with HLB MeOH. Approximately 37% of total features (1386) extracted by Infinity MeOH were not observed in HLB MeOH extracts (Fig. 3B), likely due in part to elution of particle-associated compounds from fine particles trapped on the SPE media. While we used LC/MS to focus on detection of the more polar, more hydrophilic class of urban contaminants that often represent highly bioavailable toxicants, complementary studies could also use gas chromatography or other separation and detection capabilities to characterize more volatile and hydrophobic contaminants. Notably, despite the lack of pre-filtration required for Infinity MeOH, visual inspection of total ion chromatograms (TICs) obtained by QTOF ESI⁺ full scan indicated improved baseline reduction relative to HLB MeOH (Fig. S1†).

Additionally, 10% of total features (340) eluted with DCM appeared to be unique when compared to methanol. However, sequential elution with DCM (after elution with MeOH) captured 98% of features unique to Infinity DCM extracts (Fig. S2†). Comparing the number of features observed across different peak area ranges revealed similar trends across extraction conditions (Fig. 3C). Notably, when using Infinity SPE cartridges without pre-filtration, additional detector maintenance was not required and sensitivity varied by <20% via monitoring of external reference standards. Given the strong performance and simple preparation of Infinity MeOH, further optimization of extraction strategies was not conducted. Subsequent research efforts could focus on more selective cartridges and solvent elutions (e.g. solvents with different polarities) to extract additional, or unique, subsets of contaminants.

Optimization of tissue extraction
During method development (1 g embryo and liver samples), wax-like precipitates were observed in extracts from both individual and combined adsorbents, although Florisil® did reduce such precipitation. By visual inspection, total ion chromatograms (TICs) obtained by QTOF ESI⁺ full scan (Fig. S3†) displayed variable baseline reduction across individual adsorbents and guided the selection of adsorbents for further evaluation (in conjunction with feature comparison). Florisil® and silica gel exhibited superior cleanup capabilities, as evidenced by relatively lower chromatographic baselines. Use of Florisil® and silica gel reduced the number of non-target features by >50% relative to other individual adsorbents (Table S1, †#1–6). These retained features create potential false negatives, and illustrate the challenging tradeoffs between attaining sufficient sample cleanup for good chromatography and efficient peak detection while also maintaining broad detection capabilities and minimizing the possibility of false negatives. For example, use of Florisil® increased the rate of false negatives for targeted pesticides to 21% from ≤10% for other individual adsorbents.

Early experiments suggested that a combination of multiple adsorbents might best meet selection criteria. Acidic alumina
was excluded because it retained more non-target features than other adsorbents and could induce ASE cell corrosion. Neutral alumina was not included for further optimization because minimal chromatographic baseline reduction was observed, despite good detection performance (i.e., number of non-target features). Therefore, we primarily evaluated combinations of basic alumina, silica gel, and Florisil® (Table S1, §7–18) and focused on maximizing feature detections while minimizing the adsorbent-to-sample mass ratio, in an effort to prevent false negatives and maximize peak area. Due to its previously noted matrix-cleanup capacity, 1 g of Florisil® adsorbent was included, as 5 g or 10 g Florisil® combinations resulted in a significant reduction in the total quantity of non-target features, as well as a reduction in feature peak area for features observed in extracts of all Florisil® combinations (Fig. S4†). Further, higher masses of Florisil® increased the rate of false negatives (14% for 1 g vs. 20% for 10 g). Similarly, lower masses of silica gel and basic alumina (i.e., less feature retention capacity) significantly improved extraction of non-target features (Table S1, §15 and 16). The inclusion of diatomaceous earth did not significantly impact the number of observed non-target features (~7% difference) or rate of false negatives (2% difference; Table S1, §12 and 13), and enabled homogeneous of tissue samples. The final adsorbent combination, listed top to bottom, was 7.5 g basic alumina, 5 g silica gel, and 1 g Florisil®, with 1 g diatomaceous earth as a mixing agent. To ensure sample clean-up, no further reductions in the adsorbent to sample mass ratio were tested. Extraction cell volume (33 or 66 mL) was also evaluated. Larger extraction cells required more solvent without significantly improving detection (3606 features, 33 mL vs. 3635 features, 66 mL). Therefore, 33 mL extraction cells were selected. When comparing solvents, methanol extraction yielded many more feature (3606 extraction cells were selected. When comparing solvents, methanol extraction yielded many more feature (3606 vs. 2251) and identical target (92) detections (11% false negatives) relative to acetonitrile, so methanol was selected.

