

Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis

Emma L. Schymanski¹ · Heinz P. Singer¹ · Jaroslav Slobodnik² · Ildiko M. Ipolyi² · Peter Oswald² · Martin Krauss³ · Tobias Schulze³ · Peter Haglund⁴ · Thomas Letzel⁵ · Sylvia Grosse⁵ · Nikolaos S. Thomaidis⁶ · Anna Bletsou⁶ · Christian Zwiener⁷ · María Ibáñez⁸ · Tania Portolés⁸ · Ronald de Boer⁹ · Malcolm J. Reid¹⁰ · Matthias Onghena¹¹ · Uwe Kunkel¹² · Wolfgang Schulz¹³ · Amélie Guillon¹⁴ · Naïke Noyon¹⁴ · Gaëla Leroy¹⁵ · Philippe Bados¹⁶ · Sara Bogialli¹⁷ · Draženka Stipaničev¹⁸ · Pawel Rostkowski¹⁹ · Juliane Hollender^{1,20}

Received: 30 January 2015 / Revised: 2 April 2015 / Accepted: 7 April 2015 / Published online: 15 May 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract In this article, a dataset from a collaborative non-target screening trial organised by the NORMAN Association is used to review the state-of-the-art and discuss future perspectives of non-target screening using high-resolution mass

spectrometry in water analysis. A total of 18 institutes from 12 European countries analysed an extract of the same water sample collected from the River Danube with either one or both of liquid and gas chromatography coupled with

Published in the topical collection *High-Resolution Mass Spectrometry in Food and Environmental Analysis* with guest editor Aldo Laganà.

Electronic supplementary material The online version of this article (doi:10.1007/s00216-015-8681-7) contains supplementary material, which is available to authorized users.

✉ Emma L. Schymanski
emma.schymanski@eawag.ch

✉ Juliane Hollender
juliane.hollender@eawag.ch

¹ Eawag: Swiss Federal Institute for Aquatic Science and Technology, Überlandstrasse 133, 8600 Dübendorf, Switzerland

² Environmental Institute, s.r.o., Okružná 784/42, 972 41 Koš, Slovak Republic

³ Helmholtz Centre for Environmental Research - UFZ, Permoserstraße 15, 04318 Leipzig, Germany

⁴ Umeå University, Linnaeus väg 6, 90187 Umeå, Sweden

⁵ Chair of Urban Water Systems Engineering, Technische Universität München, Am Coulombwall 8, 85748 Garching, Germany

⁶ Department of Chemistry, University of Athens, Panepistimiopolis Zografou, 157 71 Athens, Greece

⁷ Environmental Analytical Chemistry, Eberhard Karls University of Tübingen, Hoelderlinstr. 12, 72074 Tübingen, Germany

⁸ Research Institute for Pesticides and Water, University Jaume I, Avda. Sos Baynat s/n, 12071 Castellón de la Plana, Spain

⁹ Ministry of Infrastructure and the Environment (Rijkswaterstaat), Zuiderwagenplein 2, 8224 AD Lelystad, Netherlands

¹⁰ Norwegian Institute for Water Research (NIVA), Gaustadalleen 21, 0349 Oslo, Norway

¹¹ Toxicological Center, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Antwerpen, Belgium

¹² German Federal Institute of Hydrology (BfG), Am Mainzer Tor 1, 56068 Koblenz, Germany

¹³ Betriebs- und Forschungslaboratorium, Zweckverband Landeswasserversorgung, Am Spitzigen Berg 1, 89129 Langenau, Germany

¹⁴ Suez Environnement CIRSEE, 38 rue du président Wilson, 78230 Le Pecq, France

¹⁵ Veolia Research and Innovation (VERI), 1 Place de Turenne, 94 417 Saint Maurice Cedex, France

¹⁶ UR MALY Freshwater Systems, Ecology and Pollutions, Irstea, Centre de Lyon-Villeurbanne, 5 rue de la Doua-CS 70077, 69626 Villeurbanne Cedex, France

¹⁷ Department of Chemical Sciences, University of Padua, Via Marzolo, 1, 35131 Padova, Italy

¹⁸ Croatian Waters, Ulica grada Vukovara 220, 10000 Zagreb, Croatia

¹⁹ NILU - Norwegian Institute for Air Research, Instituttveien 18, 2007 Kjeller, Norway

²⁰ Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zurich, Switzerland

mass spectrometry detection. This article focuses mainly on the use of high resolution screening techniques with target, suspect, and non-target workflows to identify substances in environmental samples. Specific examples are given to emphasise major challenges including isobaric and co-eluting substances, dependence on target and suspect lists, formula assignment, the use of retention information, and the confidence of identification. Approaches and methods applicable to unit resolution data are also discussed. Although most substances were identified using high resolution data with target and suspect-screening approaches, some participants proposed tentative non-target identifications. This comprehensive dataset revealed that non-target analytical techniques are already substantially harmonised between the participants, but the data processing remains time-consuming. Although the objective of a “fully-automated identification workflow” remains elusive in the short term, important steps in this direction have been taken, exemplified by the growing popularity of suspect screening approaches. Major recommendations to improve non-target screening include better integration and connection of desired features into software packages, the exchange of target and suspect lists, and the contribution of more spectra from standard substances into (openly accessible) databases.

