Environmental Science & Technology

Article

Cite This: Environ. Sci. Technol. 2019, 53, 77−87

pubs.acs.org/est

Fall Creek Monitoring Station: Highly Resolved Temporal Sampling to Prioritize the Identification of Nontarget Micropollutants in a Small Stream

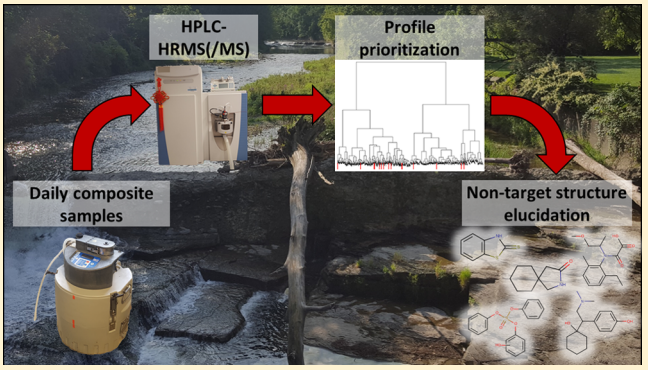
Corey M. G. Carpenter,[†] Lok Yee J. Wong,[‡] Catherine A. Johnson,[†] and Damian E. Helbling*[,†]

[†]School of Civil and Environmental Engineering, Cornell University, Ithaca, New York 14853, United States

[‡]Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853, United States

Supporting Information

ABSTRACT: The goal of this research was to comprehensively characterize the occurrence and temporal dynamics of target and nontarget micropollutants in a small stream. We established the Fall Creek Monitoring Station in March 2017 and collected daily composite samples for one year. We measured water samples by means of high-resolution mass spectrometry and developed and optimized a postacquisition data processing workflow to screen for 162 target micropollutants and group all mass spectral (MS) features into temporal profiles. We used hierarchical clustering analysis to prioritize nontarget MS features based their similarity to target micropollutant profiles and developed a high-throughput pipeline to elucidate the structures of prioritized nontarget MS features. Our analyses resulted in the identification of 31 target micropollutants and 59 nontarget micropollutants with varying levels of confidence. Temporal profiles of the 90 identified micropollutants revealed unexpected concentration−discharge relationships that depended on the source of the micropollutant and hydrological features of the watershed. Several of the nontarget micropollutants have not been previously reported including pharmaceutical metabolites, rubber vulcanization accelerators, plasticizers, and flame retardants. Our data provide novel insights on the temporal dynamics of micropollutant occurrence in small streams. Further, our approach to nontarget analysis is general and not restricted to highly resolved temporal data acquisitions or samples collected from surface water systems.



INTRODUCTION

Organic micropollutants are anthropogenic organic chemicals that are present in the environment at low concentrations.[1] Decades of monitoring studies have identified hundreds of organic micropollutants in natural and engineered water systems.[2,3] However, conventional micropollutant monitoring strategies focus on a finite number of target micropollutants,[4] a practice which is known to underestimate micropollutant exposure risk by a factor of at least 2−10, even when considering just one class of micropollutant (e.g., pesticides).[5] Additionally, infrequent grab samples do not adequately capture the expected temporal dynamics of micropollutant concentrations (e.g., seasonal runoff from agriculture) and will likely miss peak events associated with the greatest ecological risks. This problem is exacerbated in small streams, which are more sensitive to changing hydrologic conditions and have lower dilution rates than larger rivers.[6−8] Additionally, more than 110 million people in the U.S. are supplied by public drinking water systems that rely at least in part on small streams (defined as intermittent, ephemeral, and headwater streams),[9] although micropollutant monitoring in small streams has been rather limited.[10,11] New approaches are needed to more comprehensively characterize the occurrence and temporal dynamics of micropollutants in water systems, and particularly in small streams.

One way to broaden the scope of micropollutant monitoring is to complement conventional targeted screening with nontargeted screening techniques. Nontargeted screening by means of high-resolution mass spectrometry and postacquisition data processing has emerged as an effective tool to comprehensively characterize the occurrence of organic chemicals in a variety of sample types including sediments,[12] animal fat,[13] dust,[14] wastewater effluent,[15] and surface water.[16] However, two major challenges still limit the widespread use of nontargeted screening for routine micropollutant monitoring. First, high-resolution mass spectral acquisitions contain thousands of nontarget MS features. Prioritization of the

Received: September 20, 2018
Revised: November 17, 2018
Accepted: November 25, 2018
Published: November 26, 2018

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

most relevant nontarget MS features for structural elucidation is essential.[17] Most previously reported prioritization strategies are driven by analytical or chemical data; for example, others have prioritized nontarget MS features based on relative peak intensity,[18] the presence of strong isotope signatures (e.g., chlorine atoms),[19] or evidence that a nontarget MS feature is part of a homologous series.[20] Whereas these prioritization strategies have led to the successful identification of nontarget micropollutants, prioritization strategies that are coupled with features of the system being studied may lead to more generalizable conclusions about micropollutant occurrence and temporal dynamics. Second, there is no broadly accepted approach to elucidate structures of nontarget MS features. A variety of vendor software has emerged in recent years, though proprietary algorithms for structural elucidation do not afford the transparency needed to fully evaluate annotations of nontarget MS features. A number of open-source tools[21−25] and databases[26−29] have also been developed to address certain aspects of nontarget analysis and structural elucidation, though there is a need to develop and optimize data analysis pipelines to enable high throughput structural elucidation of nontarget MS features.

Another way to gain insights on micropollutant occurrence and temporal dynamics in a water system is to implement a more continuous sampling strategy. For example, intermittent sampling of surface water systems around the U.S. over several years revealed distinct temporal profiles of pesticides that peaked during the agricultural growing season.[30] Daily composite sampling of a surface water collection system revealed that antecedent and postapplication rain events trigger glyphosate transport from runoff-prone soils.[31] Recent studies in the Rhine River have demonstrated that daily composite sampling over long periods of time can reveal unexpected temporal dynamics of target micropollutants[32] and trends in nontarget MS features can be used to detect contamination events.[33]

The objectives of this study were to (i) establish a continuous monitoring station in a small stream to generate highly resolved temporal profiles of target micropollutants and nontarget MS features; (ii) use the highly resolved temporal profiles to prioritize nontarget MS features for structural elucidation; and (iii) explore the highly resolved temporal data to reveal the temporal dynamics of micropollutant occurrence and gain fundamental insights on contaminant sources, fate, and transport phenomena. We selected the drinking water intake on Fall Creek (Ithaca, NY) as the location of the monitoring station. We collected daily composite samples from Fall Creek for one year and measured the samples by means of high-resolution mass spectrometry. The data were used to develop and optimize a nontarget MS feature prioritization workflow and structural elucidation pipeline using open-source tools. Our approach led to the discovery of several types of micropollutant temporal profiles, some of which exhibited strong concentration-discharge dependencies. We detected 31 target micropollutants and elucidated the structures of 59 nontarget micropollutants with varying levels of confidence. These data provide novel insights on the temporal dynamics of micropollutant occurrence in small streams and the most comprehensive assessment of polar organic micropollutant exposure that is presently possible.

MATERIAL AND METHODS

Study Area. We selected the drinking water intake located on Fall Creek as the location of the Fall Creek Monitoring Station (FCMS). Fall Creek is a small tributary of Cayuga Lake located in Ithaca, New York and is the source of drinking water for over 30 000 people. The Fall Creek watershed upstream of the FCMS has an approximate area of 320 km^2 and over 22 000 people live within the watershed boundaries. A U.S. Geological Survey (USGS) stream gage located less than 2 km downstream from the FCMS recorded an average discharge of 7 $m^3 \cdot s^{-1}$ (median 4.0 $m^3 \cdot s^{-1}$, range 0.7−70 $m^3 \cdot s^{-1}$) during the study period,[34] reflecting both the small size of Fall Creek and variable streamflow. A GIS analysis of the watershed revealed that 29% and 16% of the land cover is defined as pasture/hay and cultivated crops, respectively.[35] Additionally, two sewage treatment plants (STPs) discharge directly into Fall Creek and 45% of the population utilizes onsite wastewater treatment and disposal systems.[36,37] Therefore, Fall Creek likely receives intermittent loadings of a variety of agricultural and STP-derived micropollutants. A map of the study area is provided in Figure S1 of the Supporting Information (SI).