Finally, we evaluated QuEChERS extraction, which is rapid, inexpensive, and has been widely applied for multi-residue analysis in fish tissues via mass spectrometry.9,38,39 The ASE in-cell extraction and cleanup method was compared against QuEChERS (Agilent, Santa Clara, CA, USA; 2 mL, 50 mg Primary Secondary Amine (PSA), 50 mg C18, 150 mg MgSO4 for fatty samples) for both liver and embryo samples. Regardless of

![Fig. 4](image_url) Comparison of extraction of observed features within different peak area ranges using accelerated solvent extraction (ASE) and QuEChERS extraction techniques for both coho salmon liver and embryo samples.

ionization modes and tissue type, ~30% more features were detected by ASE extraction relative to QuEChERS extraction (Fig. 4). Therefore, we chose the ASE in-cell extraction and cleanup methods for suspect and non-target HRMS analysis of biological tissues.

### Suspect and non-target screening workflow

HRMS ESI+/ESI− data were analyzed to identify contaminants in highway runoff and coho tissues (Fig. 1). For analysis of large, complex datasets, deconvolution is critical to prioritize candidates for subsequent MS/MS characterization, within the constraints of database capacity and logistics.40 Unlike liver and gill samples, limited sample mass was available for brain and heart, so some extractions used <1 g. To minimize any uncertainty associated with reduced sample mass, brain and heart samples were not included in the data deconvolution. Mass-Hunter Profinder was used to isolate, extract and align features (criteria described earlier), then MPP was used to align compounds by retention time and mass across the different samples, preventing the identification of unaligned ions as different features.45 We first prioritized candidates by comparing occurrence across sample groups (Venn diagrams, Fig. 1). Unique occurrences in runoff relative to controls (1024), exposed tissues relative to control tissues (1014 liver; 548 gill), and co-occurrences in the runoff and exposed fish tissues (472 liver; 492 gill) were of particular interest. Such comparisons created a more focused and shorter list of candidates for identification efforts, using the suspect screening approach outlined above.

Notably, for many possible isobaric compounds, the absence of commercial reference standards precluded confident (S1) identification. In some cases, retention time prediction based upon hydrophobicity (via log Kow prediction, data from [www.chemspider.com](http://www.chemspider.com)) helped to exclude interfering compounds and confirm identifications.40,41 With consistent chromatography, we correlated retention time and log Kow for pesticides, pharmaceuticals, neurochemicals, steroids, bio-toxins, and other contaminant standards (n = 260; Fig. 2). For example, ethoprophos (mass 242.0564; RT 9.75 min; predicted RT 9.45 min; log Kow 3.59) and 7-([b-chloroethyl] theophylline (mass 242.0570; predicted RT 3.08 min; log Kow 0.47) are isobaric within our analytical uncertainty, and yet ethoprophos detection could be confirmed using the retention time prediction. Such linear regression models can help resolve some false positives, although ionizable compounds and other structural features sometimes limit the accuracy of retention time predictions.22–24 While this approach helped to aid identification of some isobaric compounds, certain compounds with similar log Kow values (within log Kow of 2) cannot be conclusively resolved. Differentiation and identification of interfering and isobaric compounds, especially those without standards, remains a primary challenge for HRMS analysis.

### Contaminant occurrence in stormwater and exposed fish

Highway runoff is a major pollution source in high traffic areas. Contaminants originate from vehicles (such as engine oils, gas,
Table 1  Compound detections in highway runoff and runoff-exposed coho salmon tissues. Several examples of presumptive annotations, albeit at low confidence, are included as S3 features for illustrative purposes

