



# Identification of LC-HRMS nontarget signals in groundwater after source related prioritization

Karin Kiefer<sup>a,b</sup>, Letian Du<sup>a,b</sup>, Heinz Singer<sup>a</sup>, Juliane Hollender<sup>a,b,\*</sup>

<sup>a</sup> Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland

<sup>b</sup> Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zurich, Switzerland

## ARTICLE INFO

### Article history:

Received 25 November 2020

Revised 26 February 2021

Accepted 28 February 2021

Available online 2 March 2021

### Keywords:

Target screening

Nontarget screening

Micropollutant

Persistent and mobile compounds

PMOC

AcquireX

## ABSTRACT

Groundwater is a major drinking water resource but its quality with regard to organic micropollutants (MPs) is insufficiently assessed. Therefore, we aimed to investigate Swiss groundwater more comprehensively using liquid chromatography high-resolution tandem mass spectrometry (LC-HRMS/MS). First, samples from 60 sites were classified as having high or low urban or agricultural influence based on 498 target compounds associated with either urban or agricultural sources. Second, all LC-HRMS signals were related to their potential origin (urban, urban and agricultural, agricultural, or not classifiable) based on their occurrence and intensity in the classified samples. A considerable fraction of estimated concentrations associated with urban and/or agricultural sources could not be explained by the 139 detected targets. The most intense nontarget signals were automatically annotated with structure proposals using MetFrag and SIRIUS4/CSI:FingerID with a list of >988,000 compounds. Additionally, suspect screening was performed for 1162 compounds with predicted high groundwater mobility from primarily urban sources. Finally, 12 nontargets and 11 suspects were identified unequivocally (Level 1), while 17 further compounds were tentatively identified (Level 2a/3). amongst these were 13 pollutants thus far not reported in groundwater, such as: the industrial chemicals 2,5-dichlorobenzenesulfonic acid (19 detections, up to 100 ng L<sup>-1</sup>), phenylphosphonic acid (10 detections, up to 50 ng L<sup>-1</sup>), triisopropanolamine borate (2 detections, up to 40 ng L<sup>-1</sup>), O-des[2-aminoethyl]-O-carboxymethyl dehydroamlodipine, a transformation product (TP) of the blood pressure regulator amlodipine (17 detections), and the TP SYN542490 of the herbicide metolachlor (Level 3, 33 detections, estimated concentrations up to 100–500 ng L<sup>-1</sup>). One monitoring site was far more contaminated than other sites based on estimated total concentrations of potential MPs, which was supported by the elucidation of site-specific nontarget signals such as the carcinogen chlorendic acid, and various naphthalenedisulfonic acids. Many compounds remained unknown, but overall, source related prioritisation proved an effective approach to support identification of compounds in groundwater.

© 2021 The Author(s). Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## 1. Introduction

Groundwater is a major drinking water resource and therefore its quality is of high interest. Various micropollutants (MPs) from agriculture, households, or industry, entering aquifers via different pathways, affect groundwater quality (Loos et al., 2010). Whereas pesticides are applied in high amounts to agricultural soils (Pimentel 2009) and then enter aquifers mainly via seepage, pharmaceuticals, household and industrial chemicals may contaminate aquifers predominantly via leaky sew-

ers (Wolf et al., 2012) or via bank filtration from wastewater impacted surface waters (Heberer et al., 2004). Consequently, the land use in the catchment and/or hydrogeological setting (e.g. bank filtration, characteristics of top layers) influence the MP pattern at the groundwater monitoring site (Stuart et al., 2014; Ter Laak et al., 2012). Given that (i) >350,000 chemicals, and mixtures thereof, are registered in national and regional inventories (Wang et al., 2020), (ii) compounds undergo transformations in the environment and engineered systems (Kolpin et al., 2009), and (iii) “new” MPs are regularly reported to be in the water cycle, e.g. Weber et al. (2007), Schmidt and Brauch (2008), Reemtsma et al. (2013), Schlüsener et al. (2015), Zahn et al. (2016), Gago-Ferrero et al. (2018), Schulze et al. (2019),

\* Corresponding author.

E-mail address: [juliane.hollender@eawag.ch](mailto:juliane.hollender@eawag.ch) (J. Hollender).

or Zahn et al. (2019), it seems insufficient to evaluate water quality only on the basis of selected analytes.

Liquid chromatography high-resolution tandem mass spectrometry (LC-HRMS/MS) enables the detection of possibly thousands of small molecules in a water sample, both from natural and anthropogenic sources. Some of these compounds can be elucidated by broad-scope target screening where reference standards are available in the laboratory, while others are detected as part of suspect screening, i.e. searching for specific structures without having reference material at hand (Krauss et al., 2010). However, usually most LC-HRMS/MS signals remain unknown (so-called nontargets), and should therefore be further investigated. Spectral libraries such as the European MassBank (Horai et al., 2010), MassBank of North America (MoNA; <http://mona.fiehnlab.ucdavis.edu/>), and mz-Cloud ([www.mzcloud.org](http://www.mzcloud.org)) allow structural annotation of nontarget signals with high confidence level (Schymanski et al., 2014b), though their use is limited to compounds for which MS/MS spectra were made accessible by other laboratories. For compounds not present in spectral libraries, in silico fragmentation tools, e.g. MetFrag (Ruttkies et al., 2016), SIRIUS4/CSI:FingerID (Dührkop et al., 2015; Dührkop et al., 2019), and CFM-ID (Allen et al., 2014), in combination with large chemical compound databases such as PubChem (Kim et al., 2019), CompTox Chemistry Dashboard (Williams et al., 2017), or the NORMAN Suspect List Exchange ([www.norman-network.com/nds/SLE](http://www.norman-network.com/nds/SLE)) are promising alternatives for structural annotation. These tools enable the high-throughput annotation of thousands of nontargets, however, with often hundreds or even thousands of candidates. Thus, the final identification with reference material remains a major bottleneck so that so far only few nontargets have been elucidated with high confidence in water samples (Albergamo et al., 2019; Gago-Ferrero et al., 2015; Schymanski et al., 2014a; Tian et al., 2020).

As a consequence, nontargets of interest need to be prioritized before investing time and resources into structure elucidation or even purchasing reference material, using one or a combination of several approaches depending on the study context. Most studies apply filters for analytical quality control, e.g. excluding compounds detected in procedural blanks and including only compounds with high reproducibility in replicate and pooled samples (Broadhurst et al., 2018; Sangster et al., 2006). The use of pooled samples (i.e. samples comprising aliquots of all samples) is more common in metabolomics than in environmental studies. Intensity is a crucial parameter in most studies (Hug et al., 2014; Schymanski et al., 2014a), as intense signals may provide higher-quality MS/MS data supporting elucidation. Additionally, intensity correlates to some extent with concentration, though it is important to highlight that high intensity or concentration does not necessarily mean high toxicological risk. Furthermore, Cl and Br isotope patterns may point to compounds from anthropogenic origins (Chiaia-Hernandez et al., 2014; Hug et al., 2014) and homologues to larger groups of related compounds such as surfactants (Schymanski et al., 2014a). Comparing related samples such as raw water and final drinking water (Müller et al., 2011) or road dust and surface water (Seiwert et al., 2020) facilitate prioritization of persistent signals originating from a specific source. In this regard, multivariate statistical tools such as principal component analysis and hierarchical clustering can be applied, for example to samples collected along a wastewater treatment plant (Schollée et al., 2015; Schollée et al., 2018) or a riverbank transect (Albergamo et al., 2019), to determine nontargets with specific trends.

The goal of this study was to comprehensively evaluate the quality of water derived from 60 groundwater monitoring sites in Switzerland. To do so, all detected LC-HRMS signals were to be related to their potential sources (urban, agricultural), followed by structural elucidation of the most prominent ones, focusing in particular on potential MPs from urban sources. Our hypothesis was

that by combining an appropriate nontarget prioritization strategy with a highly automated structure elucidation workflow, we would find previously unreported compounds in groundwater. First, we performed an extensive target screening for 498 MPs to classify the samples according to their dominating urban and agricultural pollutants. Next, all LC-HRMS signals were classified based on their pre-dominating occurrence in urban or agricultural sources, which was assessed by comparing samples with high vs. low urban or high vs. low agricultural influence. Structure elucidation focused on nontargets from potentially urban sources because we investigated already in detail major agricultural MPs, i.e. pesticides and their transformation products (TPs), in a similar sample set (Kiefer et al., 2019). Accordingly, most intense nontargets from potentially urban sources and, in addition widespread nontargets, were annotated with candidate proposals using MetFrag and SIRIUS4/CSI:FingerID and finally elucidated with reference standards, if commercially available. To not overlook important urban MPs with lower signal intensity, we applied additionally a suspect screening for 1162 polar, and therefore mobile, compounds from mainly urban sources.

## 2. Methods

### 2.1. Groundwater samples

The applied nontarget screening approach aimed to classify nontargets with regards to their origin (urban, agricultural), based on their occurrence at monitoring sites that are i) urban impacted, ii) agriculturally impacted, or iii) show only low anthropogenic impact. Therefore, 60 monitoring sites (44 abstraction wells, 16 springs) were selected out of >500 sites from the Swiss National Groundwater Monitoring NAQUA ([www.bafu.admin.ch/naqua](http://www.bafu.admin.ch/naqua)) based on long-term monitoring data collected within NAQUA. Twenty sites contained MPs from pre-dominantly urban sources (pharmaceuticals, sweeteners), while 20 other sites showed high frequency or concentrations of MPs from agricultural sources (pesticides and their TPs). A further 20 sites exhibited comparably low anthropogenic contamination. Several sites were not clearly classifiable as urban- or agriculturally-impacted due to the occurrence of pollutants from both source types.

The 60 groundwater samples were collected in laboratory glass bottles (previously annealed at 500 °C; 1 L bottles, SIMAX Kavalier, Czech Republic) in May and August 2018 within the routine sampling of NAQUA. Samples were stored cooled for up to four weeks and then frozen at -20 °C until measurement. For quality control, pooled samples were prepared, i.e. samples consisting of an equal amount of each sample. For this, aliquots (5 or 10 mL) of each groundwater sample were transferred to laboratory glass bottles (0.5 L or 1 L) using a 5 mL brown glass vial before samples were frozen. To determine contaminations from sampling, sample storage or handling, ultrapure water (>18 MΩcm, Barnstead Nanopure Diamond system) was filled into laboratory glass bottles, transferred during sampling into another bottle (at five monitoring sites), and then stored and processed analogously to the groundwater samples as blank samples.