Sample Collection. We used an ISCO automatic sampler (6712 Full-Size Portable Sampler, Teledyne Isco) to collect daily, time-proportional composite samples directly from the raw water intake of the Cornell Water Filtration Plant. We collected approximately 1 L daily samples through Teflon-lined polyethylene tubing in 1.8 L glass bottles using a 20 min sampling interval. Additionally, we obtained weekly field blanks by collecting 1 L of nanopure water through the automatic sampler. Teflon-lined polyethylene tubing was replaced approximately every four months for precautionary purposes. We retrieved the samples from the FCMS at weekly intervals and stored them at 4 °C until preparation, which was always within 24 h of retrieval. Daily samples and weekly field blanks were collected between March 2017−2018.

Standards and Reagents. Details on the sources, preparation, and storage of authentic reference standards and reagents are provided elsewhere.[3] MS acquisition parameters for 162 target micropollutants and 33 isotope labeled internal standards (ILISs) are provided in Tables S1 and S2 of the SI.

Sample Preparation and Analysis. We prepared the samples and field blank at weekly intervals by transferring 45 mL of each sample into a 50 mL conical tube (VWR), centrifuging at 4700 rpm (4816 g) for 15 min (Sorvall Legend XTR, Thermo Scientific), amending with 0.1% (v/v) 1 M ammonium acetate buffer, and adjusting the pH to 6.5 ± 0.2 using 5% formic acid and 1.4 N ammonia. Exactly 8 mL of each pH-adjusted sample was then transferred into triplicate 10 mL glass sample vials (Chromacol, Thermo Fisher Scientific). We also prepared one quality control sample each week by diluting a mixture of reference standards to 250 $ng \cdot L^{-1}$ in nanopure water. Each sample, field blank, and quality control sample was then spiked with 2 ng each of a mixture of 33 ILISs and stored at 4 °C until analysis, typically within 1 week of preparation.

We adapted a previously described analytical method that implements large volume injection (LVI) and high performance liquid chromatography (HPLC) coupled with high resolution mass spectrometry (QExactive hybrid quadrupole orbitrap, Thermo Fisher Scientific).[38,39] Briefly, samples were injected at 5 mL volumes onto a Hypersil GOLD aQ trap column (2.1 × 20 mm, particle size 12 μm, Thermo Fisher

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

Scientific) at room temperature (21−22 °C) and eluted with a mobile phase gradient onto an XBridge C-18 analytical column (2.1 × 50 mm, particle size 3.5 μm, Waters) at 25 °C for analyte separation. Full scan mass spectra were acquired in positive ionization mode at a resolution of 140 000 at 200 m/z. Data dependent MS^2 spectra were acquired at the exact masses and retention times of all target micropollutants and prioritized MS features (see details in the following sections); the data dependent MS^2 inclusion list was continuously updated throughout the duration of the study. Additional data dependent MS^2 spectra were acquired for the most intense MS features if the inclusion list was not triggered. A total of three MS^2 scans were recorded after each full scan. Additional details on the analytical method including the mobile phase gradient and the MS and MS^2 acquisition parameters are provided in Tables S3 and S4 of the SI. We note that the sample preparation and analytical methods selected for this work constrain the scope of our nontarget analysis to include polar to semipolar organic molecules that can be ionized in positive mode electrospray ionization.

Peak Picking and Profile Generation. We developed and optimized an automated workflow for the characterization of target micropollutants and nontarget MS features using enviMass v3.413.[22] Our workflow consists of nine steps: (1) convert instrument .RAW files into .mzXML files using ProteoWizard v3.0.10827;[40] (2) identify fully resolved chromatographic peaks using the peak picking settings provided in Table S5 and assign a mass-to-charge ratio (m/z), a retention time (RT), and an intensity to each of the picked peaks; (3) recalibrate the m/z of each of the picked peaks based on the measured m/z of the ILISs in each sample (controls for mass drift); (4) exclude picked peaks that are not present in all three of the triplicate sample injections; (5) identify picked peaks in the field blank that have matching m/z and RT with picked peaks in each sample and calculate a sample-to-blind intensity ratio; (6) annotate any target micropollutants and ILISs based on expected m/z, RT, and isotopic signature; (7) group picked peaks with the same m/z and RT across samples into profiles; (8) normalize the intensities of all profiles based on the measured intensities of the ILISs (controls for variable matrix effects); and (9) identify isotopologues and adducts associated with a parent chemical and group into components. The output from this workflow is a final profile list of MS features described by their average m/z (tolerance set at 3 ppm) and RT (tolerance set at 30 s) among all of the samples, and their ILIS normalized intensity in each sample (a surrogate for concentration). Each MS feature in the profile list is further annotated as a target micropollutant, an ILIS, or a nontarget MS feature. We optimized the enviMass settings provided in Table S5 to maximize the number of true positive micropollutant annotations using data acquired from a quality control sample and the first several weeks of daily samples obtained from the FCMS. We first performed a conventional target screening using XCalibur v3.1 (Thermo Fisher Scientific) as previously described,[3] and then compared the results with the micropollutant annotations from our enviMass workflow. All settings were iteratively optimized to maximize the number of true positive micropollutant annotations in the quality control sample and to generate identical micropollutant profiles between data processed by XCalibur and enviMass among the samples from the FCMS.

Filtering and Clustering of MS Feature Profiles. As a means to prioritize profiles of nontarget MS features for structural elucidation, we first applied a series of data reduction filters including thresholds for sample-to-blind ratio (≥10), mean trend intensity (≥10^5), and RT (≥6.5 min) assigned to each MS feature profile. We also excluded all MS feature profiles associated with lower order isotopologues and adducts. Finally, we excluded all MS feature profiles that did not contain at least 30 total detections and at least 10 consecutive detections among the annual daily samples. We then grouped the remaining MS feature profiles by means of a hierarchical clustering analysis (HCA) using the *hclust* function in the R Statistical Software v3.3.3.[41] using Ward's agglomeration method and Euclidean distance matrices[3,12] based on the similarity of their ILIS normalized intensities over time; the ILIS normalized intensities of each MS feature profile were further normalized to their maxima so that the profiles were clustered based on their temporal trends and not absolute intensity. Nontarget MS features that clustered within or adjacent to localized clusters containing target micropollutants were prioritized for structure elucidation.

Structural Elucidation of Nontarget MS Features. We developed and optimized an automated pipeline to assign chemical structures to the prioritized nontarget MS features using a series of self-written R scripts and publically available R packages. We note that high resolution mass spectrometry alone can lead to putative structural assignments by means of spectral annotation, but multiple analytical techniques are required to unequivocally annotate the structure of an unknown chemical. Therefore, all chemical structures are assigned a confidence level based on previously established criteria summarized in Table S6.[42] The accurate masses (m/z) assigned to each of the prioritized nontarget MS features are considered to be exact masses of interest (level 5). For each nontarget MS feature profile, the sample with the highest ILIS normalized intensity is selected and the package RMassBank[43] is used to extract MS and MS^2 data into R. The package GenFormR[44] is then used to predict molecular formulas based on MS isotopic signatures, MS^2 fragments, and a series of user-defined atomic constraints (C, H, N, O, P, F, S, Cl, I, and Br) (level 4). We then used the package MetFragR[21] to compare measured MS^2 fragmentation patterns with in silico fragmentation patterns of all chemicals in the PubChem[45] online database with a molecular formula that matches the predicted molecular formula. In cases where an unequivocal molecular formula could not be assigned to a nontarget MS feature (i.e., more than one molecular formula assigned with a similar score), we used its accurate mass to search the PubChem database instead. The resulting list of candidate chemical structures (level 3) was ranked based on the weighted scoring of six factors: fragment score (0.30); Metfusion score[46] (0.30); number of PubChem references (0.05); number of PubChem patents (0.05); RT score (0.15); and presence in SusDat, the merged NORMAN suspect list[47] (0.15). All scores are normalized (0−1) to the top ranked candidate for each individual scoring metric. These types of scoring factors have been used in previous studies,[14,21] but the weighting factors were optimized to maximize the correct annotations of target micropollutants among the samples collected from the first few weeks of the FCMS. The most plausible structure(s) was selected based on the scoring and other chemical information (name, use-class, clustering with target micropollutant) (level 2P). Finally, the fragmentation patterns of the selected candidate structures were compared to online mass spectral libraries including MassBank of North America (MoNA)[28] and