Keywords Non-target screening · High resolution mass spectrometry · LC–MS · GC–MS · Suspect screening · Surface water

Introduction

The evolution of accurate mass (AM) high resolution mass spectrometry (HRMS) coupled to gas or liquid chromatography (GC or LC, respectively) has initiated a new trend in analytical data processing in recent years. Targeted analytical methods are now often complemented with non-target or untargeted data acquisition and screening methods, where tandem mass spectrometry (MS–MS) is used to obtain fragmentation information to support identification. Earlier articles on the use of AM HRMS (hereafter shortened to HRMS, where the accurate mass is implied) in the environmental context (e.g. Refs. [1–4]) mention the three main approaches towards substance identification in non-target HRMS analysis, which are summarised here with some additional nuances:

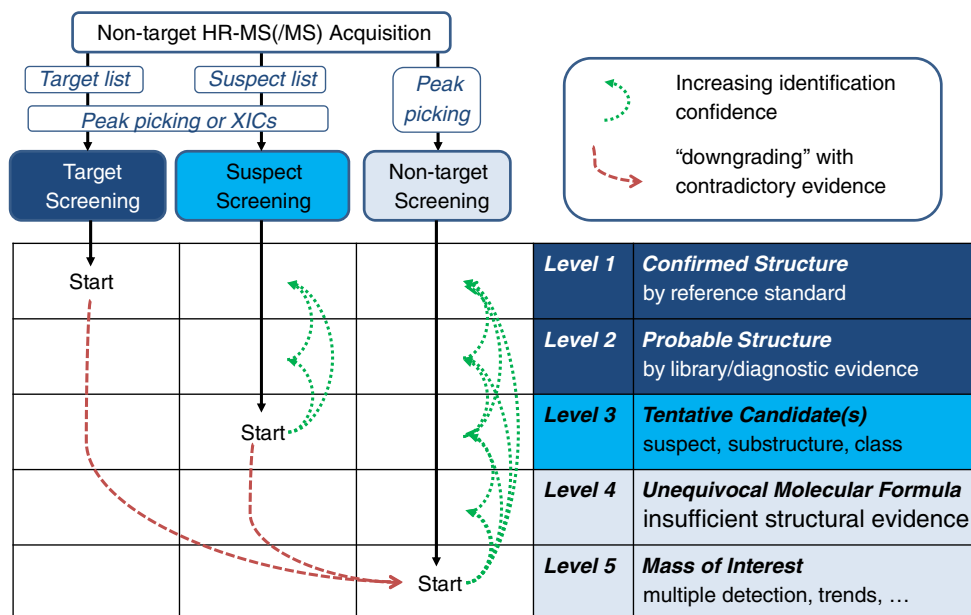
1. Target screening involves a reference standard measured in-house under the same analytical conditions such that retention time (RT), HRMS, and, where possible, (HR)MS–MS information is available for identification and confirmation. Quantitative target results should be distinguished from semi-quantitative results by using the term “quantitative target analysis/screening”.

2. Suspect screening is performed when prior information (from a variety of sources, discussed in greater detail below) indicates that a given structure may be present in the sample. Thus, although no reference standard is available, the exact mass and isotope pattern calculated from the molecular formula plus or minus the expected adduct(s) of the suspect substance can be used to screen for this substance in the sample.
3. Non-target screening involves all remaining components detected in a sample where no prior information is available. Because no structural information is available in advance, a full non-target identification starting from the exact mass, isotope, adduct, and fragmentation information needs to be performed.

The data analysis for target and suspect substances in non-target acquisition data can be performed in two main ways (or a combination of both). Traditionally in target analysis the presence or absence of each substance is determined individually using the extracted ion chromatogram (XIC). However, the evolution of non-target methods means that often a screening for target and/or suspect compounds is performed after peak detection with a suitable algorithm, such that the exact masses of the appropriate adduct of the target or suspect are searched within a given mass and (for targets) RT error. Whereas the first approach treats targets and suspects preferentially (i.e., they can be detected in cases where the peak is of insufficient quality for automated peak detection), in the latter case the target and suspect compounds are effectively a subset of all the “non-target components”. Here, the term “component” is used to refer to the group of exact masses (i.e. adduct and isotopologues) associated with one compound. Irrespective of the extraction technique, evidence from the measurement data is needed to confirm the identification, including the isotope pattern, presence of additional adducts, RT, fragmentation information, and other experimental evidence (e.g. presence of related substances, time trends). Several papers discuss this in greater detail (e.g. Refs. [1, 3–5]). Although the concept of identification points (IPs) given in EU Guideline 2002/657/EC [6] can be used to represent the evidence available for an identification where a reference standard (and thus a RT) is available, these are in urgent need of revision because they were released before HRMS(–MS) became as prominent as it is today. The concept of “identification levels” in HRMS analysis has been discussed recently in the environmental community and beyond (e.g. Refs. [7–9]), building on earlier attempts by the Metabolomics Society [10] to deal with the varying levels of confidence in identifications resulting from these three strategies for identifying substances using non-target acquisition data.

The concepts of identification strategy and confidence are merged in Fig. 1, showing that target, suspect, and non-target compounds start by definition at Levels 1 (reference standard

Fig. 1 Matrix of identification approach versus identification confidence



available), 3 (tentative candidate(s)), and 5 (no information), respectively. If sufficient MS (exact mass, isotope, adduct), MS–MS (i.e., fragmentation), and experimental information (e.g. retention behaviour, presence of related substances) is available, suspect and non-target components can gain in confidence through to Level 2 (library match and/or diagnostic fragments) and even Level 1 after purchase of the corresponding standard for identifications (green arrows in Fig. 1). Compounds with a confirmed identity then become target compounds in future investigations. However, should the evidence of the sample and of the reference standard (target) or tentative candidate (suspect) not match, then the component associated with the target or suspect should become a “non-target of interest” (Level 5)—see red arrows. If the HRMS and RT information matches a target compound but HRMS–MS is not available because of lack of intensity, this is still regarded as a target identification, but should be reported with fewer identification points (e.g. 2 IP, as opposed to 4.5 IP where HRMS–MS from an isolated precursor is available [6]).

It is important to note that this elaborate combination of strategy, confidence, and evidence is necessary with soft ionisation (SI) HRMS–MS analysis at this stage for one main reason: the lack of comprehensive spectral libraries for soft ionisation techniques, in contrast with the comparatively comprehensive GC–electron ionisation (EI) MS libraries available [11, 12] with over 200,000 substances. The reasons for this are varied, including the lack of reproducibility between SI instruments and different settings and the relative newness of the technique, and are discussed in greater detail elsewhere (e.g. Refs. [13–15]). In consequence of this difference in library resources, the identification of “unknown” environmental substances of interest measured with GC–EI

MS can often be performed with a spectral library, whereas substances measured with alternative ionisation techniques usually require the parallel approaches of target, suspect, and non-target screening. This is not just limited to LC-based techniques; the dependence of strategy on ionisation technique is revealed for the environmental context in recent work comparing quadrupole time-of-flight (QTOF) GC–MS using EI and atmospheric pressure chemical ionisation (APCI) [16] and GC and ultra-high performance LC (UPLC) coupled to QTOFMS for a universal screening approach [17].