### 2.2. Sample preparation via vacuum-assisted evaporation

To avoid losses of polar compounds during sample enrichment, samples were concentrated via vacuum-assisted evaporation, analogously to Mechelke et al. (2019) though with slight modifications. Samples were transferred to BÜCHI glass vials (1 mL appendix, previously annealed at 450 °C) at a volume of 120 mL and spiked with 35 isotope-labelled internal standards at 100 ng L<sup>-1</sup>. Samples were evaporated on a Syncore® Analyst (BÜCHI, Switzerland) for 7–8 h to approximately 1 mL at 20 mbar and 45 °C (appendix

cooled at 7–10 °C) using the back-flush unit. The concentrate volume was adjusted to 1.6 mL by adding ultrapure water using annealed glass Pasteur pipettes (747715, Brand GmbH, Germany). The BÜCHI vials were rinsed thoroughly with the concentrate to reduce analyte losses due to sorption. Finally, the concentrate was centrifuged at 3720 g (Heraeus Megafuge 1.0 R, Thermo Fisher Scientific, U.S.) in annealed vials and transferred to 1.5 mL vials (previously annealed at 500 °C; vials: 080400-XL; screw caps: 090301; BGB Analytik, Switzerland). Analogously to the samples, four calibration standards (1, 10, 100, 1000 ng L<sup>-1</sup> in ultrapure water), four spiked samples (1, 10, 100, 250 ng L<sup>-1</sup>), six pooled samples (i.e. replicates) and seven field and laboratory blank samples (ultrapure water spiked with internal standards) were prepared for quality control and target quantification. For details on spike solutions, see SI-A1.

### 2.3. LC-HRMS/MS analysis

Samples were measured in triplicate in a randomized order. After nine sample injections (three triplicates), a blank was injected, followed by a pooled sample for quality control. The concentrated samples (140 µL, i.e. 10.5 mL of the original sample) were injected to a LC system consisting of a PAL RTC autosampler (CTC Analytics, Switzerland), a reversed phase C18 column (Atlantis T3, 3 µm, 3 × 150 mm; Waters, Ireland), and a Dionex UltiMate 3000 RS pump (Thermo Fisher Scientific RS). The gradient elution started with 100% water (containing 0.1% formic acid) to achieve an optimal retention of polar compounds. Then, methanol (containing 0.1% formic acid) was added and increased to 95% from 1.5 to 18.5 min, and finally kept constant for 10 min. The flow rate was 0.3 mL min<sup>-1</sup>. For details, see Table SI-A1.

Analytes were ionized in electrospray (3.5/–2.5 kV) and detected on an Orbitrap mass spectrometer (Fusion Lumos, Thermo Fisher Scientific, U.S.) with a resolution R of 240,000 (at *m/z* 200, full width at half maximum (FWHM)) in MS1 full-scan mode (*m/z* 100–1000), followed by three to four data-dependant MS/MS full-scans (high-resolution product scans; R 30,000 FWHM at *m/z* 200; cycle time 1 s; isolation window of precursor 1 *m/z*). Internal calibration (EASY-IC™) ensured a mass accuracy of <±2 ppm in MS1 scans for 99.8% of detected target compound peaks and internal standard peaks (<±1 ppm for 98.4% of peaks). AcquireX software (Deep Scan; Thermo Fisher Scientific, U.S.) was used to increase MS/MS coverage. In the first triplicate injection, data-dependant MS/MS full-scans were triggered based on a mass list containing target compounds. Then, AcquireX performed peak picking and added detected features (*m/z* and retention time) to the mass list so that in the second triplicate injection, data-dependant MS/MS full-scans were triggered based on the mass list modified by AcquireX. Before the third injection, AcquireX shifted features, for which MS/MS scans were already acquired during the second injection, from the mass to the exclusion list so that these features were not triggered again. If no features from the mass list were detected, the MS/MS scans for the most intense signals were acquired. Triggered features were excluded dynamically for 3 s. For details, see SI-A2.

### 2.4. Target screening

For each target compound, extracted ion chromatograms were plotted with the MSnbase R package (Gatto and Lilley, 2012) and visually inspected. If the target compound was detected in groundwater samples, the concentration was determined using Trace Finder 4.1 (Thermo Fisher Scientific, U.S.) based on the peak area ratio of the target compound to that of corresponding internal standard. If no structurally identical isotope-labelled internal standard was available, an internal standard was selected with similar

retention time as the analyte and resulting in a relative recovery close to 100% in the spiked samples using an in-house R script (Schollée, 2018). For details on target quantification including determination of limit of quantification (LOQ), see SI-A3.

### 2.5. Suspect and nontarget screening

Measurement files were converted to mzXML format using MSConvert 3.0, ProteoWizard (Chambers et al., 2012) and then processed in the enviMass workflow (envibee GmbH, Switzerland). Data post-processing was conducted in the R environment, version 3.6.3 (R Core Team, 2020).

#### 2.5.1. Data pre-processing

Data pre-processing was performed using the enviMass workflow (version 4.2633) including peak picking, mass recalibration, retention time alignment, intensity normalization based on median intensity of internal standards, replicate filtering, and target and suspect annotation. Features, i.e. chromatographic peaks defined by their *m/z* and retention time, that likely resulted from the same compound (adducts, isotopologues) were grouped into so called components based on intensity correlation and *m/z* distance. The most intense feature within a component was used for further data analysis. Settings (e.g. *m/z* and retention time tolerances) were optimized until 87% of target peaks (detected with Trace Finder 4.1, Section 2.4) were found. 62% of target peaks that were not detected were <10 ng L<sup>-1</sup> or exceeded the LOQ (Section 2.4) by less than factor 5; other non-detects were mostly related to large retention time shifts (i.e. exceeding the retention time tolerance for targets) or poor peak shape. For final settings, see SI-A4.

Finally, a table containing the peak height intensity pattern of each component across the samples (so-called profiles) was exported. Profiles were prioritized for further inspection by excluding those with a retention time <3 min, average sample/blank intensity ratio <5, and maximum peak intensity <10<sup>6</sup>. For comparison, a peak height intensity of 10<sup>6</sup> corresponded to a concentration of ≥10 ng L<sup>-1</sup> for approximately 90% of targets ionizing in positive ionization mode and 35% of targets ionizing in negative ionization mode.

It should be noted that for some compounds several components might exist (e.g. positive and negative ionization, in-source fragments). Moreover, for some target compounds, two or three profiles were observed, often related to shifting retention times or peak picking artefacts. Therefore, profiles differing in *m/z* by <2 ppm and retention time <30 s were grouped using an in-house R function (Schollée et al., in preparation). Peak intensities were averaged across replicates.

#### 2.5.2. Prioritization of profiles using sample classification

To classify nontarget compounds according to their potential origin (urban, agricultural), groundwater samples were first classified based on the sum concentration of 269 targets from predominantly urban origin (pharmaceuticals and their TPs, sweeteners, industrial chemicals, biocides, illicit drugs, personal care products, and others, SI-B1) and 229 targets from predominantly agricultural origin (pesticides and their TPs, SI-B1). Samples were defined as having high urban influence if the sum concentration of urban targets was >100 ng L<sup>-1</sup>; otherwise, they were defined as samples with low urban influence. Likewise, samples for which the sum concentration of agricultural targets exceeded 100 ng L<sup>-1</sup> were defined as having high agricultural influence, and otherwise as having low agricultural influence. The cut-off of 100 ng L<sup>-1</sup> was guided by the European Union's drinking water standard for single pesticides and relevant pesticide TPs (European Commission, 1998). The sum concentration (instead of concentrations of single com-



pounds) was used so that the classification depended not only on single targets with high concentrations.

Next, the ratio of a compound's average intensity in samples with high urban/agricultural influence to that of its average intensity in samples with low urban/agricultural influence was calculated as a measure for the likelihood that the compound originated from urban sources or agricultural sources, respectively:

$$\text{Measure for urban origin} = \frac{\text{average intensity in samples with high urban influence}}{\text{average intensity in samples with low urban influence}} \quad (1)$$

$$\text{Measure for agricultural origin} = \frac{\text{average intensity in samples with high agricultural influence}}{\text{average intensity in samples with low agricultural influence}} \quad (2)$$

If the compound was not detected in one of the sample groups, it was assumed that the compound was detected in one sample with the minimum intensity observed in the whole dataset to avoid dividing by 0.

Profiles were classified as follows:

- potential urban MP: measure for urban origin  $>5$  and measure for agricultural origin  $\leq 5$ ,
- potential agricultural MP: measure for urban origin  $\leq 5$  and measure for agricultural origin  $>5$ ,
- potential urban and agricultural MP: measure for urban origin  $>5$  and measure for agricultural origin  $>5$ ,
- not classifiable: remaining profiles.

Here, the decision to use a factor of five as threshold was guided by profiles annotated as target compounds, i.e.  $>80\%$  of profiles of urban targets and  $>90\%$  of profiles of agricultural targets, respectively, showed on average more than five times higher intensities in samples with high urban/agricultural influence vs. samples with low urban/agricultural influence. However, it should be pointed out that some target compounds, which would be classified as "potential urban and agricultural MP", likely originate only from urban sources (e.g. x-ray contrast agent diatrizoic acid).

To gather more information on potential sources, the most intense profiles in each group (urban, urban/agricultural, agricultural profiles with maximum intensity  $>5 \times 10^6$ ; not classifiable profiles with maximum intensity  $>10^7$  and  $\geq 30$  detections) were retrospectively screened for in the effluent of two Swiss municipal wastewater treatment plants (24 h composite samples, dry-weather conditions). For this, extracted ion chromatograms (EICs) were generated for blank and effluent samples, measured with a comparable method (without enrichment), using the R package MSnbase (Gatto and Lilley, 2012). The EICs were then checked for peaks with intensity  $>10^5$  and deviating  $<1$  min from the average retention time of the corresponding profiles (SI-C2).

Total concentrations in each sample were estimated assuming that the compounds ionize either less efficiently than, as efficiently as, or more efficiently than the 113 target compounds (92 in positive ionization mode, 21 in negative ionization mode) spiked in to groundwater samples (1, 10, 100, 250 ng L<sup>-1</sup>). Here "less efficiently than", "as efficiently as", and "more efficiently than" correspond to the 25th, 50th and 75th percentile of target compound peak intensities. For details, see SI-A5.

### 2.5.3. Elucidation of suspects and nontargets

**Nontargets:** Structural elucidation of nontargets focused on profiles that were classified as potentially of urban or urban and agricultural origin and that had maximum intensity  $>5 \times 10^6$ , and on all profiles that had maximum intensity  $>10^7$  and  $\geq 30$  detections. For each profile, MS1 and MS/MS data were extracted from the sample with highest precursor intensity using the RMassBank package (Stravs et al., 2013). Structural proposals were then assigned using both MetFrag CL 2.4.5 (using functions from the R package ReSOLUTION, Schymanski (2020)) and

SIRIUS4/CSI:FingerID. Both in silico fragmenters were used with a list of  $>988,000$  compounds of potential environmental relevance, including those in CompTox (Schymanski, 2019), PubChem-Lite tier1 (Bolton and Schymanski, 2020), NORMAN SusDat (Norman Network et al., 2020), STOFF-IDENT (Letzel et al., 2017), the original dataset used for UBAPMT (Arp and Hale, 2020), i.e. Ex-

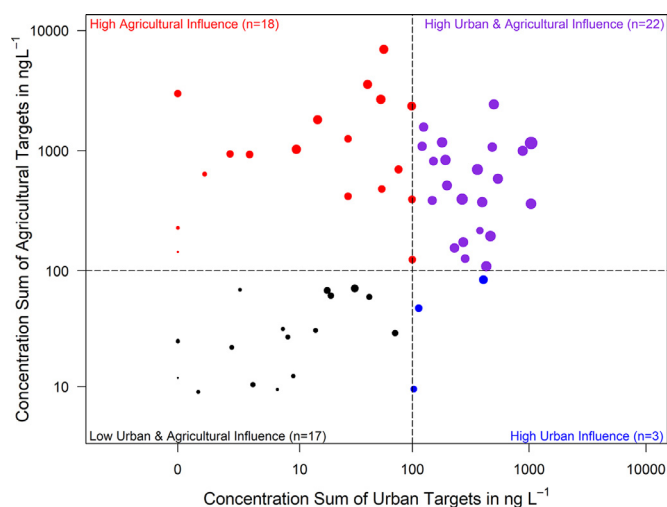
tended PMT (H.-P. Arp and S.E. Hale, personal communication), SwissPest19 (Kiefer et al., 2020b), as well as 71 additional potentially mobile pesticide transformation products (T. Poiger, personal communication). For details, see SI-A6.