<table>
<thead>
<tr>
<th>Compound</th>
<th>Formula</th>
<th>Theoretical mass [M + H]+</th>
<th>CAS #</th>
<th>RT (min)</th>
<th>log Kow</th>
<th>Detected in (confidence)</th>
<th>Highway runoff (ppm)</th>
<th>Mass error (ppm)</th>
<th>Liver (ppm)</th>
<th>Mass error (ppm)</th>
<th>Gill (ppm)</th>
<th>Mass error (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetanilide</td>
<td>C₇H₉NO</td>
<td>136.0757</td>
<td>103-84-4</td>
<td>4.45</td>
<td>1.66</td>
<td>S1</td>
<td>2.4</td>
<td>−2.4</td>
<td>S1</td>
<td>−1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prometon</td>
<td>C₁₀H₁₂N₅O</td>
<td>226.1662</td>
<td>1610-18-0</td>
<td>8.05</td>
<td>1.33</td>
<td>S1</td>
<td>4.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethoprophos</td>
<td>C₂₅H₃₅O₄PS₂</td>
<td>243.0637</td>
<td>13194-48-4</td>
<td>9.80</td>
<td>3.60</td>
<td>S1</td>
<td>−0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEET</td>
<td>C₁₂H₁₇NO</td>
<td>192.1383</td>
<td>134-62-3</td>
<td>7.24</td>
<td>1.96</td>
<td>S1</td>
<td>−0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cotinine</td>
<td>C₁₀H₁₄N₂O</td>
<td>177.1022</td>
<td>486-56-6</td>
<td>3.45</td>
<td>−0.23</td>
<td>S1</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caffeine</td>
<td>C₈H₁₀N₄O₂</td>
<td>195.0877</td>
<td>58-08-2</td>
<td>3.92</td>
<td>−0.10</td>
<td>S1</td>
<td>3.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (or 5)-Methyl-1H-benzotriazole</td>
<td>C₂H₇N₃</td>
<td>134.0713</td>
<td>136-85-6</td>
<td>4.97</td>
<td>1.8</td>
<td>S1</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N,N’-Dicyclohexyurea</td>
<td>C₁₁H₁₈N₅O</td>
<td>225.1961</td>
<td>2387-23-7</td>
<td>8.44</td>
<td>3.85</td>
<td>S2</td>
<td>0.5</td>
<td>−3.9</td>
<td>S2</td>
<td>−3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dihydroactiniodiolide</td>
<td>C₁₁H₁₈O₂</td>
<td>181.1223</td>
<td>17092-92-1</td>
<td>6.42</td>
<td>2.26</td>
<td>S2</td>
<td>0.3</td>
<td>−3.9</td>
<td>S2</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ε-Caprolactam</td>
<td>C₆H₁₀NO</td>
<td>114.0914</td>
<td>105-60-2</td>
<td>3.56</td>
<td>−0.32</td>
<td>S2</td>
<td>1.0</td>
<td>−0.2</td>
<td>S2</td>
<td>2.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(2-Phenylethyl)-acetaimide</td>
<td>C₁₀H₁₇NO</td>
<td>164.1070</td>
<td>877-95-2</td>
<td>5.08</td>
<td>1.12</td>
<td>S2</td>
<td>−0.3</td>
<td>−0.2</td>
<td>S2</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choline</td>
<td>C₂H₁₄NO</td>
<td>104.1070</td>
<td>62-49-7</td>
<td>0.70</td>
<td>−3.70</td>
<td>S2</td>
<td>1.6</td>
<td>2.8</td>
<td>S2</td>
<td>2.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triadimefon</td>
<td>C₁₄H₂₃Cl₁₂O₂</td>
<td>294.1004</td>
<td>43121-43-3</td>
<td>9.05</td>
<td>2.77</td>
<td>S2</td>
<td>−0.4</td>
<td>−4.7</td>
<td>S2</td>
<td>−4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Piperidinone</td>
<td>C₂H₆NO</td>
<td>100.0757</td>
<td>675-20-7</td>
<td>2.60</td>
<td>−0.45</td>
<td>S2</td>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexa(methoxymethyl)melamine</td>
<td>C₁₃H₁₆N₂O₆</td>
<td>391.2300</td>
<td>3089-11-0</td>
<td>6.80</td>
<td>3.07</td>
<td>S2</td>
<td>−5.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norvaline</td>
<td>C₈H₁₄NO₂</td>
<td>118.0863</td>
<td>760-78-1</td>
<td>0.84</td>
<td>0.38</td>
<td>S2</td>
<td>−2.0</td>
<td>−3.0</td>
<td>S2</td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-(3-Phenylpropyl)pyridine</td>
<td>C₁₄H₁₃N</td>
<td>198.1277</td>
<td>2110-18-1</td>
<td>4.70</td>
<td>3.60</td>
<td>S2</td>
<td>−2.0</td>
<td>−3.0</td>
<td>S2</td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimethadione</td>
<td>C₁₁H₂₀NO₅</td>
<td>130.0499</td>
<td>695-53-4</td>
<td>0.89</td>
<td>0.15</td>
<td>S3</td>
<td>−0.8</td>
<td>5.8</td>
<td>S3</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poloxalene</td>
<td>C₁₂H₁₈O₄</td>
<td>165.1121</td>
<td>9003-11-6</td>
<td>3.27</td>
<td>−1.52</td>
<td>S3</td>
<td>1.3</td>
<td>2.3</td>
<td>S3</td>
<td>2.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fenbutene</td>
<td>C₁₃H₁₇O₂</td>
<td>191.1667</td>
<td>5984-83-8</td>
<td>5.83</td>
<td>3.21</td>
<td>S3</td>
<td>−1.0</td>
<td>−1.2</td>
<td>S3</td>
<td>−1.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morforex</td>
<td>C₁₃H₁₇N₂O</td>
<td>249.1961</td>
<td>41152-17-4</td>
<td>10.32</td>
<td>1.71</td>
<td>S3</td>
<td>−2.3</td>
<td>−3.1</td>
<td>S3</td>
<td>−3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Octodrine</td>
<td>C₈H₁₄N</td>
<td>130.1590</td>
<td>543-82-8</td>
<td>3.74</td>
<td>2.69</td>
<td>S3</td>
<td>0.8</td>
<td>−0.5</td>
<td>S3</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paramethasone</td>
<td>C₁₃H₁₇Cl₂O₂</td>
<td>393.0727</td>
<td>53-33-8</td>
<td>4.23</td>
<td>1.53</td>
<td>S3</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-Despropyl ropinirole</td>
<td>C₁₀H₁₂NO</td>
<td>219.1492</td>
<td>10696-16-9</td>
<td>7.63</td>
<td>1.59</td>
<td>S3</td>
<td>−0.5</td>
<td>−0.4</td>
<td>S3</td>
<td>−0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impacarzine</td>
<td>C₁₄H₂₁N₂O₂</td>
<td>494.4429</td>
<td>41340-39-0</td>
<td>17.86</td>
<td>8.12</td>
<td>S3</td>
<td>−0.16</td>
<td>−0.15</td>
<td>S3</td>
<td>−0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Benzothiazolsulfonic acid</td>
<td>C₈H₁₄NO₂S₂</td>
<td>215.9784</td>
<td>941-57-1</td>
<td>3.65</td>
<td>1.67</td>
<td>S3</td>
<td>−0.07</td>
<td>−1.7</td>
<td>S3</td>
<td>4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3'-Aminobenzenilide</td>
<td>C₁₁H₁₂N₂O</td>
<td>213.1022</td>
<td>102-07-8</td>
<td>6.12</td>
<td>1.48</td>
<td>S3</td>
<td>−0.6</td>
<td>−3.1</td>
<td>S3</td>
<td>1.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
brake fluid, maintenance, and anti-corrosion compounds) and combustion emissions, with secondary sources related to roadways and their maintenance. While certain contaminant classes (e.g., metals, PAHs, and PAH transformation products) that can induce adverse effects in fish,\textsuperscript{42-43} are well described in stormwater\textsuperscript{11,12,14} and roadway runoff,\textsuperscript{4} numerous additional pollutants of concern for aquatic organisms have been detected in stormwater runoff using LC-based analyses, such as benzo-triazoles,\textsuperscript{44,45} benzo-hiazoles,\textsuperscript{46} and pesticides.\textsuperscript{47} For example, benzo-triazoles are used as vulcanization accelerants in tire production and are toxic to fish,\textsuperscript{48} including cytotoxicity to fish gills.\textsuperscript{49} Nevertheless, many contaminants in these complex stormwater mixtures remain uncharacterized despite observations of stormwater-derived toxicity.\textsuperscript{1}