Several strategies for the selection of candidates for suspect and non-target screening of environmental substances using HRMS–MS information have been developed and put into practice recently. Suspect screening has been performed using predicted transformation products (TPs) [18–20], registered pesticides and their TPs in Switzerland [21], surfactants in wastewater [3] and fracking fluids [22], site-specific chemicals [5], and high-consumption pesticides and pharmaceuticals in sediments [23]. Other lists, including the potentially persistent and bioaccumulative organics in commerce [24], pharmaceuticals [25], and impurities, by-products, and transformation products [26] from Howard and Muir, can also be used in suspect screening. For non-target identification, candidate discrimination criteria used include retention time information, also represented as partitioning coefficients [19, 23] or the chromatographic hydrophobicity index (CHI) [5, 27], in silico fragmentation prediction with bond dissociation methods, for example MetFrag [28], or rule-based approaches including Mass Frontier [29], and also the number of references available for a substance in such databases as ChemSpider or SciFinder as a measure of relative importance in the environment

[30, 31]. It is clear from the detailed data processing schemes in recent non-target papers (e.g. Figs. 1 and 3 in [5], Fig. 1 in [21]) that non-target screening of environmental samples is becoming increasingly complex.

In response to this trend of increasing complexity and the need articulated by members to compare and harmonise non-target screening methods in Europe, the NORMAN Association (www.norman-network.net) instigated a collaborative non-target screening trial in 2013 on a sample extract from the River Danube. Each participating institute was requested to analyse the test sample using established MS techniques in their laboratory and declare (1) how many substances were present in the sample and (2) how many of these could be provisionally identified using target, suspect, and non-target screening approaches. Analytical standards for the calculation of retention index information were also provided. After the trial, a workshop was held with all participants to develop agreements on harmonised terminology, workflows, and reporting formats. This dataset forms a unique opportunity to review the state-of-the-art of non-target acquisition and identification techniques using high resolution mass spectrometry with a comparable dataset from several environmental institutes in Europe and to comment on current and future trends.

Materials and methods

Sampling and trial participation

The sample used in the collaborative trial was collected from location JDS57, downstream of Ruse/Giurgiu (RO/BG; rkm 488; coordinates N43.890150, E26.017067) on September 18, 2013 as part of the Third Joint Danube Survey, organised by the International Commission for the Protection of the Danube River (ICPDR) [33, 34]. The sample preparation included a large-volume solid-phase extraction (LVSPE) of 1000 L water [35]. Briefly, the sampler cartridge was filled with 160 g Macherey Nagel Chromabond® HR-X (neutral resin) and 100 g each of Chromabond® HR-XAW (anionic) and HR-XCW (cationic exchange resin). The resins were extracted with 500 mL each of ethyl acetate and methanol (HR-X), and 500 mL methanol with 2 % 7 mol L⁻¹ ammonia in methanol (HR-XAW) or 500 mL methanol with 1 % formic acid (HR-XCW). The extracts were then combined, neutralised, filtered (Whatman GF/F), and reduced to a final volume of 1 L using rotary evaporation. Aliquots of 1.5 mL, equivalent to 1.5 L river water, were transferred into vials and evaporated to dryness under nitrogen. These were sent to each participant with a laboratory blank, which was created via circular pumping of 5 L LC-MS-grade water through the LVSPE to mimic

leaching from 1000 L water passing through the LVSPE. Sample stability (at least three time points over 1.5 months) and homogeneity testing (three replicates) was performed using 20 and 50 substances for GC-MS and LC-MS, respectively, to confirm the suitability of the sample for the trial. The substances are listed with the results in the Electronic Supplementary Material (ESM), Tables S1 and S2.

The samples were dispatched on December 9, 2013 with standard mixtures for use in calculating retention-index information: C₁₀ to C₂₅ alkane standards for GC-MS techniques and 10 substances for LC-MS techniques. The LC-MS standard mixture (the ChemSpider [32] identity number and the logarithm of the octanol-water partitioning coefficient (log P) calculated with ChemAxon [36] are given in brackets) consisted of metformin (3949; -1.36), chloridazon (14790; 1.11), carbetamide (133997; 1.65), monuron (8470; 1.93), metobromuron (17276; 2.24), chlorbromuron (24141; 2.85), metconazole (77764; 3.59), diazinon (2909; 4.19), quinoxifen (2635909; 4.98), and fenofibrate (3222; 5.28). All participants were requested to measure these standards with the sample and report the results by March 15, 2014. As such, this dataset formed a test set for retention information comparison, data storage, and re-processing of raw mass spectral data for retrospective analysis. Of the 26 institutes from 15 countries who received samples, 19 submitted results; seven institutes for GC-MS and 17 institutes for LC-MS (five institutes submitted for both). The participants ranged from institutes performing non-target methods for the first time through to experienced research groups. The overall time committed by the participants varied from two days to six weeks. Each institute that submitted results is represented by a number.

LC-HRMS

An overview of the LC-HRMS(-MS) methods used is given in Tables 1 and 2.