Nontargets were prioritized for confirmation based on multiple lines of evidence as in the following: (i) a positive hit in the MS/MS libraries NIST17 (National Institute of Standards and Technology, U.S. Department of Commerce), MoNA (LC-MS/MS spectra obtained from <https://mona.fiehnlab.ucdavis.edu/downloads> in December 2019) or MassBank (obtained from <https://github.com/MassBank/MassBank-data> in December 2019) using the NIST Mass Spectral Search Program (version 2.3) or in mzCloud (selected hits manually checked), (ii) performance of in silico fragmentation, (iii) peak shape, (iv) intensity and detection frequency, (v) plausibility of retention time, and (vi) availability of reference material.

**Suspects:** The suspect list comprised 1162 MS-ready structures, all with heteroatoms and exact masses  $>100$ , and was compiled from Schulze et al. (2019), KEMI Market List (Fischer, 2017), UBAPMT (Arp and Hale, 2020), and Extended PMT (H.-P. Arp and S.E. Hale, personal communication). The Extended PMT dataset was filtered for compounds classified as very mobile (vM) and "highly expected in the environment" (H.-P. Arp, personal communication, e.g. due to high production volumes). The KEMI Market List, containing  $>25,000$  chemicals that are expected on the EU market (e.g. industrial chemicals, pharmaceuticals, pesticides), was restricted to compounds that are more likely to occur in groundwater. Therefore, only compounds were selected that exhibited a high water exposure index ( $>15$ ; water exposure index ranges from 1 to 27) and were classified as (potentially) mobile or very mobile (pot M/vM, M/vM, or vM) based on speciation and logD<sub>OW</sub> (predicted by JChem for Office, version 19.22.0.548; ChemAxon Ltd.), as described by Arp and Hale (2019).

Profiles annotated with the suspect list by enviMass (Section 2.5.1) were prioritized in a similar workflow as nontarget profiles. MetFrag and SIRIUS4/CSI:FingerID were run using both the suspect list and PubChemLite tier1 (Bolton and Schymanski, 2020) as databases to check how well the measured MS/MS spectrum fits to the suspect compared to other candidates (SI-A6). Only suspects ranked amongst the top 3 candidates of SIRIUS4/CSI:FingerID were checked manually, as described for nontargets.

Prioritized suspects and nontargets were classified, in accordance with Schymanski et al. (2014b), (i) as confirmed structures where reference material was available for identity confirmation (Level 1), (ii) as probable structures where an MS/MS library match was achieved (Level 2a) or (iii) as tentative structures where tentative identifications were based solely on MS/MS interpretation (Level 3). The annotation of MS/MS fragments with structural proposals for tentatively identified compounds was supported by CFM-ID 3.0. Level 1 candidates were identified and quantified as follows. Ten selected samples were enriched and measured together with four calibration standards (1, 10, 100, 1000 ng L<sup>-1</sup>) and



**Fig. 1.** Classification of samples according to the concentration sums of urban and agricultural targets. Dashed lines mark the threshold of 100 ng L<sup>-1</sup> used for classification. Size of circles correlates with number of detected targets (1–74 per sample). Axes are log-scaled.

six spiked samples (100, 250, 1000 ng L<sup>-1</sup>) with adjusted MS/MS settings. The determined calibration model was applied to the previously measured samples for quantification. The quality of quantification was evaluated for each compound based on relative recoveries in spiked samples and the reproducibility of concentrations in the samples, which were each measured twice. Not every compound could be quantified satisfactorily so that in some cases, either concentration ranges or no concentrations are reported. Further details regarding structural confirmation, quantification and associated MS/MS spectra can be found in SI-A7, SI-A12, and SI-B4, respectively.

### 3. Results and discussion

#### 3.1. Sample classification based on targets

The extent of urban and/or agricultural influence on the 60 groundwater monitoring sites was evaluated based on the sum concentration of 269 urban and 229 agricultural target compounds (for concentrations of individual targets see SI-B1, and for targets with detections  $\geq 100$  ng L<sup>-1</sup> see Table SI-A7). Accordingly, 17 sites showed only low urban and agricultural influence, 18 sites were predominantly influenced by agricultural targets, 22 sites were influenced by agricultural and urban targets and three sites were predominantly impacted by urban targets (Fig. 1). These classifications, which were based on the here presented target screening, were consistent with classifications based on long-term monitoring data for 53 out of 60 monitoring sites (for details see SI-A8). For most monitoring sites the classification was also consistent with the land use of the catchment, i.e. sites with high urban influence were often close to settlements or to wastewater impacted surface waters (bank filtration), sites with high agricultural influence were usually in areas with intensive agricultural land use, and the catchments of sites with low urban and agricultural influence were usually dominated by grassland or forest.

Urban classification was primarily associated with the sweetener acesulfame, the biocide TP N,N-dimethylsulfamide (may also originate from the banned plant protection product tolylfluanid), the industrial chemical melamine, and the corrosion inhibitor benzotriazole, i.e. if one of these targets is removed from the dataset, then two to five sites are no longer classified as having “high urban influence” or “high urban and agricultural influence” (Fig. 1).

Analogously, the agricultural classification was most driven by a TP of the fungicide chlorothalonil (R471811), which was the only target compound detected in each sample (maximum concentration 2200 ng L<sup>-1</sup>; Kiefer et al. (2020a)). For further details, see Figure SI-A2. If samples were classified based on detection frequency of urban or agricultural targets, using a cut-off of e.g. 10 detections, then 44 sites would be classified in the same way (Figure SI-A3). Here it should be noted that some pesticides may also be used as biocides (e.g. N,N-dimethylsulfamide (TP), triazine herbicides) and some pharmaceuticals are also used as veterinary drugs. Therefore, these targets may be related to urban and agricultural activities, potentially resulting in a wrong classification of some sites.

#### 3.2. Classification of nontargets

In the 60 groundwater samples, 6504 intensity profiles across samples (hereafter “compounds”) were detected with maximum intensity  $>10^6$  (Table 1), including 4800 nontargets and 98 targets in positive ionization mode and 1573 nontargets and 33 targets in negative ionization mode. The total number of compounds was likely slightly lower, due to ambiguities that arise during componentisation of isotopologues and adducts, e.g. ten target compounds were associated with several profiles. Furthermore, some compounds might be detected in both ionization modes, such as 15 target compounds. Of the 6373 nontargets, 4027 (63%) were found in less than five out of six pooled sample replicates or the intensity showed a relative standard deviation of  $>50\%$  (in pooled samples with detections, see Table 1), indicating that either the concentration was too low to be reproducibly detected (potentially due to dilution during mixing of sample aliquots) or that the peak shape was not reproducible. Moreover, for 296 compounds, the maximum intensity in the pooled samples was even higher than the maximum intensity in the groundwater samples. This may be related to contamination during sample handling or peak picking artefacts (e.g. noise).

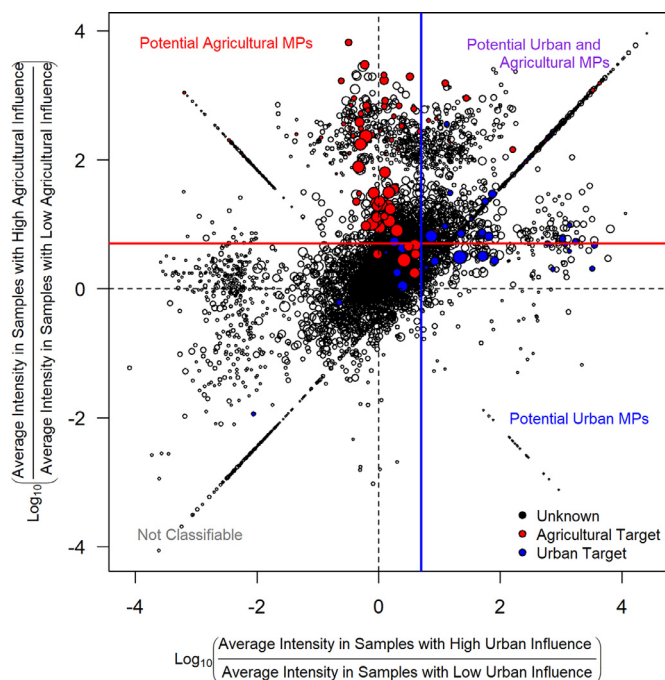
From the 6373 nontargets, 331 (5%) were classified as potential urban MPs, 945 (15%) as potential agricultural MPs, and 1892 (30%) as potential MPs from urban and agricultural sources. The remaining 3205 (50%) could not be assigned to one of these groups (not classifiable) and might be compounds of natural origins (Table 1, Fig. 2). More than 90% of profiles annotated as target compound were classified correctly, which was expected because classification criteria ( $>5$  times higher intensity in contaminated samples) were guided by the target compounds.

The “not classifiable” target compounds (centre of Fig. 2) were unclassified due to one or more of the following reasons: (i) a compound was incorrectly annotated (different compound with similar retention time); (ii) a compound had a noisy EIC, resulting in too many detections by enviMass; (iii) a compound originates from multiple sources (the sweetener saccharin may originate from wastewater or from pig manure (Buerge et al., 2011) and can therefore occur at less contaminated sites); (iv) a compound was detected in only one sample with low urban and agricultural influence (site-specific); (v) a compound had shifting retention time and was therefore not correctly grouped; (vi) compounds that have probably been spread ubiquitously, such as banned triazine pesticides and their TPs; and (vii) a compound eluted very early and was therefore affected by ion suppression. This early-eluting compound was melamine, an industrial chemical and TP of the larvicide cyromazine, which is often applied to manure (ECHA, 2015). Based on concentrations (determined with a structurally identical internal standard), melamine would be correctly classified as an MP from urban and agricultural sources. This compound was, however, hardly retained by the applied chromatographic method (retention time 3.4 min), eluting together with salts and other highly polar compounds, leading to strong ion suppression effects and,

**Table 1**

Classification of target and nontarget compounds as well as characterisation of nontarget compounds regarding maximum intensity, retention time, and detections in the 60 groundwater samples. Numbers in brackets correspond to the percentage of profiles that are less relevant for structure elucidation (e.g. false positives): (i) profiles with early retention time (3–4 min, i.e. salts and very polar compounds), (ii) profiles, for which a very broad (>5 min) peak shape is likely (i.e. if >2 profiles with retention time >10 min were grouped; potential natural organic matter (NOM), see SI-A9 for explanation), or (iii) profiles, for which the maximum intensity in the pooled samples was higher than the maximum intensity in the samples. The category “reproducible in pooled samples” provides the number of nontargets detected in at least five out of six pooled sample replicates with relative standard deviation <50% (calculated in pooled samples with detections).