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

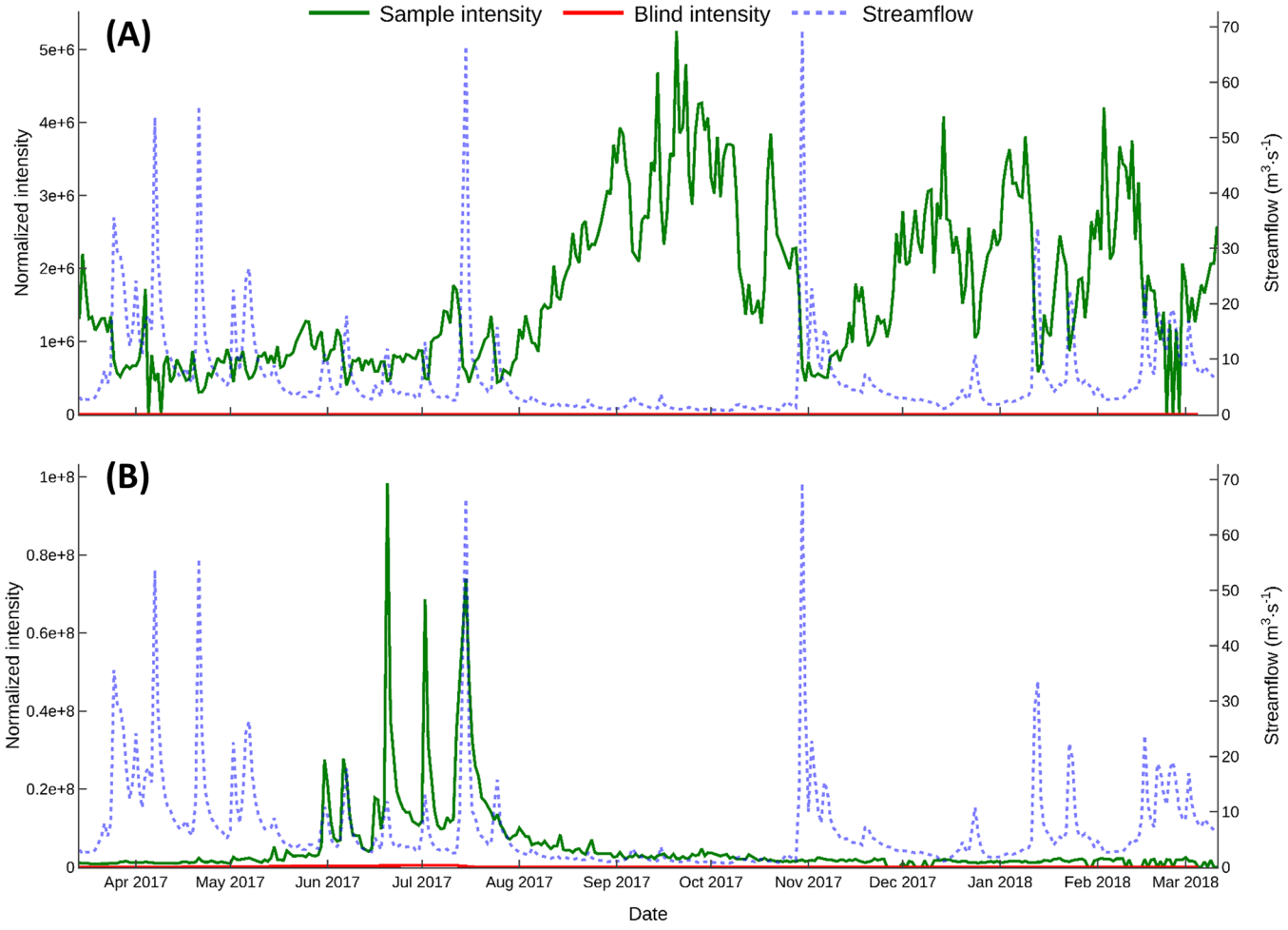


Figure 1. Micropollutant temporal trend profiles and streamflow for target micropollutants (A) STP-derived desvenlafaxine and (B) agriculture-derived atrazine.

mzCloud[26] (level 2L) or an in-house library that included major diagnostic MS^2 fragments and RTs for over 600 micropollutants (level 1L). If available, a pure reference standard was obtained for unambiguous confirmation (level 1). The complete pipeline to assign chemical structures to the prioritized nontarget MS features is available for download at github.com/cmc493.

RESULTS AND DISCUSSION

Peak Picking and Profile Generation. Our sample set consisted of 361 daily composite samples (4 samples were lost throughout the year) and 51 field blanks. Our optimized peak picking workflow identified 18.3 million fully resolved chromatographic peaks among the 361 samples, some of which represent the same constituent identified in multiple samples. The workflow also includes a replicate filter, which excludes picked peaks that are not present in all three of the triplicate sample injections; the replicate filter excluded approximately 25% of the picked peaks, leaving 13.8 million peaks to be included in the resulting 300 309 profiles of MS features.

MS Feature Profiles of Target Micropollutants. Our optimized peak picking and profile generation workflow resulted in 31 target micropollutant annotations among the 162 target micropollutants included in the study. Although some of the target micropollutants were sporadically detected in few samples and did not generate a continuous temporal profile, many target micropollutants were more ubiquitously present and generated temporal profiles that provide new insights into temporal dynamics of micropollutant abundance in small streams. For example, some frequently detected target micropollutants that are often associated with STP outfalls (e.g., desvenlafaxine, fexofenadine, and lamotrigine) had relatively high normalized intensities during periods of low streamflow and relatively low normalized intensities during periods of high streamflow, as demonstrated in Figure 1A for desvenlafaxine. This apparent negative association between micropollutant abundance and streamflow suggests a continuous loading into Fall Creek and subsequent dilution during periods of wet weather. Conversely, the abundance of other frequently detected target micropollutants that are often associated with agricultural activities (e.g., atrazine, metolachlor, simazine) exhibited a strong positive association with streamflow during the agricultural season (approximately June through August) and had low to no abundance throughout the remainder of the study period, as demonstrated in Figure 1B for atrazine. These data provide evidence that micropollutants associated with agricultural activities can be mobilized during precipitation events and their abundance increases even as the amount of water in Fall Creek increases, reflecting significantly increased mass loading during runoff events. Whereas these general types of associations between streamflow and the

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

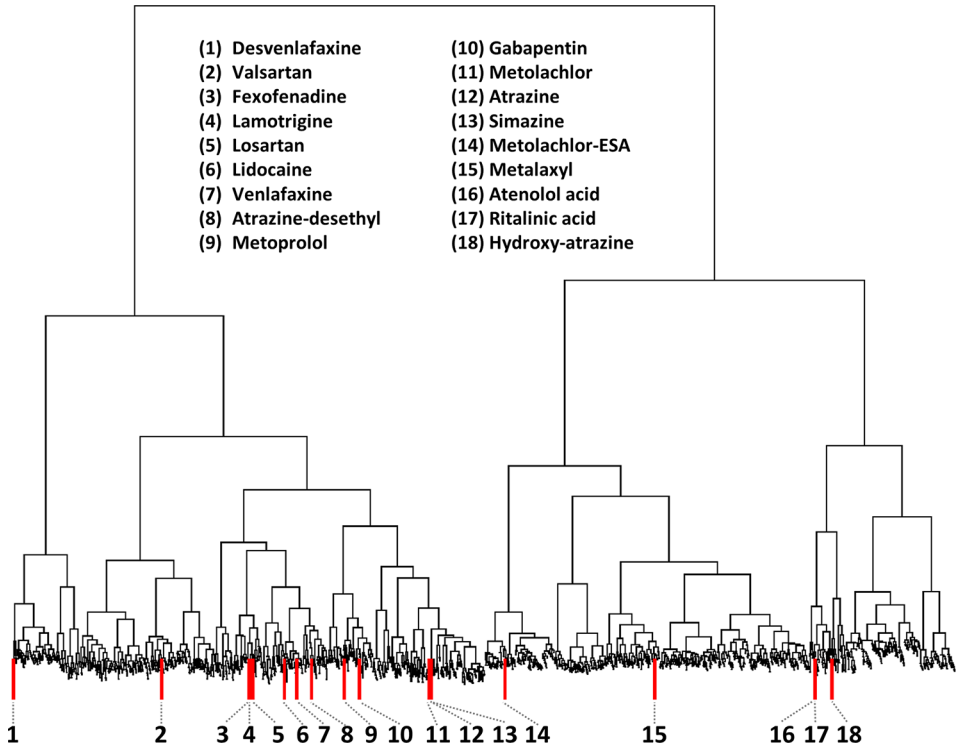


Figure 2. Dendrogram of filtered profiles (n = 1981). Marked profiles indicate target micropollutants.

abundance of STP-derived[48−50] and agriculture-derived[7,51,52] micropollutants have been previously observed, the data reported here offer the first description of the dynamics of micropollutant abundance in daily composite samples collected from a small stream over an annual period. These data have value for evaluating the dynamics of chemical exposure in small streams and could inform decision making at water utilities using small streams as a drinking water source. The temporal profiles of the other 29 target micropollutants are plotted along with streamflow in Figures S2−S30 of the SI.