One of the 17 LC participants used a serial coupling of zwitterionic hydrophilic interaction (HILIC) and reversed-phase (RP) chromatography (LCLC; HILIC-C₁₈) [37], and the remainder used C₁₈ reversed-phase columns (one C₈) with either high-performance liquid chromatography (HPLC) or long ultra-HPLC (UHPLC) runs. The solvent was usually water-methanol or water-acetonitrile, with no or varied modifiers (formic acid or ammonium acetate, formate or hydroxide). Between 2 and 100 µL of the sample was injected. Electrospray ionisation (ESI) only was used, with different collision-induced dissociation (CID) or higher-energy CID (HCD) energies, and participants who measured in both positive and negative mode did this in separate runs. All-ion approaches (fragmentation without precursor selection) were used by some participants. The scan ranges started between

Table 1 Summary of liquid chromatography methods used by LC participants

Participant	Instrument and model	Column	Dimensions (mm×mm, μm)	Solvent	Injection volume (μL)	Flow; run time (mL min ⁻¹); (min)
1	AB Sciex TripleTOF 5600	Phenomenex Luna C ₁₈ (2)	2.0×150; 3.0	H ₂ O/ACN (FA)	10	0.2; 33
2	Thermo Q Exactive Orbitrap	Waters Xbridge C ₁₈	2.1×50; 3.5	H ₂ O/MeOH (FA)	20	0.2; 30
3	Agilent ToF 6230	ZIC-HILIC, Poroshell C ₁₈	2.1×150; 5.0 3.0×50; 2.7	H ₂ O/ACN (NH ₄ ac)	10	0.05-0.4; 58
4	Agilent 6550 iFunnel Q-TOF LC-MS	Waters Acquity UPLC HSS T3 and ZORBAX ECLIPSE PLUS C ₁₈ RRHD	2.1×150; 1.8 2.1×100; 1.7	H ₂ O/MeOH (NH ₄ ac)	100	0.4; 20
5	Waters Micromass Xevo G2 QTOF	Acquity UPLC BEH C ₁₈	2.1×100; 1.7	H ₂ O/MeOH (FA)	50	0.3; 18
6	AB Sciex TripleTOF 5600	Phenomenex AQ C ₁₈ (Guard), Agilent Zorbax Eclipse Plus C ₁₈	2.0×4.0; n/a 2.1×150; 3.5	H ₂ O/ACN (FA)	5	0.3; 37
7	Waters Xevo G2-S QTOF	Waters Acquity HSS C ₁₈ , BEH C ₁₈	2.1×150; 1.8 2.1×100; 1.7	H ₂ O/ACN (NH ₄ fa, FA); H ₂ O/MeOH (NH ₄ ac)	5	0.4-0.45; 15
8	Thermo LTQ Orbitrap Discovery	Waters Atlantis HSS T3	2.1×150; 3.5	H ₂ O/MeOH (FA)	10	0.2; 19
9	Thermo LTQ Orbitrap XL	Phenomenex Kinex C ₁₈	3.0×100; 2.6	H ₂ O/MeOH (FA)	10	0.2; 41
10	Thermo LTQ Orbitrap Discovery	Thermo Hypersil Gold	2.1×100; 3.0	H ₂ O/MeOH (FA)	20	0.2; 35
11	Agilent 6530 AM QTOF-LC-MS	Eclipse Plus C ₈	2.1×150; 3.5	H ₂ O/ACN (NH ₄ ac, FA)	10	0.25; 29
12	Agilent AM Q-TOF LC-MS 6520	Phenomenex Kinex C ₁₈	2.1×100; 1.7 2.1×100; 2.6	H ₂ O/ACN (FA, NH ₄ OH)	40	0.3; 38
13	Agilent 6550 iFunnel Q-TOF LC-MS	Zorbax Extend C ₁₈	2.1×50; 1.8	H ₂ O/ACN (FA)	2	0.5; 46
14 ^a	Bruker maxis impact	Dionex Acclaim	2.1×100; 2.2	H ₂ O/MeOH (NH ₄ fa, FA, NH ₄ ac)	10	0.2-0.48; 19
15	Thermo Q Exactive Orbitrap	Hypersil Gold aQ	2.1×100; 1.9	H ₂ O/MeOH (none, FA, NH ₄ fa)	5	0.3; 35
16	Waters Xevo G2-S Q-TOF	Waters C ₁₈ BEH	2.1×100; 1.7	H ₂ O/MeOH (NH ₄ fa, pH 5)	3	0.45; 13
17	Agilent 6550 QTOF LC-MS	Agilent Poroshell HPH C ₁₈ (+Guard)	2.1×5; 2.7 2.1×150; 2.7	H ₂ O/MeOH (FA pos only)	5	0.3; 35

ACN=acetonitrile, H₂O=water, FA=formic acid, MeOH=methanol, NH₄ac=ammonium acetate, NH₄OH=ammonium hydroxide, NH₄fa=ammonium formate. Oven temperatures ranged from 22 to 50 °C. n/a=not available

^a Quantitative target analysis was performed using a Thermo TSQ Access Triple-Quad instrument

Table 2 Summary of mass spectrometry and data processing procedures used by LC participants