	Urban	Urban and Agricultural	Agricultural	Not Classifiable	Total
<b>All Compounds</b>	345 (26%)	1929 (8%)	1013 (21%)	3217 (48%)	6504 (30%)
<b>Targets</b>					
Urban Targets	14 (0%)	21 (0%)	1 (0%)	6 (17%)	42 (2%)
Agricultural Targets	0	16 (0%)	67 (1%)	6 (0%)	89 (1%)
<b>Nontargets</b>					
Total (without Targets)	331 (27%)	1892 (8%)	945 (22%)	3205 (48%)	6373 (31%)
Retention Time <4 min	68 (100%)	81 (100%)	140 (100%)	1127 (100%)	1416 (100%)
Reproducible in Pooled Samples	122 (20%)	517 (9%)	318 (19%)	1389 (43%)	2346 (31%)
Potential NOM	9 (100%)	34 (100%)	44 (100%)	244 (100%)	327 (100%)
Intensity >5 × 10 <sup>6</sup>	56 (38%)	323 (12%)	86 (24%)	394 (72%)	859 (42%)
Intensity >10 <sup>7</sup>	21 (19%)	156 (15%)	30 (30%)	191 (76%)	398 (46%)
Intensity >5 × 10 <sup>7</sup>	2 (0%)	11 (0%)	3 (33%)	38 (82%)	54 (59%)
≥30 Detections	10 (40%)	24 (17%)	51 (24%)	862 (44%)	947 (43%)
Intensity >10 <sup>7</sup> & ≥30 Detections	6 (33%)	12 (25%)	14 (21%)	97 (86%)	129 (71%)
Intensity >10 <sup>7</sup> & <30 Detections	15 (13%)	144 (14%)	16 (38%)	94 (67%)	269 (34%)
Site-specific (1 Detection)	23 (17%)	727 (6%)	110 (39%)	106 (65%)	873 (16%)



**Fig. 2.** Compounds were classified based on their occurrence and intensity in samples with urban or agricultural influence. Each circle represents one compound; size of circle correlates with detection frequency in samples (1–60). Compounds to the right of/above the blue/red lines ( $\log_{10}(5)$ ) show, on average, more than five times higher intensities in samples with high urban/agricultural influence than in samples with low urban/agricultural influence. Compounds on the diagonals are artefacts resulting from the replacement of non-detects with the minimum intensity observed across the whole dataset (see Section 2.5.2).

in turn, low intensity (which is used for automatic classification), which hampered proper classification. In total, 20% of all detected compounds eluted between 3 and 4 min. Half of these compounds are located in the centre of Fig. 2, indicating either their ubiquitous occurrence or their possible misclassification, as in the case of melamine.

In addition, many compounds with broad late-eluting peaks, positive mass defect values and higher  $m/z$  were located in the

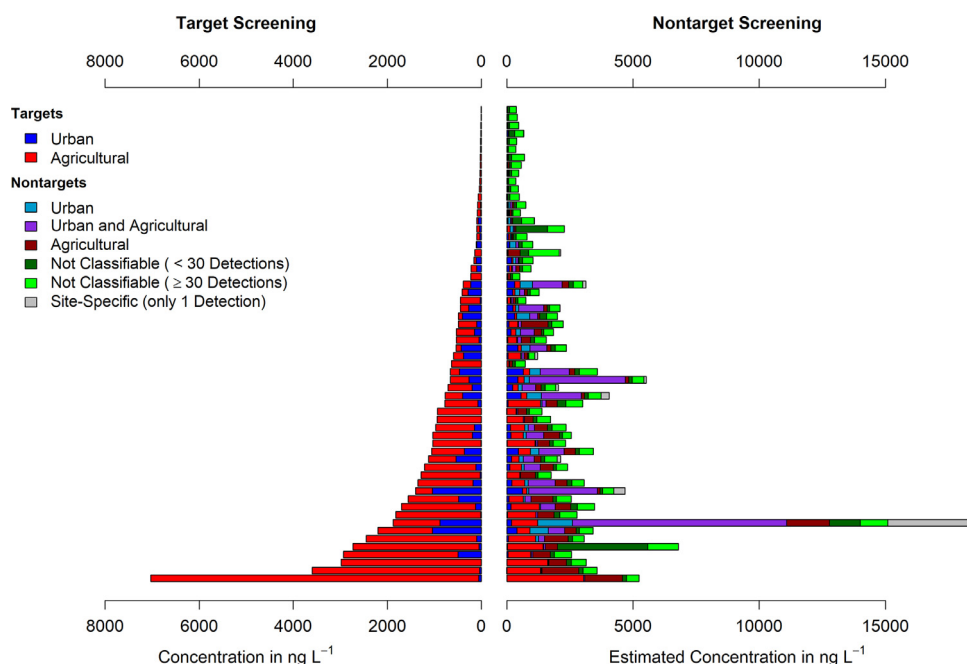
centre of Fig. 2. Many were assigned as potential natural organic matter and occurred either at each site or were randomly distributed (see SI-A9 for details; Table 1).

Further characterization and elucidation focused on the more intense compounds classified as potential MPs from urban or urban and agricultural sources (maximum intensity >5 × 10<sup>6</sup>) and the most prominent compounds from the remaining dataset (>10<sup>7</sup>, ≥30 detections). From the resulting 488 compounds in the positive and negative mode (SI-B2), 480 compounds had at least one MS/MS spectrum, demonstrating the effectiveness of AcquireX (Section 2.3). AcquireX increased the MS/MS coverage for the 488 nontargets by 39% (for suspects an increase of 73% was observed; detailed discussion in SI-A10). Overall, 409 compounds were annotated with between 1 and 576 candidate proposals using MetFrag and/or SIRIUS4/CSI:FingerID. Most compounds without candidate proposals (80%) eluted before 4 min and therefore might be artefacts from the LC-HRMS/MS analysis.

To get further evidence for an anthropogenic origin, 562 compounds were checked for characteristic isotope patterns such as Cl (SI-C1) and for their occurrence in effluent samples of two municipal wastewater treatment plants (SI-C2, including 74 agricultural compounds with maximum intensity >5 × 10<sup>6</sup>, which were not part of the 488 compounds). At least 39 of the 488 compounds contained one or more Cl atoms (see SI-B2 and discussion in SI-A9) based on their isotope pattern. In terms of occurrence, 38% of the 56 urban compounds were detected in at least one of the two effluent samples, only 12% of the 323 urban/agricultural compounds and 12% of the 86 agricultural compounds showed a peak in the effluent samples. Interestingly, also a high percentage of the 97 not classifiable compounds (39%) were found in the wastewater effluents (SI-B2). However, the majority of these not classifiable compounds (70%) eluted between 3 and 4 min, i.e. they were possibly misclassified or occur naturally (e.g. salts).

### 3.3. Characterization of groundwater quality at monitoring sites

Fig. 3 illustrates the total estimated concentrations of detected compounds at the 60 monitoring sites, after excluding potential false positives (see figure caption). The total concentrations were estimated assuming that the compounds ionize, on average, as efficiently as the target compounds (Section 2.5.2).



**Fig. 3.** Total concentrations determined in target screening (left) and estimated total concentrations determined from nontarget screening (right) for each monitoring site. The colour code indicates the potential source of the contamination. Concentrations were estimated assuming that the nontargets ionize on average as efficiently as target compounds (same ionization efficiency for each nontarget). Therefore, concentrations of urban and agricultural targets determined in the target screening may differ from estimated concentrations of targets determined in the nontarget screening. To reduce false positives, compounds were excluded if they (i) were identified as potential natural organic matter (broad peak, retention time > 10 min, SI-A9), (ii) had retention time < 4 min, (iii) had a maximum intensity in the pooled samples that was larger than the maximum intensity in the samples (296 compounds), or (iv) were not reproducibly detected in the pooled samples (< 5 detections or relative standard deviation > 50%). The early-eluting compounds were excluded because they are likely to be misclassified (intensity does not necessarily correlate with concentration) or to be of natural origins (e.g. salts). The exclusion of compounds not reproducibly detected in the pooled samples may lead to an underestimation of site-specific pollution. See Figure SI-A6 to Figure SI-A11 for similar plots based on detections or estimated concentrations, with and without excluded compounds.

The target screening, shown for comparison, demonstrates that anthropogenic activities affect the monitoring sites to very different extents. Total target concentrations ranged from 12 to 7000 ng L<sup>-1</sup>. For most monitoring sites, the nontarget screening confirmed the contamination trend observed in the target screening. Monitoring sites with few target detections also had less nontarget detections than sites where several targets were found (Figure SI-A7). Most not classifiable nontargets were detected in positive ionization mode (Figure SI-A8). To estimate which percentage of total contamination was explained by the target screening for 498 MPs, we compared the estimated concentration of targets with the estimated concentration of nontargets from each source (without not classifiable/site-specific nontargets, i.e. only 1 detection, and without potential false positives). According to this approximation, the targets would explain 4 to 72% of the total contamination in individual samples (median: 34%; pooled samples: 30–33%; based on detections: 8–28%, median: 16%, pooled samples 10–11%). Assuming that nontargets ionize as efficiently as target compounds, 46 nontargets had concentrations > 100 ng L<sup>-1</sup> in at least one groundwater sample. However, it should be kept in mind that these estimates are subject to various uncertainties: (i) the classification might be erroneous; (ii) the true ionization efficiency of individual compounds might differ considerably from the ionization efficiency assumed by our quantification approach (same ionization efficiency for all compounds); (iii) matrix effects (ion suppression and enhancement) may influence signal intensities and thus the estimated concentrations in individual samples; (iv) some potential MPs might have been detected several times, e.g. in both ionization modes or due to insufficient componentisation (see Section 3.2).

Despite these uncertainties, nontarget screening indicated that the contamination from agriculture at sites with high agricultural influence might be considerably higher than assumed based

solely on target screening. Indeed, based on the roughly estimated compound concentrations, 21–96% of potential agricultural MPs, detected at sites with high agricultural influence, would not be explained by targets (median: 49%, pooled samples: 48–52%; based on detections: 62–89%, median 77%, pooled samples: 81%). This was at first glance surprising since we had previously performed a suspect screening for most registered pesticides, including their TPs, using samples from partially the same monitoring sites (Kiefer et al., 2019). The suspect list comprised more than 1000 pesticide TPs, compiled from various sources and mostly observed within the European pesticide registration. However, for many pesticides applied in Switzerland between 2005 and 2017, relatively few TPs could be gathered for the suspect list; for 26% of pesticides transformation data was unavailable or inaccessible. Considering that for some pesticides, e.g. chlorothalonil, more than 20 TPs are known (EFSA, 2018), it is likely that the suspect list lacked important TPs, which were therefore not detected using the suspect screening approach.