Filtering and Clustering of MS Feature Profiles. As a means to prioritize profiles of nontarget MS features for structural elucidation, we applied a series of data reduction filters to exclude profiles that do not meet certain quality control metrics. We first assigned conservative thresholds for sample-to-blind intensity ratio (≥10), median trend intensity (≥10^5), and RT (≥6.5 min) based on a meta-analysis of these properties in the full profile list (Figure S31). The sample-to-blind ratio filter excluded 28% of the profiles that were present in both the samples and field blanks at median sample-to-blind intensity ratios less than 10. The mean trend intensity filter excluded 45% of the remaining profiles with mean trend intensities less than 10^5. The RT filter excluded 6% of the remaining profiles containing picked peaks that eluted near the solvent front and had poor chromatographic retention and peak shape; we note that some very polar micropollutants (e.g., metformin) may have been excluded from the profile list after applying this filter. We next excluded approximately 1% of the remaining profiles associated with lower order isotopologues and adducts; it is worth noting that most of the profiles excluded here would not have been excluded during any other filtering or prioritization step and would therefore have been included in the final profile list. Finally, as a means to prioritize the remaining profiles containing a continuous temporal profile for at least 3% of the study period, we excluded all MS feature profiles that did not contain at least 30 total detections and at least 10 consecutive detections among the annual daily samples. The final list contains 1981 filtered MS feature profiles (0.66% of the total number of profiles), including 18 of the 31 target micropollutants that were originally annotated; most of the target micropollutants that were excluded from the final list were removed during the final filtering step because they were only sporadically detected throughout the study period. More details on our overall data reduction approach are provided in the SI.

We then grouped the 1981 filtered MS feature profiles by means of HCA. The resulting dendrogram is presented in Figure 2, which highlights the locations of the 18 target micropollutants. We hypothesized that the highly resolved temporal profiles of the target micropollutants could be used as a means to prioritize nontarget MS features for structural elucidation; we expect that micropollutants represented by closely clustered MS features will have similar sources, use-classes, and fate and transport properties within the watershed.[3,53] For example, nontarget MS features that are clustered closely to atrazine, metolachlor, and simazine (Figure 2) are likely to be micropollutants that are also related to agricultural activities and have similar transport behavior. We prioritized 115 nontarget MS feature profiles (6% of the filtered profiles) that were clustered within or adjacent to localized clusters containing target micropollutants. To the best of our knowledge, this is the first study to prioritize nontarget MS features based on highly resolved temporal profiles and their relationship with respect to the profiles of target micropollutants. Inferring the use-classes or sources of the nontarget micropollutants aids in the identification of the unknown chemical structures by narrowing the breadth of candidate structures.

Structural Elucidation of Nontarget MS Features. Each of the nontarget MS features included in the dendrogram

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

Table 1. List of Detected Target and Nontarget Micropollutants (Level 3 or above) With Associated Molecular Formula or CAS Number (if Available) and Identification Confidence Level

name	molecular formula or CAS No.	level	name	molecular formula or CAS No.	level
desvenlafaxine	93413−62−8	1	**simazine**	122−34−9	1
rac-threo-dihydrobupropion	99102−04−2	1	esculentic acid	464−92−6	2P
gabapentin-lactam	64744−50−9	1	2-methylthio-benzothiazole	615−22−5	1
2-mercapto-benzothiazole	149−30−4	1	NT270 (mercapto-benzothiazole substructure)	$C_{11}H_{11}NO_3S_2$	3
NT240 (benzothiazolone substructure)	$C_{11}H_{13}NO_3S$	3	NT344 (mercapto-benzothiazole substructure)	$C_{15}H_{21}NO_2S_3$	3
valsartan	137862−53−4	1	**metolachlor-ESA**	171118−09−5	1
8-benzyloxy-2,3-dihydro-1,4-benzodioxin-5-carboxylic acid	69114−85−8	2P	propazine-2-hydroxy	7374−53−0	1L
4-butoxy-*N*,*N*-bis (2-ethoxyethyl) benzenesulfonamide	$C_{18}H_{31}NO_5S$	2P	**metalaxyl**	57837−19−1	1
fexofenadine	83799−24−0	1	metolachlor-OXA	152019−73−3	1
lamotrigine	84057−84−1	1	nootkatone	4674−50−4	2L
NT264 (guanidine substructure)	$C_{16}H_{29}N_3$	3	NT212 (hydroxy-s-triazine substructure)	$C_9H_{17}N_5O$	3
losartan	114798−26−4	1	2-amino-benzothiazole	136−95−8	1
lidocaine	137−58−6	1	**atenolol acid**	56392−14−4	1
lauric diethanolamide	120−40−1	2P	**ritalinic acid**	19395−41−6	1
methyl diethyl-dithiocarbamate	686−07−7	1	**hydroxy-atrazine**	2163−68−0	1
NT180 (thiocarbamate substructure)	$C_6H_{13}NOS_2$	3	2,2′-dithiobis-benzothiazole	120−78−5	1
NT148 (thiocarbamate substructure)	$C_6H_{13}NOS$	3	1,2-dihydro-2,2,4-trimethylquinoline	147−47−7	1
(4S)-4-Amino-5-(dibutylamino)-5-oxopentanoic acid	$C_{13}H_{26}N_2O_3$	2P	triethyl phosphate	78−40−0	1
venlafaxine	93413−69−5	1	*targets removed during profile filtering*		
dimethyl phthalate	131−11−3	1	**1-methyl-benzotriazole**	136−85−6	1
atrazine-desethyl	6190−65−4	1	**benzotriazole**	95−14−7	1
triphenyl phosphate	115−86−6	1	**bupropion**	34841−39−9	1
NT343 (#-hydroxyphenyl diphenyl phosphate)	$C_{18}H_{15}O_5P$	3	**caffeine**	58−08−2	1
NT222 (#-anilinoquinazoline)	$C_{14}H_{11}N_3$	3	**diethyl phthalate**	84−66−2	1
metoprolol	37350−58−6	1	**fluconazole**	86386−73−4	1
gabapentin	60142−96−3	1	**irbesartan**	138402−11−6	1
metolachlor	51218−45−2	1	**methocarbamol**	532−03−6	1
alachlor-OXA	171262−17−2	1	**prometon**	1610−18−0	1
metolachlor-2-hydroxy	131068−72−9	2P	**propazine**	139−40−2	1
atrazine	1912−24−9	1	**sitagliptin**	486460−32−6	1
nuciferine *N*-oxide	104385−30−0	2P	**trimethoprim**	738−70−5	1
			warfarin	2610−86−8	1

[a]Organized to present micropollutants in the order in which they were clustered (from left to right) in the dendrogram provided in Figure 2; Figures S32−S90 in the SI describe the temporal trends and the MS information for each nontarget micropollutant; bolded names represent target micropollutants.

are described by their average m/z and RT among all of the samples and are considered to be exact masses of interest (level 5). To facilitate the structural elucidation of each of the prioritized nontarget MS feature profiles, we developed and applied an automated pipeline to assign chemical structures using a series of self-written R scripts and publically available R packages. Similar approaches have been described for structural elucidation of nontarget micropollutants in other environments,[15,18] though our pipeline offers high-throughput data processing and structural elucidation, is fully transparent and customizable, and has been made publically available for free use within the R environment. Our approach allowed us to identify 59 nontarget micropollutants with an increased level of confidence (in addition to the 31 target micropollutants previously identified), including 14 confirmed nontarget micropollutant structures (level 1 or 1L), 8 probable structures (level 2L or 2P), 9 tentative candidates (level 3), and 28 unequivocal molecular formulas (level 4). The nontarget micropollutants identified with a confidence of level 3 or higher are listed in Table 1, along with the target micropollutants with which they clustered in the dendrogram.