In the current study, we analyzed highway runoff samples collected after dry summer periods. These typically have higher levels of accumulated contaminants and represent the worst-case scenario for contaminant concentration and subsequent toxicity to exposed organisms.\textsuperscript{30-32} Stormwater derived contaminants were detected in water samples and in the tissues of runoff-exposed coho salmon (Table 1). Notably, exposure periods were short, only 1–3 hours prior to coho mortality, so tissue analysis really represents initial, rapid contaminant uptake processes in these fish. In both sample types, we often collected MS/MS spectra for many high priority features, but lacked reference standards and MS/MS library spectra to preclude conclusive identification of such features (S3). The need for continued development of standardized, searchable libraries of MS/MS data, particularly for contaminants expected in systems such as urban stormwater, is acute. An additional 17 features were identified (seven S1; 10 S2) via reference standard or database comparison, including both natural compounds and anthropogenic contaminants. For example, ethoprophos, an organophosphate insecticide, was detected in the highway runoff (S1, confirmed by standard), but not fish tissue (Fig. S51). DEET (N,N-diethyl-meta-toluamide), a common active ingredient in insect repellents widely reported as an environmental contaminant,\textsuperscript{53} caffeine, and cotinine were also detected in highway runoff (but not fish tissues) through suspect screening and confirmed with reference standards. Such detections align well with previous observations of near ubiquitous occurrence of diverse emerging organic contaminants (particularly pharmaceuticals and pesticides) in urban stormwater and surface waters.\textsuperscript{54-56}