Whether groundwater quality is indeed more affected by agriculture than by urban activities (as indicated by the target screening) remains unclear. Many nontarget compounds were jointly assigned to both urban and agricultural sources (“potential urban and agricultural MPs”; Fig. 3; purple), though some of these probably originate from only urban sources, as shown for some identified compounds (see Table 2 and Section 3.4.2). One monitoring site differed strongly from the remaining sites in terms of estimated concentrations, total number of nontarget compounds and the number of site-specific compounds (Fig. 3, Figure SI-A6 to Figure SI-A11). Strikingly, this site was not suspicious in the target screening, except for high concentrations of the pesticide cycluron (140 ng L<sup>-1</sup>; SI-B1), which was banned in Switzerland in 2005 and in the EU in 2003 (European Commission, 2002). The elucidation



**Table 2**

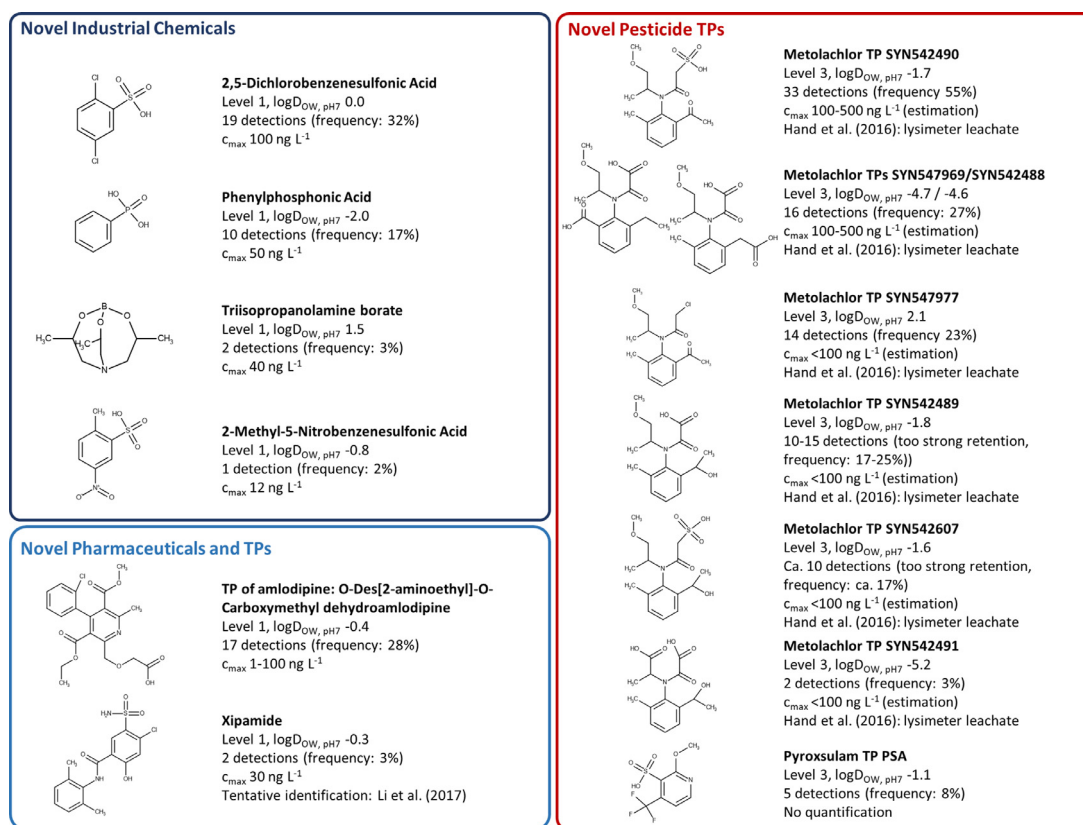
Identified suspects and nontargets, reported previously in literature and ordered according to confirmation confidence and detection frequency in groundwater. The Pub-Chem Compound ID (CID) is provided as identifier. Further identifiers (InChIKey, SMILES) and quantification results are listed in SI-B4. For MS/MS fragments, see SI-A12. See SI-B2 and SI-B3 for all prioritized nontargets and suspects, including annotated candidates. References point to studies reporting prior detections in environmental samples; n.q. = no quantification.

Compound	Screening	Classification	Maximum concentration in ng L <sup>-1</sup>	No. of detections	Use/sources	Literature on environmental occurrence
<b>Trifluoroacetic acid*</b> Level 1, CID 6422	Suspect	Not Classifiable	>5000	60	Various sources	Berg et al. (2000), Scheurer et al. (2017)
<b>Trifluoromethanesulphonic acid*</b> Level 1, CID 62406	Suspect	Agricultural	90	53	Industrial chemical	Zahn et al. (2016)
<b>Atrazine-desethyl-desisopropyl</b> Level 1, CID 18831	Nontarget	Agricultural	90	51	Pesticide TP	BMASGK (2018)
<b>Perfluoropropanesulfonic Acid</b> Level 1, CID 9859771	Nontarget	Urban & Agricultural	5–30	51	Perfluorinated compound	Mak et al. (2009)
<b>Oxypurinol*</b> Level 1, CID 135398752	Nontarget	Urban	300	20	Pharmaceutical TP	Funke et al. (2015)
<b>Methyldiphenylphosphine oxide**</b> Level 1, CID 75041	Nontarget	Urban & Agricultural	n.q.	23	Industrial chemical	Brand et al. (2018)
<b>Edetic acid (EDTA)*</b> Level 1, CID 6049	Suspect	Urban	n.q.	20	Industrial chemical	Schmidt et al. (2004)
<b>2-Acrylamido-2-methyl-1-propanesulfonic acid (AMPS)*</b> Level 1, CID 65360	Suspect	Urban & Agricultural	90	13	Wide-spread use (industry and households)	Schulze et al. (2019)
<b>Perfluorobutylsulphonamide</b> Level 1, CID 10958205	Nontarget	Urban & Agricultural	n.q.	13	Perfluorinated compound	Chu et al. (2015) in fish
<b>Methenamine*</b> Level 1, CID 4101	Suspect	Urban	<200	11	Wide-spread use (industry and households)	Knepper et al. (1999)
<b>Dimethylbenzenesulfonic acid (isomers)*</b> Level 1	Suspect	Urban & Agricultural	80	10	Industrial chemical	Betowski et al. (1996)
<b>p-Toluenesulfonic acid*</b> Level 1, CID 6101	Suspect	Not Classifiable	200–700	7	Industrial chemical	Crathorne et al. (1984)
<b>5-Methoxy-2H-benzotriazole</b> Level 1, CID 119717	Nontarget	Urban & Agricultural	3	7	Benzotriazole derivate or TP thereof	Liu et al. (2013), Huntscha et al. (2014) tentative structure Heberer (2002)
<b>Propyphenazone</b> Level 1, CID 3778	Nontarget	Urban & Agricultural	10	5	Pharmaceutical	
<b>Pyrimidinol (2-Isopropyl-6-methyl-4-pyrimidone)</b> Level 1, CID 135444498	Nontarget	Urban & Agricultural	60	2	Pesticide / biocide TP (diazinon TP)	Diaz-Cruz and Barcelo (2006)
<b>Fluometuron</b> Level 1, CID 16562	Suspect	Urban & Agricultural	40	2	Pesticide / biocide	Herrero Hernández et al. (2012)
<b>Sulisobenzene</b> Level 1, CID 19988	Suspect	Site-specific	50–100	1	UV filter, various uses	Rodil et al. (2008)
<b>Iopromide TP 643</b> Level 2a, CID 139597923	Nontarget	Urban & Agricultural	n.q.	26	Pharmaceutical TP	Schulz et al. (2008)
<b>Iopromide TP 701 A</b> Level 2a, CID 139596314	Nontarget	Urban & Agricultural	n.q.	23	Pharmaceutical TP	Schulz et al. (2008)
<b>Iomeprol TP 629</b> Level 2a, CID 23189998	Nontarget	Urban	n.q.	18	Pharmaceutical TP	Kormos et al. (2009)
<b>Triphenylphosphine oxide*</b> Level 2a, CID 13097	Nontarget	Urban & Agricultural	n.q.	16	Industrial chemical (synthesis by-product)	Knepper et al. (1999)
<b>Hexa(methoxymethyl)-melamine</b> Level 2a, CID 62479	Suspect	Not Classifiable	n.q.	7	Industrial chemical	Dsikowitzky and Schwarzbauer (2015), Bobeldijk et al. (2002)
<b>Metolachlor TP CGA357704</b> Level 2a, CID 71312482			<100	6	Pesticide TP	Reemtsma et al. (2013)
<b>Chlorendic acid</b> Level 2a, CID 8266	Nontarget	Site-specific	n.q.	1	Industrial chemical, TP of organochlorine pesticides	Ying et al. (1986), IPCS (1996)
<b>Isomer of 5,6-Dimethyl-2H-benzotriazole</b> Level 3	Nontarget	Urban	1–10	15	Benzotriazole derivate or TP thereof	Huntscha et al. (2014), Trcek et al. (2018)
<b>Isomer of 5-Methoxy-2H-benzotriazole</b> Level 3	Nontarget	Urban & Agricultural	10–30	14	Benzotriazole derivate or TP thereof	Liu et al. (2013), Huntscha et al. (2014) tentative structure
<b>Naphthalenedisulfonic Acids (various isomers)*</b> Level 3	Nontarget	Urban & Agricultural	>500	1	Industrial chemical	Jekel and Gruenheid (2005), Knepper et al. (1999)

\* Contamination during sample processing cannot be fully excluded. Only detections are reported, meaning that the concentration in groundwater samples was at least twice as high as the maximum concentration in 18 blank samples. Where quantification was not possible, an intensity threshold was instead applied, requiring a sample intensity at least five times higher than the maximum intensity recorded in blank samples.

\*\* The METLIN MS/MS library reports different MS/MS fragments than recorded in this study.





**Fig. 4.** Novel MPs elucidated through nontarget and suspect screening. For more details, including structural identifiers and quantification results in individual samples, see SI-B4. For MS/MS fragments, see SI-A12. Two metolachlor TPs were too strongly retained by LC, i.e. partially eluting in subsequent samples. logD<sub>OW, pH7</sub> (water-n-octanol distribution coefficient considering the speciation at pH 7) was predicted with JChem for Office (Version 17.1.2300.1455; ChemAxon Ltd.).

of some nontargets supports the hypothesis that this monitoring site is highly contaminated (Section 3.4).