An example of how the pipeline is used to elucidate the structure of a nontarget MS feature (NT242, m/z_{avg} = 242.1306, RT_{avg} = 12.2 min) is provided in Figure 3. First, a nontarget MS feature profile is selected based on its proximity to a target micropollutant in the dendrogram and its temporal profile is displayed (Figure 3A); NT242 clustered with desvenlafaxine suggesting it may be a STP-derived micropollutant. Then, the sample with the highest ILIS normalized intensity (October 5) is selected and the associated MS and MS^2 data are automatically loaded into R from a local .mzXML file to display the extracted ion chromatogram (Figure 3B). The RT of the nontarget MS feature is used to extract the associated MS^1 scan to display the isotopic signature (Figure 3C) and the nearest MS^2 scan to display the fragmentation pattern (Figure 3D). The average m/z of the nontarget MS

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

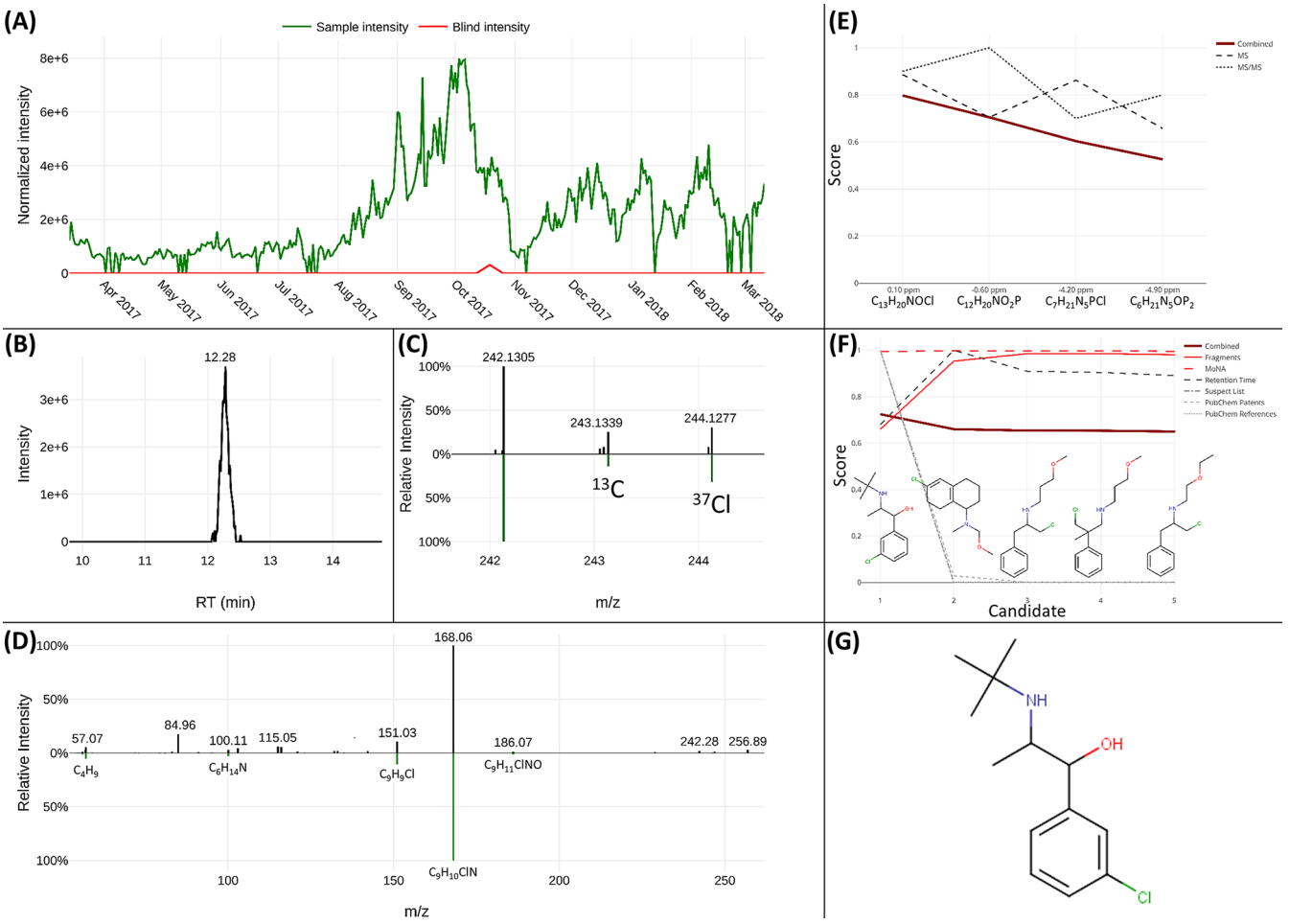


Figure 3. Identification of NT242 following the structure elucidation workflow: (A) temporal trend profile; (B) extracted ion chromatogram; (C) measured MS spectra (top) and theoretical MS spectra (bottom); (D) measured MS^2 fragmentation pattern (top) and in silico MS^2 annotated fragments (bottom); (E) molecular formula prediction; (F) candidate structure ranking; and (G) confirmed structure. NT242 was confirmed as rac threo-dihydrobupropion (level 1) using an authentic reference standard which matched the RT_{avg} of 12.2 min, the MS spectra (m/z_{avg} = 242.1306 for $[M + H]^+$, Δm = 0.13 ppm), the theoretical abundance (13%) of the ^{13}C monoisotopic mass, the theoretical abundance (32%) of the ^{37}Cl monoisotopic mass, and the MS^2 fragments (m/z = 57.07, 151.03, and 168.06).

feature is then automatically used to predict the molecular formula within a 5 ppm mass deviation (Δm), and the MS^1 and MS^2 data are used to score the prediction.[44] The top four scored molecular formulas for NT242 are provided in Figure 3E; the top scored molecular formula is $C_{13}H_{20}NOCl$ (Δm = 0.10 ppm).

Next, the pipeline implements MetFrag and uses the top scored molecular formula to obtain candidate structures from the PubChem online database, which are further filtered to remove nonconnected compounds (e.g., salts) and lower-order isotopes.[21] The final list of candidate structures is then subjected to in silico fragmentation[21] and ranked based on a weighted score that we optimized to maximize the scoring of the 31 target micropollutants previously identified. The fragment score (0.30) is determined by how well the measured MS^2 fragments are explained by the in silico MS^2 fragments. The Metfusion score (0.30) is determined by how well the measured MS^2 fragments match with fragments of similar structures in the MoNA spectral library.[46] The PubChem score is determined by the total number of PubChem references (0.05) and patents (0.05) for a candidate structure, which is a metric that can significantly improve the accuracy of nontarget structure elucidation.[21,54−56] The RT score (0.15) is determined by how well the expected RT of a candidate structure aligns with the RT of the nontarget MS feature. We used the data from our 162 target micropollutants to develop a linear relationship between the octanol−water partition coefficient (logP) and measured RT. The expected logP of the nontarget MS feature is predicted and compared to the estimated logP of the candidate structure for scoring.[21,57] The suspect score (0.15) is determined based on whether the candidate structure is included in SusDat, the merged NORMAN suspect list;[47] we reasoned that candidate structures that have been previously detected or suspected as a water contaminants should receive a higher score. The top five candidate structures for NT242 and their scoring are provided in Figure 3F. More details on our structural elucidation pipeline and the weighting factors are provided in the SI.

The most plausible candidate structure(s) based on scoring and other chemical information (name, use-class, clustering with target micropollutant) is selected from the ranked MetFrag candidate list. The top scored candidate structure for NT242 is rac threo-dihydrobupropion (Figure 3G), a transformation product of the pharmaceutical bupropion. This is a plausible identification because the temporal profile of

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

NT242 clustered in close proximity to the pharmaceutical desvenlafaxine and bupropion was one of the 31 target micropollutants identified in this study. The acquired MS data for plausible candidate structures can be compared with online and in-house mass spectral libraries, though no MS data for rac threo-dihydrobupropion was found in any library. Therefore, an authentic reference standard was acquired for rac threo-dihydrobupropion and the data acquired for the authentic standard matched the data acquired for NT242, confirming the level 1 identification. To the best of our knowledge, rac threo-dihydrobupropion has not been previously reported as a micropollutant. It is interesting to note that we measured a continuous temporal profile of rac threo-dihydrobupropion in Fall Creek, but only an intermittent profile for bupropion; if one assumes that the two micropollutants have the same sources, this suggests that there may be differential fate and transformation behavior of the two micropollutants. Temporal profiles and the analytical data supporting the identification of the 59 nontarget micropollutants are provided in Figures S32–S90 of the SI.