Participant	Scan range	(reference <i>m/z</i>)	Ionisation	Fragment method	Target software	Suspect, nontarget procedure
1	100–1200	30,000 (400)	ESI+ only	CID, 40	Peakview, Multiquant	PeakView, Markerview. MS, isotopes, RT, MS–MS, manual peak check
2	100–1000	140,000 (200)	ESI±	HCD various, 50 (DD); merged HCD 30–70 (DIA)	Xcalibur, Trace Finder, nontarget	ExactFinder 2.0, nontarget, MetFusion, internal lists. MS, RT, MS–MS, library, prediction, interpretation
3	100–1700	12,000 (1000)	ESI±	In-source fragmentation at 100 V	MassHunter Qual (B.06.00)	MassHunter Qual, Profinder (B.06.00). MS, RT, internal list, MS–MS, library, STOFF-IDENT DB, pred, isotopes
4	100–1000	56,014 (922)	ESI±	CID, 10, 20, 40, All Ions	MassHunter/Quant	MassHunter Quant and Qual (v B.06.00, Build 6.0.633.0). FindByFormula, MassProfiler, MS–MS, isotope, PCDDL
5	50–1000	20,000 (556)	ESI±	CID, 15–40 Ramp, MS ^E	Chromalynx XS	Chromalynx XS, home-made database, MassBank. MS, MS ^E , prediction
6	100–1200	31,000 (327)	ESI±	CID, 40; merged 25, 40, 55, DDA	MasterView	STOFF-IDENT, DAIOS, internal list of substances. MS, (RT, MS–MS)
7	50–700, 50–1200	22,500 (956)	ESI±	CID, 10–55 Ramp	Unifi	Unifi. RT, MS, MS–MS, Prediction
8	50–1000	30,000 (400)	ESI+ only	Not used	–	mzMine, ChemSpider
9	100–1000	100,000 (400)	ESI±	HCD, variable	Xcalibur	Xcalibur. MS, RT, MS–MS, RT(CHD), fragment (MetFrag) prediction, manual interp.
10	80–1500	30,000 (400)	ESI±	CID, 35	Exact Finder	Exact Finder. MS, RT, MS–MS
11	50–1300	8000 (600)	ESI±	CID, 10, 20	MassHunter	Pragst library. MS, MS–MS
12	50–2000	18,000 (311)	ESI±	CID, 15, 40	MassHunter Qual B04.00	MetIn database (pesticide, forensic). MS, MS–MS
13	50–1200	20,000 (622)	ESI+ only	CID, 20, 40	MassHunter B6.0	ForensicsTox, Pesticides, MassBank, DAIOS. MS, RT, MS–MS, prediction
14 ^a	50–1000	40,000 (431)	ESI±	CID, 25	LCQuan, Target Analysis, Data Analysis	Target and Data Analysis. MS, isotopes, MS–MS, RT (KNN-GA-SVM), fragment (MetFrag, SmartFormula3D) prediction
15	70–1000	70,000 (200)	ESI±	HCD, 50	Trace Finder	Trace Finder, Sieve, in-house library, Thermo library (with and without RT). MS, RT, MS–MS, library & prediction
16	50–1000	20,000 (556)	ESI+ only	CID, 10–45 Ramp, MS ^E	MassLynx, Chromalynx	MassLynx/Chromalynx, in-house library. MS, RT, MS–MS
17	50–1500	23,000 (119)	ESI±	CID, variable	MassHunter B06 SPI	MassHunter. MS, RT, MS–MS, prediction. Agilent ForensicTox library

ESI=electrospray ionisation, CID=collision-induced dissociation, HCD=higher energy CID, DD=data-dependent acquisition, DIA, Allions, MS^E=fragmentation without precursor isolation, RT=retention time, DB=database(s)

^a Quantitative target analysis was performed using a Thermo TSQ Access Triple-Quad instrument

m/z 50 and 100, ending between m/z 1000 and 2000, and the resolution ranged from \sim 18,000 (at m/z 300) to 140,000 (at m/z 200). Five instruments from the Orbitrap family were used, and 12 were from the TOF family (different vendors). The data processing for the target compounds was generally performed using vendor software, while the peaks that were not assigned as targets were identified through suspect and non-target screening, performed with a much wider variety of software and methods. While many participants also quantified target substances, this is not the focus of this article.

Many different information sources were used by participants (shown in Table 3) including open and vendor libraries with mass spectra available, extensive compound databases, selective compound databases containing water-body relevant substances (STOFF-IDENT and DAIOS), as well as suspect lists from literature. These were generally supported or rejected via a combination of accurate mass (including isotope and/or adduct information), RT, MS–MS, library match, and predicted properties. Approaches to consider retention time information included the log P method using standards provided to each participant (Grosse et al., in prep.), often in combination with STOFF-IDENT [41], the CHI approach [27], and a QSPR approach (Aalizadeh et al., in prep) using support vector machines (SVM) and k-nearest neighbours (kNN). Software used for fragmentation prediction included MetFrag [28], MetFusion [55], Mass Frontier [29], MassFragment in UNIFI and MassLynx (Waters [56]), SmartFormula 3D (Bruker Daltronics [57]), and Molecular Structure correlator [58, 59].

GC–MS

A summary of the GC–MS methods is given in the ESM, Tables S3 and S4. One participant submitted high resolution data, and three of seven GC participants used GC \times GC. Two used chemical ionisation (CI) methods in addition to EI: APCI and positive and negative CI. Four participants chose a large volume injection to obtain sufficient signal intensity (see results) and the scan range started in the range m/z 35–50, ending in the range m/z 350–800. A range of different solvents were used. Target compound identification was usually performed using vendor software. Non-target identification was performed using the NIST database (Table 3) either with or without AMDIS [60], combined in most cases with the Kovats retention index.

Data reporting and comparison

Participants were requested to submit their results in a data collection template (DCT), a multi-tab spreadsheet, to ensure sufficient information was available for evaluation. Pre-treatment steps to increase comparability included reclassification of methods and reported identifications to be

consistent with the definition of target, suspect, and non-target provided in the introduction. Several substance identifiers were used to perform the comparison, including the SMILES code [61], Chemical Abstract Services (CAS) number, names, molecular formulas, and the InChI Key [62]. OpenBabel [63] was used to perform conversions, and ChemSpider [32] and PubChem [40] were used to fill data gaps. Entries for the same target repeated multiple times per participant were merged to create one entry only, with multiple identifications for the same substance and peak as a target, suspect, and/or non-target selected according to the hierarchy target>suspect>non-target. However, multiple identities for the same peak (e.g. co-eluting targets) were not merged.

Results and discussion

Stability testing

Different batches of the sample used in the trial were subjected to stability and homogeneity testing to determine whether the sample was suitable for such a trial and provide a baseline for comparison; tables of results are provided in the ESM, Table S1 (GC) and S2 (LC). For GC–MS, 20 substances were selected to cover the full run time and polarity range. Homogeneity testing was performed on Lots 15, 33, and 50 (three replicates of each) with the coefficient of variation usually below 1 % and with a maximum of 7 %. For the stability (tested on Days 1, 10, and 30), the highest relative concentration was 1.033 on Day 30. For LC–MS, 50 substances were selected and while the homogeneity (three aliquots) had coefficients of variation in the range 1–16 %, very similar values were obtained for reproducibility testing (three injections). Stability was tested on Days 1, 7, 15, and 48 and revealed no degradation of any compound selected. Generally an increasing trend to higher factors (Day X/Day 1) was observed. The rather high variation of the stability factors (in the range 0.5–2) was caused predominantly by measurement sensitivity and was regarded as sufficient for the trial because the main focus here is not a comparison of quantitative results.