### 3.4. Identification of nontargets and suspects

Structural elucidation efforts were especially successful where (i) nontargets were annotated with relatively few structural candidates (12 of 21 unequivocally or tentatively identified nontargets had <10 candidates), (ii) MS/MS spectra of the correct candidate were available in libraries or literature (13 nontargets), (iii) useful metadata was accessible (e.g. information on field of application for candidates), and finally (iv) reference material could be purchased.

Using 29 reference standards, we confirmed 11 suspects and 12 nontargets (Level 1) and rejected three suspects and three nontargets. Moreover, five nontargets and two suspects could be identified as probable structures by a library spectrum match (Level 2a), while a further five nontargets and five suspects were assigned tentative structures (Level 3). Fig. 4 presents novel compounds (Level 1 or Level 3) and Table 2 compounds, which were already reported to be detected in environmental samples. Further details, including quantification results and MS/MS spectra, are given in SI-A12 and SI-B4. In addition, MS/MS spectra will be uploaded to MassBank ([www.massbank.eu](http://www.massbank.eu)).

#### 3.4.1. Novel micropollutants

**2,5-dichlorobenzenesulfonic acid** is pre-registered under REACH and was predicted as likely to be carcinogenic and persistent in the environment (ECHA 2020). 2,5-dichlorobenzenesulfonic acid was detected in 19 out of 60 samples, mostly in the low ng L<sup>-1</sup> range, in one sample at 100 ng L<sup>-1</sup>. We did not find evi-

dence for a prior detection in environmental samples in available literature.

**Phenylphosphonic acid** is registered under REACH (10–100 t a<sup>-1</sup>, ECHA (2020)) and was detected at 10 monitoring sites with concentrations of up to 50 ng L<sup>-1</sup>. Schulze et al. (2019) compared various analytical approaches for the analysis of polar compounds. Phenylphosphonic acid could only be analysed by reversed-phase LC, but the tested enrichment methods, all based on solid phase extraction, were unsuitable for this industrial chemical. The occurrence in environmental samples after evaporative enrichment is here reported for the first time.

**O-des[2-aminoethyl]-O-carboxymethyl dehydroamlodipine** is a TP of the blood pressure regulator amlodipine, which is approved in Switzerland. The TP was detected in nearly one third of the samples at concentrations <100 ng L<sup>-1</sup> and is probably reported here for the first time in environmental samples.

**Triisopropanolamine borate** was detected at two monitoring sites at concentrations of up to 40 ng L<sup>-1</sup> and was to our knowledge not reported before in the environment. The sulphonamide diuretic drug **xipamide** was found in two samples at 30 ng L<sup>-1</sup>. Previously, Li et al. (2017) identified xipamide only tentatively in European wastewater-impacted rivers. Interestingly, xipamide is not approved in Switzerland and was therefore also not on the target list (PharmaWiki, 2020). One nontarget compound, detected at five monitoring sites and classified as a potential urban and agricultural MP, was tentatively identified as the pesticide TP **pyroxsulam TP PSA** (reference standard not available).

Amongst the top ten most intense potential agricultural MPs with a detection frequency >50%, we tentatively identified **metolachlor TP SYN542490** (reference standards not available). Assuming a similar ionization efficiency as metolachlor-ESA and meto-

lachlor TP CGA 368208, maximum concentrations were estimated to be in the range of 100–500 ng L<sup>-1</sup>, which is comparable to the concentrations of the target compounds metolachlor-ESA and metolachlor TP CGA 368208. S-metolachlor is currently under review for renewal in the EU. In contrast to our previously performed target and suspect screening (Kiefer et al., 2019), which included nine metolachlor TPs, here we had access to nine further metolachlor TPs (including TP SYN542490; T. Poiger, personal communication). They were previously observed in a lysimeter study by the pesticide producer (Hand et al., 2016). After the detection of SYN542490, we manually screened for the remaining TPs and tentatively identified the TPs CGA357704 (reported previously by Reemtsma et al. (2013)), SYN542607, SYN547977, SYN542489, SYN542491, and one or both of the isomers, namely SYN547969 and SYN542488; all of which were detected at lower intensity and lower frequency compared to SYN542490.

### 3.4.2. Evaluation of pre-classification of elucidated compounds

The identification efforts focused on compounds classified as potential MPs from urban sources and urban and agricultural sources. Probably all elucidated compounds with this pre-classification originate from urban or industrial sources, except for the tentatively identified TP of the herbicide pyroxsulam, namely pyroxsulam TP PSA. Furthermore, we investigated the most intense and wide-spread potential agricultural pollutants leading to the identification of TPs of the pesticides metolachlor and atrazine, confirming the correct pre-classification. Interestingly, in addition to the mentioned pesticide TPs, we also detected the short-chain perfluorinated compound trifluoromethanesulfonic acid (53 detections), which has previously been reported by Zahn et al. (2016) and Schulze et al. (2019) as a wide-spread MP with concentrations in the ng L<sup>-1</sup> to µg L<sup>-1</sup> range. The highest concentrations in our study (up to 90 ng L<sup>-1</sup>) were detected in samples with high agricultural influence, explaining its classification as a potential agricultural MP. In contrast to another short-chain perfluorinated compound, trifluoroacetic acid, for which various sources have been reported including the degradation of pesticides (Scheurer et al., 2017), trifluoromethanesulfonic acid is to our knowledge not known as being of potential agricultural origin.

Four elucidated suspects could not be related to urban and/or agricultural sources, i.e. were “not classifiable”. The industrial chemical hexa(methoxymethyl)melamine showed, on average, only four times higher intensity in samples with high urban influence compared to the average intensity in samples with low urban influence and was therefore “not classifiable” (threshold was five, Section 2.5.2). In the case of p-toluenesulfonic acid, contamination during sample handling or analysis led to false positive detections in a few samples, including the blank samples. This compound is a known background contaminant (Schulze et al., 2019). On average its intensity was 28 times higher in samples than in blanks, meaning it was not removed during pre-processing (Section 2.5.1). Similarly, 1,3-diphenylguanidine was detected in all samples and in 19 of these had an intensity five times greater than in blanks. However, in spiking experiments (100 ng L<sup>-1</sup>) we observed 3–4 times higher intensities in the sample matrix than in ultrapure water (ion enhancement), suggesting that the detections with relatively high intensity might represent contaminations. Therefore, 1,3-diphenylguanidine is not reported in Table 2. A further “not classifiable” compound was trifluoroacetic acid, which is a ubiquitously spread pollutant and may enter groundwater via diffuse sources (atmospheric deposition, pesticide application) or point sources (industrial emissions to rivers) (Berg et al., 2000; Freeling et al., 2020; Scheurer et al., 2017). Accordingly, trifluoroacetic acid was detected in all samples, though reliable quantification was not possible.

Many identified nontargets and suspects were pre-classified as “potential urban and agricultural MP”, although they likely originate only from urban sources (e.g. TPs of x-ray contrast agent iopromide, benzotriazole derivatives, Table 2). This imprecise classification results from the large number of monitoring sites, which were influenced by both urban and agricultural activities, whereas only few sites showed a high urban but low agricultural influence (Fig. 1). To achieve a better classification of urban MPs, more monitoring sites with primarily urban influence would be needed for the workflow, but this was difficult to obtain due to the Swiss small-scale structured landscape.

The characterization of the monitoring sites (Section 3.3) highlighted one site in particular, due to its especially high contamination in terms of both the number and estimated concentrations of detected nontargets (Fig. 3, Figure SI-A6 to Figure SI-A11). Out of 46 nontargets with estimated concentrations >100 ng L<sup>-1</sup> (Section 3.3; 10 of the 46 were identified), 23 were detected at this site. The hypothesis that this site might be highly contaminated was further supported by the elucidation of various compounds, being either site-specific or showing highest concentrations at this site, such as the industrial chemicals naphthalenedisulfonic acids (various isomers, >500 ng L<sup>-1</sup>) and p-toluenesulfonic acid (200–700 ng L<sup>-1</sup>), the diazinone TP pyrimidinol (60 ng L<sup>-1</sup>), the industrial chemicals dimethylbenzenesulfonic acids (various isomers, 80 ng L<sup>-1</sup>) and 2,5-dichlorobenzenesulfonic acid (100 ng L<sup>-1</sup>), and chlorendic acid (Level 2a). The ECHA classifies chlorendic acid as a carcinogen category 1B (ECHA, 2020). Potential sources include the degradation of flame-retardant polyesters or organochlorine pesticides such as endosulfan, heptachlor, or aldrin (IPCS, 1996). The reason for this contamination is unknown, but the groundwater from this site is not currently used for drinking water production.

## 4. Conclusions

By combining an appropriate compound prioritization strategy with a highly automated structural elucidation workflow, we were able to characterize groundwater quality in a more comprehensive manner than previously possible using targeted methods and in doing so, to identify as yet unreported MPs.

- Nontargets were prioritized based on their potential origin for structural elucidation. Categorisation of nontargets also provided rough estimates of the number and the concentrations of thus far overlooked contaminants.
- A combination of computational tools supported the structural elucidation process. AcquireX improved MS/MS coverage for nontargets and suspects, while MetFrag and SIR-IUS4/CSI:FingerID (together with an extensive compound database) resulted in 23 unequivocally identified and 17 tentatively identified compounds. 13 of these compounds are novel.
- Structural elucidation was most successful for compounds (i) with MS/MS spectra in libraries or literature, (ii) that were assigned limited numbers of candidate structures, (iii) that had accessible metadata, and (iv) for which reference material was available.
- Despite the high degree of automation of the structural elucidation workflow, structural elucidation itself remains a major bottleneck in transforming unknowns into known compounds. Moreover, elucidation of compounds not presently in any database, e.g. so far not observed TPs, is highly challenging and was not covered at all by our elucidation approach.
- One groundwater sample was revealed to be much more polluted than assumed based on target screening, highlighting the relevance of comprehensive screening approaches for evaluating water quality.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This study was funded by the Swiss Federal Office for the Environment (FOEN). Samples were provided by the Swiss cantons and by the Swiss National Monitoring NAQUA. We thank Miriam Reinhardt, FOEN, project management NAQUA, and all those involved in sample collection and transport. Furthermore, we are grateful to Philipp Longrée and Bernadette Vogler, Eawag, for their excellent support in the laboratory. We are especially thankful to Emma Schymanski, Luxembourg Centre for Systems Biomedicine, for an overview on available compound lists, to Martin Loos, envibee GmbH, for his support with enviMass, and to Jennifer Schollée and Michael Stravs, Eawag, for sharing R functions and helpful discussions. Besides, we received valuable feedback to the manuscript from Miriam Reinhardt and Ronald Kozel, FOEN. Special thanks go to Jennifer Schollée and Martin Jones, Eawag, for proofreading and critical comments.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.watres.2021.116994](https://doi.org/10.1016/j.watres.2021.116994).