Suspect screening techniques may be especially valuable for helping to focus analytical efforts on specific mechanisms of toxicity and high priority contaminant classes. For example, acetanilide, used for dye production and rubber vulcanization, was detected by a suspect screening approach using the Agilent Metlin database. Once prioritized as a compound of interest due to its metabolic toxicity profile, it was identified (S1) both in runoff and in liver and gill of runoff-exposed coho (Fig. 5 and S6), with mass errors <5 ppm. Acetanilide was not observed in the control exposure, control fish tissues, or lab controls. Via comparison with a calibration curve (5 μg L\textsuperscript{-1} to 1 mg L\textsuperscript{-1}), we estimate ~20 ng g\textsuperscript{-1} and ~6 ng L\textsuperscript{-1} of acetanilide in the liver tissues and acetanilide in the runoff, respectively, neglecting possible matrix effects. We also observed acetanilide in the gill, but the estimated concentration was out of the calibration range and was not reported. A bioaccumulation factor for acetanilide in the liver was calculated as ~3000 L kg\textsuperscript{-1}, indicating its potential for bioaccumulation in aquatic organisms.

To the best of our knowledge, acetanilide has not been previously reported in urban waters or exposed fish, with the exception of one report of acetanilide in industrial wastewater from a specialty chemicals manufacturer.\textsuperscript{57} Though we currently have little insight into its toxicological implications, acetanilide or its metabolic products are linked to methemoglobinemia\textsuperscript{46} and cancer\textsuperscript{58} in humans, and can impact aquatic species.\textsuperscript{59} We also detected 4 (or 5)-methyl-1H-benzotriazole (S1), a commonly used corrosion inhibitor that has demonstrated aquatic toxicity at elevated concentrations\textsuperscript{60,62} and has been previously linked to roadway sources,\textsuperscript{63} in highway runoff. Such detections, though limited in number, pose hypotheses for subsequent toxicological investigation, and also may aid source identification efforts if they are strongly correlated with toxicants. They also highlight the value of integrating specific contaminant attributes such as sources (e.g., rubber vulcanization accelerants and other roadway-related compounds) or toxicological modes of action (e.g., metabolic poisons) into comprehensive screening databases,\textsuperscript{64} particularly in cases when biological studies provide available evidence of modes of action. Such capabilities can help to better link chemical and biological analysis of water quality and improve screening capabilities for impacted receiving waters as potential contaminant detections quickly translate to source and hazard outcomes.