LC–HRMS overview

In total, 17 institutes submitted LC–HRMS results and 15 of these reported target substances (for which the identity was confirmed with a reference standard). The breakdown of targets, suspects, and non-targets is given in Table 4, including additional categories (whether the target was quantified or not, isomer mix, identified non-target, and formula only) to reflect the identification confidence and how the substances were identified. Participants 6 and 13, with large numbers of assigned formulas, were able to do this automatically through their software. The distribution reflects the target and suspect

Table 3 Summary of information sources used by collaborative trial participants

Database or library name	State as used during the trial		Current state
	Total compounds	Compounds with spectra	Compounds at March 2015
ChemSpider [32]	32 million	–	32 million
DAIOS [38, 39]	1,404	>1,000 ^a	1,404
PubChem [40]	63,105,228	–	68,479,719
STOFF-IDENT [41]	7,864 ^b	–	7,864
MassBank MS/MS [42–44]	–	3,350	3,350
mzCloud [45]	–	1,956	2,510
NIST EI-MS [11, 46]	–	212,961 ^c	242,477
NIST MS/MS [11, 46]	–	4,628	8,171
Wiley Registry of Mass Spectral Data (EI-MS) [47]	–	289,000 [12]	638,000
Agilent Broecker, Herre & Pragst Toxicology/Forensics ^f [48, 49]	8,998 ^c	3,497	8,998
Agilent Pesticide Library LC/Q-TOF MS/MS ^f [50]	1,664	~700 ^c	1,664
Agilent Pesticide Library GC/Q-TOF EI-MS ^f	750	750	750
Agilent METLIN Synthetic Substance Library ^g	64,092 ^c	~10,000 ^c	64,092
Agilent METLIN Scripps Online Database ^{f,g} [51, 52]	83,135	12,171 ^c	240,566
Agilent Veterinary Drug Library ^f	1,684	770	1,684
Bruker ToxScreen (incl. Pesticide Screener) ^g [53]	–	704 ^{ad}	1753
Sciex / AB Sciex LC/MS/MS Meta Library ^g [54]	–	2,381 ^c	2,381
Thermo Environmental Food Safety (EFS) with/without retention time (RT) ^g	–	447 ^p ; 278 ⁿ ; 454 ^{dp} ; 90 ^{dn}	732
Thermo toxicology ^g	–	618 ^p ; 36 ⁿ	654
Waters database with RT ^g	–	730 ^{de}	730
In-house Libraries without spectra (two participants)	2,000; 1,600 [17]	–	2,000; 1,600
In-house Libraries with spectra (two participants)	–	526 ^d ; 63 ^d	526; 63
In-house Libraries with spectra for some substances	2,200 ^d	835 ^{ad}	2,200
–	7,815	1500 ^{ap} ; 500 ^{an}	7,815
–	3,000	350 ^d	3,000
Surfactant List [3]	394	–	394

^a Indicates that fragment information but not spectra with intensities were available (e.g. Q1, Q2, Q3, ...)

^b Retention time information and measured standards used to select best matches by log D values

^c Maximum number of substances; participants used different versions with slightly varying numbers

^d Indicates retention time information included, measured on the same system or an identical set-up

^e Retention information transformed using 40 standards

^f Confirmed by Agilent Technologies (Faye, pers. comm.)

^g Numbers provided by participants

^p Positive ionisation mode

ⁿ Negative ionisation mode

lists and the databases available at the different institutes, and reveals the huge amount of information that can be gained by increasing the exchange of knowledge between institutes. The time invested by the participants (two days to six weeks of the total three months given for the trial) is also reflected in these results.

In total 625 target results were reported, corresponding to 347 unique compounds, and the majority (451 results) were quantified. A total of 631 suspect and tentatively identified non-target substances were reported, 553 of which had a

defined structure associated with the peak (the remaining substances were reported as a substance class or unspecified isomer). Altogether 649 unique compounds were identified as target, suspect or non-target. The most frequently reported included (targets+suspects) carbamazepine (13+2), atrazine (9+4), sulfamethoxazole (11+1), DEET (6+6), metformin (10+2), terbutylazine (6+6), caffeine (7+5), atrazine-desethyl (8+3), and tramadol (7+3). Metoprolol, terbutylazine-desethyl, phenazone (also known as antipyrine), 4 and 5-methylbenzotriazole, atrazine-2-hydroxy,

Table 4 Results summary for LC-HRMS participants, broken down into categories and identification confidence [7]

Participant	Quantified target (Level 1)	Semi/non-quantified target (Level 1)	Suspect (Level 2–3)	Isomer mix, unspecified structure (Level 3)	Identified non-target (Level 2–3)	Formula only (Level 4)	Non-target peaks remaining (Level 5)
1	25	–	68	–	–	–	6,776
2	92	14	8	40	1	–	8,535
3	–	4	70	–	–	–	–
4	125	43	61	9	1	–	–
5	18	5	8	–	–	–	17
6	53	7	57	–	–	1,316	712
7	–	–	33	–	–	–	3,174
8	–	–	–	14	–	30	3
9	68	4	–	–	2	10	26
10	1	11	4	–	–	–	–
11	–	4	7	–	–	–	8
12	4	–	4	3	26	5	57
13	–	9	66	12	2	1,218	133
14	23	53	78	–	2	9	1,649
15	33	13	–	–	–	1	273
16	9	–	18	–	–	–	3
17	–	7	37	–	–	–	–

venlafaxine, isoproturon, telmisartan, perfluorooctane sulfonate (PFOS), and metolachlor were all reported nine times. One institute additionally reported an isobaric substance (propisochlor) as a suspect instead of metolachlor. A further seven and four substances were reported eight and seven times, respectively. These results indicate that the “well known compounds” are obviously more common target compounds and found more often, but the fact that many parent and TPs occur in this list indicates that many groups consider the formation of TPs already in their target screening. Terbutylazine was the most frequently reported of the isobaric compounds; see Fig. 2 and text below for more discussion on co-eluting isobars. Of the 24 substances specified as suspects for the trial (detected at this location in previous surveys [64] and given in the ESM, Table S5), all but three were detected at least once by the LC participants; one of these missing three was reported by a GC participant (bisphenol A), and the remaining two were not reported by any participant (naproxen and *N,N*-dimethyl-1-decanamide).