Other Nontarget Micropollutants. Our workflow resulted in the identification of 59 nontarget micropollutants in Fall Creek, with varying levels of confidence. Several notable classes of nontarget micropollutants were identified. First, six of the nontarget micropollutants could be classified as rubber vulcanization accelerators, which are rarely reported as environmental micropollutants. For example, 2-mercaptobenzothiazole (MBT, level 1) has been reported in industrial wastewaters and stormwater, but to the best of our knowledge, has not been reported in surface water systems as it is easily transformed in the environment.[58−60] We also identified putative transformation products of MBT including 2,2′-dithiobisbenzothiazole (DTBT, level 1) and 2-methylthiobenzothiazole (MTBT, level 1), along with other compounds with MBT substructures (NT270 and NT344, level 3) and a benzothiazoline substructure (NT240, level 3). MTBT has been reported in wastewater,[61,62] and DTBT has rarely been detected in surface water.[63] The source of rubber vulcanization accelerators in the Fall Creek watershed is unknown, though we speculate that vehicular tire wear and subsequent transport in road runoff may be the source.[64,65] These findings are particularly relevant as a recent study has suggested that micropollutants derived from tire wear may be linked to high toxicity events in small streams.[53]

Next, five micropollutants that could be classified as plasticizers or flame retardants were identified in Fall Creek including diethyl phthalate (level 1, target micropollutant), dimethyl phthalate (level 1), triethyl phosphate (level 1), triphenyl phosphate (level 1), and hydroxyphenyl diphenyl phosphate (level 3). Phthalate esters, including diethyl phthalate and dimethyl phthalate, have been previously reported as environmental micropollutants and are associated with multiple sources including urban runoff and wastewater effluents.[66] Likewise, organophosphorous plasticizers and flame retardants such as triethyl phosphate and triphenyl phosphate have also been reported as environmental micropollutants in surface waters and are mainly associated with wastewater sources.[67,68] To the best of our knowledge, hydroxyphenyl diphenyl phosphate is reported here for the first time and is likely a transformation product or manufacturing impurity of triphenyl phosphate, although its source cannot be determined from our data.

A number of nontarget pesticides and pesticide transformation products were also identified with temporal profiles that fall into three distinct clusters. First, alachlor-OXA (level 1) and metolachlor-2-hydroxy (level 2P) were identified based on their clustering near atrazine (level 1, target micropollutant), metolachlor (level 1, target micropollutant), and simazine (level 1, target micropollutant). This cluster of pesticides exhibits a strong positive association with streamflow during the agricultural season (approximately June through August) and has low to no abundance throughout the remainder of the study period. Second, another cluster of pesticide transformation products were more ubiquitously present throughout the study period and exhibited little association with streamflow. These include atrazine-hydroxy (level 1, target micropollutant), metolachlor-ESA (level 1, target micropollutant), and propazine-2-hydroxy (level 1L). We suspect that the source of these transformation products is from groundwater recharge.[61] Finally, a third cluster of pesticides and pesticide transformation products exhibited a positive association with streamflow throughout the study period, suggesting persistence and continued mobilization throughout the year. These include metalaxyl (level 1, target micropollutant), metolachlor-OXA (level 1), and a hydroxy-*s*-triazine-containing compound (NT212, level 3). These data allow us to identify three distinct clusters of agriculture-derived micropollutants that provide insights on their relative persistence and fate throughout an annual period. Our results indicate that different pesticides and pesticide transformation products have varying concentration-discharge dependencies throughout the year, which is an important finding for exposure assessment and source control.

Finally, our sampling strategy resulted in samples being stored for varying amounts of time (between 0 and 6 days) inside the automatic sampler. Whereas most micropollutants were stable during this storage, some exhibited evidence of degradation during storage. This manifests as a periodic sawtooth pattern in the temporal profile, as shown in Figure S47 of the SI. For example, several micropollutants with thiocarbamate substructures including methyl diethyldithiocarbamate (MeDDC, level 1), NT180 (level 3), and NT148 (level 3) exhibited this behavior. MeDDC is a human metabolite of disulfiram, a pharmaceutical used to treat alcoholism; disulfiram was not detected in a retrospective screening of the high-resolution mass spectral acquisition. To the best of our knowledge, MeDDC has not been reported as a micropollutant. Thio- and dithiocarbamates are found in many fungicides and their instability during storage suggests limited persistence in Fall Creek.

Insights and Environmental Implications. A primary goal of FCMS is to enable a continuous and comprehensive characterization of organic micropollutant occurrence in a drinking water source. The data presented here represent results from the first year of samples collected from the FCMS and describe our efforts to establish high-throughput sample preparation, sample analysis, and postacquisition data processing workflows to meet this goal. Whereas we used temporal profiles to prioritize nontarget MS features for structural elucidation, our data processing workflow and pipeline can be used to elucidate chemical structures of any nontarget MS feature in high-resolution mass spectral acquisitions from any type of sample including dust, soil, sediment, blood, serum, or wastewater. Our data reveal new insights on the dynamics of micropollutant occurrence in a small stream. For example,

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

important and distinct concentration-discharge relationships were noted for both STP-derived and agriculture-derived micropollutants, though these general relationships alone are insufficient to explain the temporal dynamics of specific micropollutants. Concentration−discharge relationships are expected to be masked or equalized in larger surface water systems, but small streams are clearly vulnerable to hydrological events within the watershed and further research is warranted to study associations among temporal micropollutant profiles and various watershed features. This work is a first step toward improving our ability to characterize the dynamics of exposure risk in small streams and to predict peak events while simultaneously considering multiple contaminants.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.8b05320.

Details on the study site, analytical method, data analysis workflow, and identification of nontarget micropollutants (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +1 607 255 5146; fax: +1 607 255 9004; e-mail: damian.helbling@cornell.edu.

ORCID

Corey M. G. Carpenter: 0000-0002-9168-932X

Damian E. Helbling: 0000-0003-2588-145X

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Chris Dennis Environment Foundation. The authors thank Chris Bordlemay and everyone at the Cornell University Water Filtration Plant for allowing us to deploy and maintain the Fall Creek Monitoring Station inside the facility. The authors thank Dr. Martin Loos for his technical support with enviMass.

■ REFERENCES

(1) Schwarzenbach, R. P.; Escher, B. I.; Fenner, K.; Hofstetter, T. B.; Johnson, C. A.; Von Gunten, U.; Wehrli, B. The Challenge of Micropollutants in Aquatic Systems. *Science* **2006**, *313* (5790), 1072−1077.

(2) Kolpin, D. W.; Furlong, E. T.; Meyer, M. T.; Thurman, E. M.; Zaugg, S. D.; Barber, L. B.; Buxton, H. T. Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams, 1999−2000: A National Reconnaissance. *Environ. Sci. Technol.* **2002**, *36* (6), 1202−1211.

(3) Carpenter, C.; Helbling, D. E. Widespread Micropollutant Monitoring in the Hudson River Estuary Reveals Spatiotemporal Micropollutant Clusters and Their Sources. *Environ. Sci. Technol.* **2018**, *52* (11), 6187−6196.

(4) Bradley, P. M.; Journey, C. A.; Romanok, K. M.; Barber, L. B.; Buxton, H. T.; Foreman, W. T.; Furlong, E. T.; Glassmeyer, S. T.; Hladik, M. L.; Iwanowicz, L. R.; Jones, D. K.; Kolpin, D. W.; Kuivila, K. M.; Loftin, K. A.; Mills, M. A.; Meyer, M. T.; Orlando, J. L.; Reilly, T. J.; Smalling, K. L.; Villeneuve, D. L. Expanded Target-Chemical Analysis Reveals Extensive Mixed-Organic-Contaminant Exposure in U.S. Streams. *Environ. Sci. Technol.* **2017**, *51* (9), 4792−4802.

(5) Moschet, C.; Wittmer, I.; Simovic, J.; Junghans, M.; Piazzoli, A.; Singer, H.; Stamm, C.; Leu, C.; Hollender, J. How a Complete Pesticide Screening Changes the Assessment of Surface Water Quality. *Environ. Sci. Technol.* **2014**, *48* (10), 5423−5432.

(6) Szöcs, E.; Brinke, M.; Karaoglan, B.; Schäfer, R. B. Large Scale Risks from Agricultural Pesticides in Small Streams. *Environ. Sci. Technol.* **2017**, *51* (13), 7378−7385.