Implications

HRMS typically detects dozens to thousands of compounds with a broad range of physicochemical properties in a single sample. This screening capability is the most valuable characteristic of this analytical technique, but laborious and slow data analysis currently precludes high throughput. In this study, we developed analytical methods and a data reduction workflow to screen for contaminants in paired highway runoff and runoff-exposed fish. In particular, we focused on the use of contaminant uptake and bioavailability processes to guide our workflow development and data prioritization, and to provide biological relevance to our identification of novel contaminants in highway runoff and runoff-exposed fish tissues. Generally, our detection of a range of contaminants (with differing structures, polarities, and sources) provides evidence that the extraction and analytical methods reported here are appropriate for broad-spectrum screening of water quality, although further work is needed to confirm a larger number of identifications in these samples and understand their full implications for ecosystem health.

Suspect analysis techniques that utilize in-house and commercial databases proved to be especially valuable because they can focus analytical efforts on potentially high risk detections and specific modes of action.\textsuperscript{64} To integrate site-specific ecotoxicological data into contaminant screening, we
recommend a continued focus on carefully integrating HRMS analysis with bioanalytical results (e.g. bioassays, toxicological modes of action) as an aspect of toxicant identification efforts, especially for screening potential toxicological hazards of complex chemical mixtures. The acetanilide detection presents an example of how a focus on a specific mode of action could both prioritize HRMS detection efforts and guide development of toxicology-related hypotheses that are testable with bioassays. Particularly for suspect and non-target screening efforts focused around a specific biological or toxicological question, database development should focus on organizing and indexing potential environmental contaminants across common biological modes of action or adverse toxicological potentials into comprehensive (and broadly available) screening databases to better exploit HRMS screening capabilities for complex mixtures. The need for additional MS/MS data that represent a more diverse suite of environmental contaminants to aid identifications is particularly acute.

As we improve our understanding of urban stormwater composition, quantifying the sources and loads of
contaminants is critical to guiding stormwater quality management actions and protecting aquatic ecosystems.\textsuperscript{55} Our non-target data clearly indicate the presence of many uncharacterized compounds in highway runoff, an important source of contaminants to urban waters. More importantly, many non-target features are detected in exposed fish tissues even after short exposure periods, indicating bioavailability and therefore a potential for adverse effects. One important limitation of the approach reported here is that many bioavailable and toxicologically relevant non-target features are detected in highway runoff, but are not necessarily observed in fish due to \textit{in vivo} biotransformations. Accounting for such features is challenging, as many information gaps on the chemical and ecotoxicological properties of emerging contaminants and their transformation products still exist.\textsuperscript{65} For instance, PAH metabolites formed in the aquatic environment or \textit{in vivo} biotransformation often exhibit a wide range of polarities and can cause adverse toxicological outcomes, yet these same compounds can lack authentic reference standards.\textsuperscript{41} Such metabolites are excellent candidates for HRMS workflows that combine detection, fractionation, and toxicity testing for confident identification.\textsuperscript{41,66} Additionally, consensus QA/QC protocols and workflows for non-target and suspect screening of contaminants in the environmental samples would be valuable. We propose the use of laboratory controls (including biological controls) to better prioritize non-target features for identification, as well as the inclusion of representative background ions and representative reference standards with a broad spectrum of physicochemical properties to better track HRMS method performance.

In conclusion, we developed HRMS methods to aid the chemical and biological characterization of urban stormwaters. These efforts integrate simple extraction procedures, non-target HRMS analysis, and a first-pass assessment of feature bioavailability and toxicological relevance in screening of complex, environmental mixtures. Our results demonstrate detection of several thousand distinct chemical features in runoff from a high-traffic arterial, the vast majority of which remain unidentified and uncharacterized in terms of aquatic toxicity. Our screening efforts detected many HRMS features in stormwater, including a significant number that also occur in the gill, liver, and other tissues of stormwater-exposed fish, indicating their bioavailability and uptake. The combination of high-end HRMS analytical chemistry and aquatic toxicology is a promising tool in identifying potential pollutants in complex environmental mixtures and guiding future source identification and control efforts.

**Conflicts of interest**

There are no conflicts to declare.

**Acknowledgements**

We thank the NOAA Ecotox Program, Suquamish Tribe’s Grovers Creek hatchery, and Washington State University research teams for their assistance with runoff and tissue collection. The authors also acknowledge the anonymous reviewers whose comments greatly improved this work.

**References**