## References

- Albergamo, V., Schollée, J.E., Schymanski, E.L., Helmus, R., Timmer, H., Hollender, J., de Voogt, P., 2019. Nontarget screening reveals time trends of polar micropollutants in a riverbank filtration system. *Environ. Sci. Technol.* 53 (13), 7584–7594.
- Allen, F., Greiner, R., Wishart, D., 2014. Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification. *Metabolomics* 11 (1), 98–110.
- Arp, H.P. and Hale, S.E., 2020. S36 | UBAPMT | Potential Persistent, Mobile and Toxic (PMT) substances (Version NORMAN-SLE-S36.0.2.0) [Data set], Zenodo. 10.5281/zenodo.3637611.
- Arp, H.P.H., Hale, S.E., 2019. UBA Texte 126/2019: REACH: Improvement of Guidance Methods For the Identification and Evaluation of PM/PMT Substances. German Environment Agency (UBA), Dessau-Roßlau, Germany.
- Berg, M., Müller, S.R., Mühlemann, J., Wiedmer, A., Schwarzenbach, R.P., 2000. Concentrations and mass fluxes of chloroacetic acids and trifluoroacetic acid in rain and natural waters in Switzerland. *Environ. Sci. Technol.* 34 (13), 2675–2683.
- Betowski, L.D., Kendall, D.S., Pace, C.M., Donnelly, J.R., 1996. Characterization of groundwater samples from superfund sites by gas chromatography/mass spectrometry and liquid chromatography/mass spectrometry. *Environ. Sci. Technol.* 30 (12), 3558–3564.
- BMASGK, 2018. Austrian Report on Drinking Water 2017 (Österreichischer Trinkwasserbericht 2017). [https://www.verbrauchergesundheits.gv.at/lebensmittel/trinkwasser/Trinkwasserbericht\\_2017.pdf?7i5641](https://www.verbrauchergesundheits.gv.at/lebensmittel/trinkwasser/Trinkwasserbericht_2017.pdf?7i5641).
- Bobeldijk, I., Stoks, P.G.M., Vissers, J.P.C., Emke, E., van Leerdam, J.A., Muilwijk, B., Berbee, R., Noij, T.H.M., 2002. Surface and wastewater quality monitoring: combination of liquid chromatography with (geno)toxicity detection, diode array detection and tandem mass spectrometry for identification of pollutants. *J. Chromatogr. A* 970 (1), 167–181.
- Bolton, E. and Schymanski, E.L., 2020. PubChemLite tier0 and tier1 (Version PubChemLite.0.2.0) [Data set], Zenodo. 10.5281/zenodo.3611238.
- Brand, S., Schlüsener, M.P., Albrecht, D., Kunkel, U., Strobel, C., Grummt, T., Ternes, T.A., 2018. Quaternary (triphenyl-) phosphonium compounds: environmental behavior and toxicity. *Water Res.* 136, 207–219.
- Broadhurst, D., Goodacre, R., Reinke, S.N., Kuligowski, J., Wilson, I.D., Lewis, M.R., Dunn, W.B., 2018. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* 14 (6), 72.
- Buerge, I.J., Keller, M., Buser, H.R., Müller, M.D., Poiger, T., 2011. Saccharin and other artificial sweeteners in soils: estimated inputs from agriculture and households, degradation, and leaching to groundwater. *Environ. Sci. Technol.* 45 (2), 615–621.
- Chambers, M.C., Maclean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egerton, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T.A., Brusniak, M.-Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S.L., Nuwaysir, L.M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E.W., Moritz, R.L., Katz, J.E., Agus, D.B., MacCoss, M., Tabb, D.L., Mallick, P., 2012. A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* 30, 918.
- Chiaia-Hernandez, A.C., Schymanski, E.L., Kumar, P., Singer, H.P., Hollender, J., 2014. Suspect and nontarget screening approaches to identify organic contaminant records in lake sediments. *Anal. Bioanal. Chem.* 406 (28), 7323–7335.
- Chu, S., Letcher, R.J., McGoldrick, D.J., Backus, S.M., 2015. A new fluorinated surfactant contaminant in biota: perfluorobutane sulfonamide in several fish species. *Environ. Sci. Technol.* 50 (2), 669–675.
- Crathorne, B., Fielding, M., Steel, C.P., Watts, C.D., 1984. Organic compounds in water: analysis using coupled-column high-performance liquid chromatography and soft-ionization mass spectrometry. *Environ. Sci. Technol.* 18 (10), 797–802.
- Diaz-Cruz, M.S., Barcelo, D., 2006. Highly selective sample preparation and gas chromatographic-mass spectrometric analysis of chlorpyrifos, diazinon and their major metabolites in sludge and sludge-fertilized agricultural soils. *J. Chromatogr. A* 1132 (1–2), 21–27.
- Dsikowitzky, L., Schwarzbauer, J., 2015. Hexa(methoxymethyl)melamine: an Emerging contaminant in German Rivers. *Water Environ. Res.* 87 (5), 461–469.
- Dührkop, K., Shen, H., Meusel, M., Rousu, J., Bocker, S., 2015. Searching molecular structure databases with tandem mass spectra using CSI:fingerID. *Proc. Natl. Acad. Sci. U. S. A.* 112 (41), 12580–12585.
- Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A.A., Melnik, A.V., Meusel, M., Dorrestein, P.C., Rousu, J., Bocker, S., 2019. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* 16 (4), 299–302.
- ECHA, 2015. Opinion of the Biocidal Products Committee on the application for approval of the active substance cyromazine for product type 18; ECHA/BPC/087/2015. <https://echa.europa.eu/documents/10162/a281e819-013d-40f5-aae3-719aac9489b7>. 21/07/2020.
- ECHA, 2020. European Chemicals Agency. <https://www.echa.europa.eu/information-on-chemicals>. Accessed: Sep 2020.
- EFSA, 2018. Peer review of the pesticide risk assessment of the active substance chlorothalonil. EFSA J. 16 (1), 1–40.
- European Commission, 1998. Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. Official J. Eur. Commun. (L330).
- European Commission, 2002. Commission regulation (EC) No 2076/2002 of 20 November 2002 extending the time period referred to in Article 8(2) of Council Directive 91/414/EEC and concerning the non-inclusion of certain active substances in Annex I to that Directive and the withdrawal of authorisations for plant protection products containing these substances. Official J. Eur. Commun. 319 (4).
- Fischer, S., 2017. S17 | KEMIMARKET | KEMI Market List MS-ready (12/05/2017) (Version NORMAN-SLE-S17.0.1.3) [Data set], Zenodo. 10.5281/zenodo.3653175.
- Freeling, F., Behringer, D., Heydel, F., Scheurer, M., Ternes, T.A., Nodler, K., 2020. Trifluoroacetate in precipitation: deriving a benchmark data set. *Environ. Sci. Technol.* 54, 11210–11219.
- Funke, J., Prasse, C., Lutke Eversloh, C., Ternes, T.A., 2015. Oxypurinol - A novel marker for wastewater contamination of the aquatic environment. *Water Res.* 74, 257–265.
- Gago-Ferrero, P., Schymanski, E.L., Bletsou, A.A., Aalizadeh, R., Hollender, J., Thomaidis, N.S., 2015. Extended suspect and non-target strategies to characterize emerging polar organic contaminants in raw wastewater with LC-HRMS/MS. *Environ. Sci. Technol.* 49 (20), 12333–12341.
- Gago-Ferrero, P., Krettek, A., Fischer, S., Wiberg, K., Ahrens, L., 2018. Suspect screening and regulatory databases: a powerful combination to identify emerging micropollutants. *Environ. Sci. Technol.* 52 (12), 6881–6894.
- Gatto, L., Lilley, K.S., 2012. MSnbase - an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics* (28) 288–289.
- Hand, L.H., Marshall, S.J., Saeed, M., Earll, M., Hadfield, S.T., Richardson, K., Rawlinson, P., 2016. High-resolution accurate mass spectrometry as a technique for characterization of complex lysimeter leachate samples. *Environ. Toxicol. Chem.* 35 (6), 1401–1412.
- Heberer, T., 2002. Tracking persistent pharmaceutical residues from municipal sewage to drinking water. *J. Hydrol. (Amsterdam)* 266 (3–4), 175–189.
- Heberer, T., Mechlini, A., Fanck, B., Knappe, A., Massmann, G., Pekdeger, A., Fritz, B., 2004. Field studies on the fate and transport of pharmaceutical residues in bank filtration. *Ground Water Monitoring Remediation* 24 (2), 70–77.
- Herrero Hernández, E., Pose Juan, E., Álvarez Martín, A., Andrades, M., Rodríguez Cruz, M., Sánchez-Martín, M.J., 2012. Pesticides and degradation products in groundwaters from a vineyard region: optimization of a multiresidue method based on SPE and GC-MS. *J. Sep. Sci.* 35 (24), 3492–3500.
- 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., 2010. MassBank: a public repository for sharing mass spectral data for life sciences. *J. Mass Spectrometry* 45 (7), 703–714.
- Hug, C., Ulrich, N., Schulze, T., Brack, W., Krauss, M., 2014. Identification of novel micropollutants in wastewater by a combination of suspect and nontarget screening. *Environ. Pollut.* 184, 25–32.
- Huntscha, S., Hofstetter, T.B., Schymanski, E.L., Spahr, S., Hollender, J., 2014. Bio-transformation of benzotriazoles: insights from transformation product identification and compound-specific isotope analysis. *Environ. Sci. Technol.* 48 (8), 4435–4443.