(7) Spycher, S.; Mangold, S.; Doppler, T.; Junghans, M.; Wittmer, I.; Stamm, C.; Singer, H. Pesticide Risks in Small Streams - How to Get as Close as Possible to the Stress Imposed on Aquatic Organisms. *Environ. Sci. Technol.* **2018**, *52* (8), 4526−4535.

(8) Lorenz, S.; Rasmussen, J. J.; Süß, A.; Kalettka, T.; Golla, B.; Horney, P.; Stähler, M.; Hommel, B.; Schäfer, R. B. Specifics and Challenges of Assessing Exposure and Effects of Pesticides in Small Water Bodies. *Hydrobiologia* **2017**, *793* (1), 213−224.

(9) U.S. EPA. *Geographic Information Systems Analysis of the Surface Drinking Water Provided by Intermittent, Ephemeral, and Headwater Streams in the U.S.*; 2009.

(10) Neale, P. A.; Munz, N. A.; Aït-Aïssa, S.; Altenburger, R.; Brion, F.; Busch, W.; Escher, B. I.; Hilscherova, K.; Kienle, C.; Novák, J.; Seiler, T. B.; Shao, Y.; Stamm, C.; Hollender, J. Integrating Chemical Analysis and Bioanalysis to Evaluate the Contribution of Wastewater Effluent on the Micropollutant Burden in Small Streams. *Sci. Total Environ.* **2017**, *576*, 785−795.

(11) Munz, N. A.; Burdon, F. J.; de Zwart, D.; Junghans, M.; Melo, L.; Reyes, M.; Schonenberger, U.; Singer, H. P.; Spycher, B.; Hollender, J.; Stamm, C. Pesticides Drive Risk of Micropollutants in Wastewater-Impacted Streams during Low Flow Conditions. *Water Res.* **2017**, *110*, 366−377.

(12) Chiaia-Hernandez, A. C.; Günthardt, B. F.; Frey, M. P.; Hollender, J. Unravelling Contaminants in the Anthropocene Using Statistical Analysis of Liquid Chromatography−high-Resolution Mass Spectrometry Nontarget Screening Data Recorded in Lake Sediments. *Environ. Sci. Technol.* **2017**, *51* (21), 12547−12556.

(13) Shaul, N. J.; Dodder, N. G.; Aluwihare, L. I.; Mackintosh, S. A.; Maruya, K. A.; Chivers, S. J.; Danil, K.; Weller, D. W.; Hoh, E. Nontargeted Biomonitoring of Halogenated Organic Compounds in Two Ecotypes of Bottlenose Dolphins (Tursiops Truncatus) from the Southern California Bight. *Environ. Sci. Technol.* **2015**, *49* (3), 1328−1338.

(14) Moschet, C.; Anumol, T.; Lew, B. M.; Bennett, D. H.; Young, T. M. Household Dust as a Repository of Chemical Accumulation: New Insights from a Comprehensive High-Resolution Mass Spectrometric Study. *Environ. Sci. Technol.* **2018**, *52* (5), 2878−2887.

(15) Schollée, J. E.; Schymanski, E. L.; Avak, S. E.; Loos, M.; Hollender, J. Prioritizing Unknown Transformation Products from Biologically-Treated Wastewater Using High-Resolution Mass Spectrometry, Multivariate Statistics, and Metabolic Logic. *Anal. Chem.* **2015**, *87* (24), 12121−12129.

(16) Hollender, J.; Schymanski, E. L.; Singer, H.; Ferguson, P. L. Non-Target Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go? *Environ. Sci. Technol.* **2017**, *51* (20), 11505−11512.

(17) Schymanski, E. L.; Singer, H. P.; Slobodnik, J.; Ipolyi, I. M.; Oswald, P.; Krauss, M.; Schulze, T.; Haglund, P.; Letzel, T.; Grosse, S.; Thomaidis, N. S.; Bletsou, A.; Zwiener, C.; Ibáñez, M.; Portolés, T.; De Boer, R.; Reid, M. J.; Onghena, M.; Kunkel, U.; Schulz, W.; Guillon, A.; Noyon, N.; Leroy, G.; Bados, P.; Bogialli, S.; Stipaničev, D.; Rostkowski, P.; Hollender, J. Non-Target Screening with High-Resolution Mass Spectrometry: Critical Review Using a Collaborative Trial on Water Analysis. *Anal. Bioanal. Chem.* **2015**, *407* (21), 6237−6255.

(18) Schymanski, E. L.; Singer, H. P.; Longrée, P.; Loos, M.; Ruff, M.; Stravs, M. A.; Ripollés Vidal, C.; Hollender, J. Strategies to Characterize Polar Organic Contamination in Wastewater: Exploring the Capability of High Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2014**, *48* (3), 1811−1818.

(19) Chiaia-Hernandez, A. C.; Schymanski, E. L.; Kumar, P.; Singer, H. P.; Hollender, J. Suspect and Nontarget Screening Approaches to