The participants were requested to report intensities in the sample and the blank using the DCT, and different approaches were used to consider the results of the blank sample. All but one participant reported intensities in the sample, and all but four reported intensities in the blank. Dibutylphthalate and/or isomers were reported eleven times by seven participants; three participants reported this in both the sample and blank, two in the sample only, one in the blank only, and one did not report either intensity. Of the other frequently reported substances, metformin and terbutylazine were reported in the blank once (several orders of magnitude lower than in the sample), carbamazepine, desethylatrazine, and atrazine twice (two or more orders of magnitude lower), DEET three times (one to two orders of magnitude lower), and caffeine four times (a factor of 2–600 lower). Most of the reported alkyl phosphate species (e.g. tributyl phosphate and triphenyl phosphate, see ESM Table S5) were found in both sample and blank, as were the polyethylene glycol suspects. The participants who applied a blank subtraction used intensity cut-offs at different ratios to determine whether the substance was regarded as present in the blank or not, and this affected the non-target peak lists provided. As an example, one participant subtracted all non-target peaks where the same peak was detected in the blank and the intensity was within two orders of magnitude of the intensity measured in the blank. Because only one each of blank and sample was provided, it was challenging to determine whether the intensities were sample-specific, especially given that sample enrichment was used. As such, both false positive and false negative results are possible.

Target and suspect screening was performed using both peak lists and XICs. Generally, the use of XICs gave more hits because this also captured target and/or suspect substances where the peak was not suitable for the peak picking

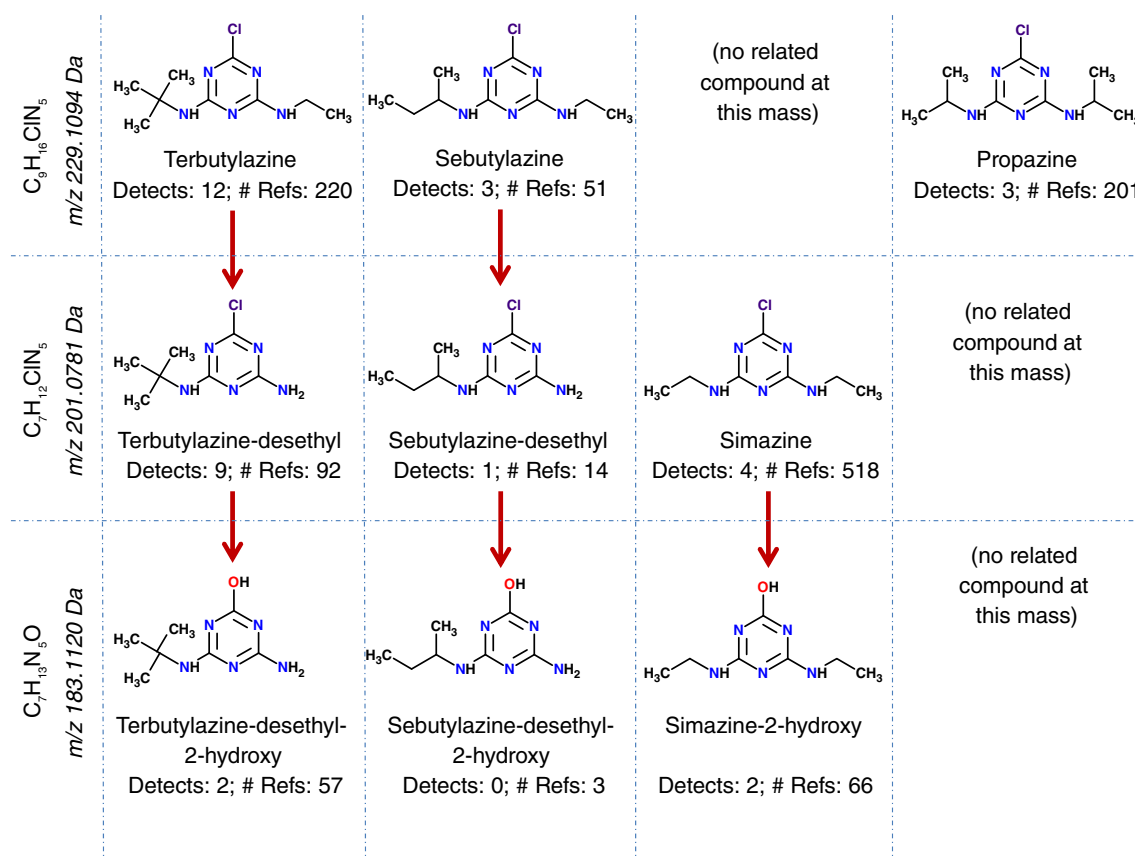


Fig. 2 Isobaric parents and transformation products in the collaborative trial sample, the number of times reported, and the number of references in ChemSpider [32]

algorithm used. However, the peak picker and the parameters used have a large effect on the quality of peak picking results and this would be interesting to investigate more closely in the future. In this trial, the strategy and results depended strongly on the software used by the participant, because most used vendor software to perform target and suspect screening. Many participants mentioned using manual or visual confirmation of the target and/or suspect results.