- IPCS, 1996. Environmental Health Criteria 185, Chlorendic Acid and Anhydride. International Programme On Chemical Safety. World Health Organization.
- Jekel, M., Gruenheid, S., 2005. Bank filtration and groundwater recharge for treatment of polluted surface waters. *Water Supply* 5 (5), 57–66.
- Kiefer, K., Müller, A., Singer, H., Hollender, J., 2019. New relevant pesticide transformation products in groundwater detected using target and suspect screening for agricultural and urban micropollutants with LC-HRMS. *Water Res.* 165.
- Kiefer, K., Bader, T., Minas, N., Salhi, E., Janssen, E.M.L., von Gunten, U., Hollender, J., 2020a. Chlorothalonil transformation products in drinking water resources: widespread and challenging to abate. *Water Res.* 183.
- Kiefer, K., Müller, A., Singer, H. and Hollender, J., 2020b. S60 | SWISSPEST19 | Swiss Pesticides and Metabolites from Kiefer et al. 2019 (Version S60.0.1.1) [Data set], Zenodo. 10.5281/zenodo.3766352.
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B.A., Thiessen, P.A., Yu, B., Zaslavsky, L., Zhang, J., Bolton, E.E., 2019. PubChem 2019 update: improved access to chemical data. *Nucleic. Acids. Res.* 47 (D1), D1102–D1109.
- Knepper, T.P., Sacher, F., Lange, F.T., Brauch, H.J., Karrenbrock, F., Roerden, O., Lindner, K., 1999. Detection of polar organic substances relevant for drinking water. *Waste Manag.* 19 (2), 77–99.
- Kolpin, D.W., Battaglin, W.A., Conn, K.E., Furlong, E.T., Glassmeyer, S.T., Kalkhoff, S.J., Meyer, M.T., Schnoebelen, D.J., 2009. In: Boxall, A.B.A. (Ed.), *Transformation Products of Synthetic Chemicals in the Environment*. Springer, Berlin, Heidelberg, pp. 83–100 ed.
- Kormos, J.L., Schulz, M., Wagner, M., Ternes, T.A., 2009. Multistep approach for the structural identification of biotransformation products of iodinated x-ray contrast media by liquid chromatography/hybrid triple quadrupole linear ion trap mass spectrometry and <sup>1</sup>H and <sup>13</sup>C nuclear magnetic resonance. *Anal. Chem.* 81 (22), 9216–9224.
- Krauss, M., Singer, H., Hollender, J., 2010. LC-high resolution MS in environmental analysis: from target screening to the identification of unknowns. *Anal. Bioanal. Chem.* 397 (3), 943–951.
- Letzel, T., Grosse, S. and Sengel, M., 2017. S2 | STOFFIDENT | HSWT/LfU STOFF-IDENT Database of Water-Relevant Substances (Version NORMAN-SLE-S2.0.1.0) [Data set], Zenodo. 10.5281/zenodo.2621452.
- Li, Z., Kaserzon, S.L., Plassmann, M.M., Sobek, A., Gomez Ramos, M.J., Radke, M., 2017. A strategic screening approach to identify transformation products of organic micropollutants formed in natural waters. *Environ. Sci. Process. Impacts* 19 (4), 488–498.
- Liu, Y.S., Ying, G.G., Shareef, A., Kookana, R.S., 2013. Biodegradation of three selected benzotriazoles in aquifer materials under aerobic and anaerobic conditions. *J. Contam. Hydrol.* 151, 131–139.
- Loos, R., Locoro, G., Comero, S., Contini, S., Schwesig, D., Werres, F., Balsaa, P., Gans, O., Weiss, S., Blaha, L., Bolchi, M., Gawlik, B.M., 2010. Pan-European survey on the occurrence of selected polar organic persistent pollutants in ground water. *Water Res.* 44 (14), 4115–4126.
- Mak, Y.L., Taniyasu, S., Yeung, L.W.Y., Lu, G., Jin, L., Yang, Y., Lam, P.K.S., Kannan, K., Yamashita, N., 2009. Perfluorinated compounds in tap water from china and several other countries. *Environ. Sci. Technol.* 43 (13), 4824–4829.
- Mechelke, J., Longree, P., Singer, H., Hollender, J., 2019. Vacuum-assisted evaporative concentration combined with LC-HRMS/MS for ultra-trace-level screening of organic micropollutants in environmental water samples. *Anal. Bioanal. Chem.* 411, 2555–2567.
- Müller, A., Schulz, W., Ruck, W.K., Weber, W.H., 2011. A new approach to data evaluation in the non-target screening of organic trace substances in water analysis. *Chemosphere* 85 (8), 1211–1219.
- Norman Network, Aalizadeh, R., Alygizakis, N.A., Schymanski, E.L., Slobodnik, J. and Fischer, S., 2020. S0 | SUSDAT | Merged NORMAN Suspect List: susDat (Version NORMAN-SLE-S0.0.2.2) [Data set], Zenodo. 10.5281/zenodo.3695732.
- PharmaWiki, 2020. Xipamid. <https://www.pharmawiki.ch/wiki/index.php?wiki=Xipamid>. Accessed: 14/11/2020.
- Pimentel, D., 2009. Integrated Pest Management: Innovation-Development Process. In: Peshin, R., Dhawan, A.K. (Eds.). In: *Integrated Pest Management: Innovation-Development Process, Volume 1*. Springer, Netherlands, Dordrecht, pp. 83–87.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/>.
- Reemtsma, T., Alder, L., Banasiak, U., 2013. Emerging pesticide metabolites in groundwater and surface water as determined by the application of a multi-method for 150 pesticide metabolites. *Water Res.* 47 (15), 5535–5545.
- Rodil, R., Quintana, J.B., López-Mahía, P., Muniategui-Lorenzo, S., Prada-Rodríguez, D., 2008. Multiclass determination of sunscreen chemicals in water samples by liquid chromatography–tandem mass spectrometry. *Anal. Chem.* 80 (4), 1307–1315.
- Ruttkies, C., Schymanski, E.L., Wolf, S., Hollender, J., Neumann, S., 2016. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J. Cheminform.* 8 (3).
- Sangster, T., Major, H., Plumb, R., Wilson, A.J., Wilson, I.D., 2006. A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabolonomic analysis. *Analyst* 131 (10), 1075–1078.
- Scheurer, M., Nödler, K., Freiling, F., Janda, J., Happel, O., Riegel, M., Müller, U., Storck, F.R., Fleig, M., Lange, F.T., Brunsch, A., Brauch, H.-J., 2017. Small, mobile, persistent: trifluoroacetate in the water cycle – overlooked sources, pathways, and consequences for drinking water supply. *Water Res.* 126, 460–471.
- Schlüsener, M.P., Kunkel, U., Ternes, T.A., 2015. Quaternary triphenylphosphonium compounds: a new class of environmental pollutants. *Environ. Sci. Technol.* 49 (24), 14282–14291.
- Schmidt, C.K., Fleig, M., Sacher, F., Brauch, H.-J., 2004. Occurrence of aminopolycarboxylates in the aquatic environment of Germany. *Environ. Pollution* 131 (1), 107–124.
- Schmidt, C.K., Brauch, H.-J., 2008. N,N-Dimethylsulfamide as precursor for N-nitrosodimethylamine (NDMA) formation upon ozonation and its fate during drinking water treatment. *Environ. Sci. Technol.* 42 (17), 6340–6346.
- Schollée, J., Schymanski, E., Avak, S., Loos, M., Hollender, J., 2015. Prioritizing unknown transformation products from biologically-treated wastewater using high-resolution mass spectrometry, multivariate statistics, and metabolic logic. *Anal. Chem.* 87 (24), 12121–12129.
- Schollée, J.E., 2018. TFAalyzeR, version 1.0.1, Analysis of tracefinder target screening. Zenodo doi:10.5281/zenodo.3234748.
- Schollée, J.E., Bourgin, M., von Gunten, U., McArdell, C.S., Hollender, J., 2018. Non-target screening to trace ozonation transformation products in a wastewater treatment train including different post-treatments. *Water Res.* 142, 267–278.
- Schollée, J.E., Hollender, J. and McArdell, C.S., in preparation. Characterizing advanced treatment with ozone and activated carbon at multiple wastewater treatment plants through non-target screening.
- Schulz, M., Löffler, D., Wagner, M., Ternes, T.A., 2008. Transformation of the X-ray contrast medium iopromide in soil and biological wastewater treatment. *Environ. Sci. Technol.* 42 (19), 7207–7217.
- Schulze, S., Zahn, D., Montes, R., Rodil, R., Quintana, J.B., Knepper, T.P., Reemtsma, T., Berger, U., 2019. Occurrence of emerging persistent and mobile organic contaminants in European water samples. *Water Res.* 153, 80–90.
- Schymanski, E., Singer, H., Longrée, P., Loos, M., Ruff, M., Stravs, M., Ripollés Vidal, C., Hollender, J., 2014a. Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry. *Environ. Sci. Technol.* 48 (3), 1811–1818.
- Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014b. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ. Sci. Technol.* 48 (4), 2097–2098.
- Schymanski, E.L., 2019. MetFrag Local CSV: compTox (7 March 2019 release) Wastewater MetaData File (Version WWMetaData\_4Oct2019) [Data set], Zenodo. 10.5281/zenodo.3472781.
- Schymanski, E.L., 2020. ReSOLUTION: SOLUTIONS for High ReSOLUTION Mass Spectrometry. R package. <https://rdrr.io/github/schymane/ReSOLUTION/>.
- Seiwert, B., Klockner, P., Wagner, S., Reemtsma, T., 2020. Source-related smart suspect screening in the aqueous environment: search for tire-derived persistent and mobile trace organic contaminants in surface waters. *Anal. Bioanal. Chem.*
- Stravs, M.A., Schymanski, E.L., Singer, H.P., Hollender, J., 2013. Automatic recalibration and processing of tandem mass spectra using formula annotation. *J. Mass Spectrom.* 48 (1), 89–99.
- Stuart, M.E., Lapworth, D.J., Thomas, J., Edwards, L., 2014. Fingerprinting groundwater pollution in catchments with contrasting contaminant sources using microorganic compounds. *Sci. Total Environ.* 468–469, 564–577.
- Ter Laak, T.L., Puijker, L.M., van Leerdam, J.A., Raat, K.J., Kolkman, A., de Voogt, P., van Wezel, A.P., 2012. Broad target chemical screening approach used as tool for rapid assessment of groundwater quality. *Sci. Total Environ.* 427–428, 308–313.
- Tian, Z., Peter, K.T., Gipe, A.D., Zhao, H., Hou, F., Wark, D.A., Khangaonkar, T., Kolodziej, E.P., James, C.A., 2020. Suspect and Nontarget Screening for Contaminants of Emerging Concern in an Urban Estuary. *Environ. Sci. Technol.* 54 (2), 889–901.
- Trcek, B., Zigon, D., Zidar, V.K., Auersperger, P., 2018. The fate of benzotriazole pollutants in an urban oxic intergranular aquifer. *Water Res.* 131, 264–273.
- Wang, Z., Walker, G.W., Muir, D.C.G., Nagatani-Yoshida, K., 2020. Toward a global understanding of chemical pollution: a first comprehensive analysis of national and regional chemical inventories. *Environ. Sci. Technol.* 54 (5), 2575–2584.
- Weber, W.H., Seitz, W., Schulz, W., 2007. Detection of the metabolites desphenyl-chloridazone and methyl-desphenyl-chloridazone in surface, ground and drinking water (Nachweis der Metaboliten Desphenyl-Chloridazon und Methyl-desphenyl-chloridazon in Oberflächen-, Grund- und Trinkwasser). *Vom Wasser* 105, 7–14.
- Williams, A.J., Grulke, C.M., Edwards, J., McEachran, A.D., Mansouri, K., Baker, N.C., Patlewicz, G., Shah, I., Wambaugh, J.F., Judson, R.S., Richard, A.M., 2017. The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *J. Cheminform.* 9 (1), 61.
- Wolf, L., Zwiener, C., Zemmann, M., 2012. Tracking artificial sweeteners and pharmaceuticals introduced into urban groundwater by leaking sewer networks. *Sci. Total Environ.* 430, 8–19.
- Ying, W.-C., Bonk, R.R., Lloyd, V.J. and Sojka, S.A., 1986. Biological treatment of a landfill leachate in sequencing batch reactors. 5 (1), 41–50.
- Zahn, D., Fromel, T., Knepper, T.P., 2016. Halogenated methanesulfonic acids: a new class of organic micropollutants in the water cycle. *Water Res.* 101, 292–299.
- Zahn, D., Mucha, P., Zilles, V., Touffet, A., Gallard, H., Knepper, T.P., Fromel, T., 2019. Identification of potentially mobile and persistent transformation products of REACH-registered chemicals and their occurrence in surface waters. *Water Res.* 150, 86–96.