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

Identify Organic Contaminant Records in Lake Sediments. *Anal. Bioanal. Chem.* **2014**, *406* (28), 7323−7335.
(20) Thurman, E. M.; Ferrer, I.; Blotevogel, J.; Borch, T. Analysis of Hydraulic Fracturing Flowback and Produced Waters Using Accurate Mass: Identification of Ethoxylated Surfactants. *Anal. Chem.* **2014**, *86* (19), 9653−9661.
(21) Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S. MetFrag Relaunched: Incorporating Strategies beyond in Silico Fragmentation. *J. Cheminf.* **2016**, *8* (1), 1−16.
(22) Loos, M. *EnviMass*, Beta Version 3.4. Zenodo, 2016.
(23) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **2006**, *78* (3), 779−787.
(24) Allen, F.; Pon, A.; Wilson, M.; Greiner, R.; Wishart, D. CFM-ID: A Web Server for Annotation, Spectrum Prediction and Metabolite Identification from Tandem Mass Spectra. *Nucleic Acids Res.* **2014**, *42* (W1), W94−W99.
(25) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinf.* **2010**, *11* (1), 395.
(26) HighChem. *mzCloud*.
(27) Stein, S. E. Chemical Substructure Identification by Mass Spectral Library Searching. *J. Am. Soc. Mass Spectrom.* **1995**, *6* (8), 644−655.
(28) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T. MassBank: A Public Repository for Sharing Mass Spectral Data for Life Sciences. *J. Mass Spectrom.* **2010**, *45* (7), 703−714.
(29) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. Metlin: A Metabolite Mass Spectral Database. *Ther. Drug Monit.* **2005**, *27* (6), 747−751.
(30) Ryberg, K. R.; Vecchia, A. V.; Gilliom, R. J.; Martin, J. D. Pesticide Trends in Major Rivers of the United States, 1992−2010. In *Scientific Investigations Report 2014−5135*; U.S. Geological Survey, 2014.
(31) Richards, B. K.; Pacenka, S.; Meyer, M. T.; Dietze, J. E.; Schatz, A. L.; Teuffer, K.; Aristilde, L.; Steenhuis, T. S. Antecedent and Post-Application Rain Events Trigger Glyphosate Transport from Runoff-Prone Soils. *Environ. Sci. Technol. Lett.* **2018**, *5* (5), 249−254.
(32) Ruff, M.; Mueller, M. S.; Loos, M.; Singer, H. P. Quantitative Target and Systematic Non-Target Analysis of Polar Organic Micro-Pollutants along the River Rhine Using High-Resolution Mass-Spectrometry - Identification of Unknown Sources and Compounds. *Water Res.* **2015**, *87*, 145−154.
(33) Schlüsener, M. P.; Kunkel, U.; Ternes, T. A. Quaternary Triphenylphosphonium Compounds: A New Class of Environmental Pollutants. *Environ. Sci. Technol.* **2015**, *49* (24), 14282−14291.
(34) U.S. Geological Survey. National Water Information System data www.waterdata.usgs.gov/nwis.
(35) Homer, C. G.; Dewitz, J. A.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N. D.; Wickham, J. D.; Megown, K. Completion of the 2011 National Land Cover Database for the Conterminous United States-Representing a Decade of Land Cover Change Information. *Photogramm. Eng. Remote Sensing* **2015**, *81* (5), 345−354.
(36) U.S. EPA. EPA Facility Registry Service (FRS): Wastewater Treatment Plants www.catalog.data.gov/dataset/epa-facility-registry-service-frs-wastewater-treatment-plants.
(37) Hollingshead, N. *Residental Wastewater Systems*; Cayuga Lake Watershed Network; 2007.
(38) Carpenter, C. M. G.; Helbling, D. E. Removal of Micropollutants in Biofilters: Hydrodynamic Effects on Biofilm Assembly and Functioning. *Water Res.* **2017**, *120*, 211−221.
(39) Pochodylo, A. L.; Helbling, D. E. Emerging Investigators Series: Prioritization of Suspect Hits in a Sensitive Suspect Screening Workflow for Comprehensive Micropollutant Characterization in Environmental Samples. *Environ. Sci. Water Res. Technol.* **2017**, *3* (1), 54−65.
(40) Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open Source Software for Rapid Proteomics Tools Development. *Bioinformatics* **2008**, *24* (21), 2534−2536.
(41) R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008.
(42) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48* (4), 2097−2098.
(43) Stravs, M. A.; Schymanski, E. L.; Singer, H. P.; Hollender, J. Automatic Recalibration and Processing of Tandem Mass Spectra Using Formula Annotation. *J. Mass Spectrom.* **2013**, *48* (1), 89−99.
(44) Meringer, M.; Reinker, S.; Zhang, J.; Muller, A. MS/MS Data Improves Automated Determination of Molecular Formulas by Mass Spectrometry. *MATCH Commun. Math Comput. Chem.* **2011**, *65* (2), 259−290.
(45) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. PubChem. Substance and Compound Databases. *Nucleic Acids Res.* **2016**, *44* (D1), D1202−D1213.
(46) Gerlich, M.; Neumann, S. MetFusion: Integration of Compound Identification Strategies. *J. Mass Spectrom.* **2013**, *48* (3), 291−298.
(47) Norman network. Norman Suspect List Exchange http://www.norman-network.com/?q=node/236.
(48) Campanha, M. B.; Awan, A. T.; de Sousa, D. N. R.; Grosseli, G. M.; Mozeto, A. A.; Fadini, P. S. A 3-Year Study on Occurrence of Emerging Contaminants in an Urban Stream of São Paulo State of Southeast Brazil. *Environ. Sci. Pollut. Res.* **2015**, *22* (10), 7936−7947.
(49) Osorio, V.; Proia, L.; Ricart, M.; Pérez, S.; Ginebreda, A.; Cortina, J. L.; Sabater, S.; Barceló, D. Hydrological Variation Modulates Pharmaceutical Levels and Biofilm Responses in a Mediterranean River. *Sci. Total Environ.* **2014**, *472* (2014), 1052−1061.
(50) Kolpin, D. W.; Skopec, M.; Meyer, M. T.; Furlong, E. T.; Zaugg, S. D. Urban Contribution of Pharmaceuticals and Other Organic Wastewater Contaminants to Streams during Differing Flow Conditions. *Sci. Total Environ.* **2004**, *328* (1−3), 119−130.
(51) Doppler, T.; Camenzuli, L.; Hirzel, G.; Krauss, M.; Lück, A.; Stamm, C. Spatial Variability of Herbicide Mobilisation and Transport at Catchment Scale: Insights from a Field Experiment. *Hydrol. Earth Syst. Sci.* **2012**, *16* (7), 1947−1967.
(52) Fairbairn, D. J.; Arnold, W. A.; Barber, B. L.; Kaufenberg, E. F.; Koskinen, W. C.; Novak, P. J.; Rice, P. J.; Swackhamer, D. L. Contaminants of Emerging Concern: Mass Balance and Comparison of Wastewater Effluent and Upstream Sources in a Mixed-Use Watershed. *Environ. Sci. Technol.* **2016**, *50* (1), 36−45.
(53) Peter, K. T.; Tian, Z.; Wu, C.; Lin, P.; White, S.; Du, B.; McIntyre, J. K.; Scholz, N. L.; Kolodziej, E. P. Using High-Resolution Mass Spectrometry to Identify Organic Contaminants Linked to Urban Stormwater Mortality Syndrome in Coho Salmon. *Environ. Sci. Technol.* **2018**, *52*, 10317−10327.
(54) McEachran, A. D.; Sobus, J. R.; Williams, A. J. Identifying Known Unknowns Using the US EPA's CompTox Chemistry Dashboard. *Anal. Bioanal. Chem.* **2017**, *409* (7), 1729−1735.
(55) Little, J. L.; Cleven, C.; Brown, S. Identification of "Known Unknowns" Utilizing Accurate Mass Data and Chemical Abstracts Service Databases. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (1), 348−359.

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87

(56) Little, J. L.; Williams, A. J.; Pshenichnov, A.; Tkachenko, V. Identification of "Known Unknowns" Utilizing Accurate Mass Data and Chemspider. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (1), 179−185.
(57) Hu, M.; Müller, E.; Schymanski, E. L.; Ruttkies, C.; Schulze, T.; Brack, W.; Krauss, M. Performance of Combined Fragmentation and Retention Prediction for the Identification of Organic Micropollutants by LC-HRMS. *Anal. Bioanal. Chem.* **2018**, *410* (7), 1931−1941.
(58) De Wever, H.; Besse, P.; Verachtert, H. Microbial Transformations of 2-Substituted Benzothiazoles. *Appl. Microbiol. Biotechnol.* **2001**, *57* (5−6), 620−625.
(59) Liao, C.; Kim, U. J.; Kannan, K. A Review of Environmental Occurrence, Fate, Exposure, and Toxicity of Benzothiazoles. *Environ. Sci. Technol.* **2018**, *52* (9), 5007−5026.
(60) Grebel, J. E.; Mohanty, S. K.; Torkelson, A. A.; Boehm, A. B.; Higgins, C. P.; Maxwell, R. M.; Nelson, K. L.; Sedlak, D. L. Engineered Infiltration Systems for Urban Stormwater Reclamation. *Environ. Eng. Sci.* **2013**, *30* (8), 437−454.
(61) Hug, C.; Ulrich, N.; Schulze, T.; Brack, W.; Krauss, M. Identification of Novel Micropollutants in Wastewater by a Combination of Suspect and Nontarget Screening. *Environ. Pollut.* **2014**, *184*, 25−32.
(62) Blum, K. M.; Andersson, P. L.; Renman, G.; Ahrens, L.; Gros, M.; Wiberg, K.; Haglund, P. Non-Target Screening and Prioritization of Potentially Persistent, Bioaccumulating and Toxic Domestic Wastewater Contaminants and Their Removal in on-Site and Large-Scale Sewage Treatment Plants. *Sci. Total Environ.* **2017**, *575*, 265−275.
(63) Baken, K. A.; Sjerps, R. M. A.; Schriks, M.; van Wezel, A. P. Toxicological Risk Assessment and Prioritization of Drinking Water Relevant Contaminants of Emerging Concern. *Environ. Int.* **2018**, *118* (May), 293−303.
(64) Pal, A.; He, Y.; Jekel, M.; Reinhard, M.; Gin, K. Y. H. Emerging Contaminants of Public Health Significance as Water Quality Indicator Compounds in the Urban Water Cycle. *Environ. Int.* **2014**, *71*, 46−62.
(65) Unice, K. M.; Bare, J. L.; Kreider, M. L.; Panko, J. M. Experimental Methodology for Assessing the Environmental Fate of Organic Chemicals in Polymer Matrices Using Column Leaching Studies and OECD 308 Water/Sediment Systems: Application to Tire and Road Wear Particles. *Sci. Total Environ.* **2015**, *533*, 476−487.
(66) Zeng, F.; Wen, J.; Cui, K.; Wu, L.; Liu, M.; Li, Y.; Lin, Y.; Zhu, F.; Ma, Z.; Zeng, Z. Seasonal Distribution of Phthalate Esters in Surface Water of the Urban Lakes in the Subtropical City, Guangzhou, China. *J. Hazard. Mater.* **2009**, *169* (1−3), 719−725.
(67) Kim, U. J.; Kannan, K. Occurrence and Distribution of Organophosphate Flame Retardants/Plasticizers in Surface Waters, Tap Water, and Rainwater: Implications for Human Exposure. *Environ. Sci. Technol.* **2018**, *52* (10), 5625−5633.
(68) Wei, G. L.; Li, D. Q.; Zhuo, M. N.; Liao, Y. S.; Xie, Z. Y.; Guo, T. L.; Li, J. J.; Zhang, S. Y.; Liang, Z. Q. Organophosphorus Flame Retardants and Plasticizers: Sources, Occurrence, Toxicity and Human Exposure. *Environ. Pollut.* **2015**, *196*, 29−46.

DOI: 10.1021/acs.est.8b05320
Environ. Sci. Technol. 2019, 53, 77−87