Analytical methods in LC–HRMS non-target screening

The results summarised in Table 4 reveal that non-target acquisition and screening has developed at an incredibly rapid pace in recent years, with many institutes now regularly screening several hundred compounds using multi-residue methods. The participants agreed that generic methods, based on C_{18} columns, generic gradients, and either HPLC or long-run UHPLC, were best for non-target approaches, to allow sufficient time for fragment acquisition. Many institutes also analysed the sample several times to gain additional information, running e.g. positive and negative separately rather than performing polarity switching, acquiring MS and all-ion MS–MS data to gain all fragments at once, running with and without internal standards, or acquiring data-dependent MS–MS

of target compounds first, followed by re-measurements to acquire MS–MS of suspects and non-targets of interest. This reveals that the number of compounds being investigated is almost on the edge of instrument capabilities, and there is a great interest in the development of appropriate data analysis strategies for all-ion MS–MS data.

One participant used HILIC combined with RP-LC to improve the separation of the highly polar substances, and the peak distribution shown in Fig. S1a in the ESM reveals the potential for greater separation of the polar substances that otherwise may elute in the dead volume of RP columns, shown in Fig. S1b in the ESM. The identified substances reported by this participant from the HILIC column (metformin at 16.85 min, melamine at 7.48 min) were also detected by other participants with the C_{18} column (12 and four times in total, respectively) in the range 0.3–2.9 min. This reveals that HILIC methods may help address the limitations of non-target screening of very polar compounds with RP columns alone, e.g. interferences observed with non-targets in the low RT range [3] and in regions of high matrix interference [23], but this has to be investigated further.

Although there was a great variety in the scan range used by the participants, all but one identified substance fell in the range m/z 100–900. The target compound piperazine ($[M+H]^+$ m/z 87.0921) was reported by one participant. The lowest

m/z suspects were five single-hit substances in the range m/z 100–114, whereas the first multiply-reported substance (benzotriazole) was detected at m/z $[M+H]^+$ 120.0556. The highest m/z target reported was iohexol (m/z 821.8879), and one suspect was reported above this, namely tilmicosin (m/z 869.5690). The conclusion to harmonise the non-target methods was to suggest a minimum scan range of m/z 100–1000, but to measure with the largest scan range possible without negatively affecting the overall analytical performance, and the results reveal that this would have captured the most relevant substances in this sample.

Several other interesting aspects could be investigated with this comprehensive dataset. Although many options are now available for ionisation, many participants chose to report substances exclusively in positive ESI mode (excluding negative mode), and no LC participant used alternative ionisation methods, for example APCI or APPI. A desire for better functionality to compare and merge peak lists measured using complementary ionisation techniques was expressed to increase the ease of use of multiple ionisation techniques for non-target screening, because this is not yet time-efficient for routine analysis. Similar functionality is also needed to merge results from complementary chromatographic approaches (e.g. HILIC-RP). The delicate balance between sample volume injected and chromatographic performance was also visible. Although the participant with the highest sample injection volume (200 μ L) also reported the most target compounds, neither melamine nor metformin were included, indicating that the high amount of solvent may have hindered the detection of these highly polar substances for chromatographic reasons. From the available data, it is clear that comprehensive non-target screening approaches can be applied with great success using either Orbitrap or TOF-based instruments. Instrument-specific aspects, including which compounds can only be measured successfully with the higher resolution of the Orbitraps, which ionise better or worse in specific instruments, which methods perform best for low intensity substances, which solvents and modifiers are better, and the effect of detection limits on detection etc., would require a custom-made spiked environmental sample to create known case studies. The parallel submission of a spiked sample was encouraged by all participants for future trials, to investigate some of these phenomena more closely. However, the nature of the dataset available enables a detailed analysis of several specific topics in non-target screening, which are illustrated with several examples in the following text.

Example: isobaric substances

Isobaric substances (where isobaric is used to refer to substances with the same exact mass) form some of the most challenging cases in high-throughput screening. With high-

accuracy data, isobaric substances often have the same molecular formula (but not always, because some can have the same exact mass within the instrumental accuracy [65]), and in environmental screening these can often be very similar substances structurally. The case of terbutylazine, sebutylazine, simazine, propazine, and their TPs, shown in Fig. 2, is a good example. Fragmentation information can be used to distinguish propazine ($2 \times -C_3H_7$) from terbutylazine and sebutylazine (both $-C_2H_5$, $-C_4H_9$) in the top row of Fig. 2, or simazine ($2 \times -C_2H_5$) from the desethyl TPs (both $-C_4H_9$) in the second row. As a result of the structural similarity, these substances also often co-elute, especially in generic chromatographic methods including those used in non-target methods. Two participants reported both simazine and desethylterbutylazine for the same m/z and RT, but with diagnostic fragments indicating the presence of both. In this sample, neither fragmentation nor retention information could separate terbutylazine and TPs from sebutylazine and TPs with the generic chromatography used, although clearly more institutes reported the terbutylazine-related substances. This is logical in terms of substance use: terbutylazine is widely used, whereas neither sebutylazine nor propazine are currently registered in the EU. However, propazine is present as a by-product in other triazines and could be observed when other triazine compounds are present if concentrations were sufficient. It is also interesting to reflect on the prioritisation that the number of references would give here: most institutes reported terbutylazine (10) and its TP terbutylazine-desethyl (9) instead of the isobaric simazine (4); the former were also provided as suspects (ESM Table S5); terbutylazine-desethyl was also reported by one GC participant. Without knowing terbutylazine to be present, sorting candidates purely by the number of references would clearly favour simazine (518 references versus 92 in ChemSpider, see Fig. 2). Thus care must also be taken with this strategy, and additional knowledge (especially the presence of related substances, fragmentation information, and retention time) should also be considered. This reinforces the point made by many participants that a better integration of identification strategies is necessary for non-target screening approaches and that the incorporation of use data, rather than the number of references alone, is desirable in the future.

Example: co-eluting substances

The case of 4 and 5-methylbenzotriazole summarises many challenges with the data evaluation. All participants reported these target substances together, agreeing that they required special methods to be separated (e.g. Ref. [66]) and could not be separated chromatographically with a generic (multi-residue) screening method. Because the data could not be compared by name (entries included 4/5-Methyl-Benzotriazole, 4 & 5-Methyl-Benzotriazole, 4- and 5-

methylbenzotriazole through to Tolyltriazole), or CAS numbers, SMILES or InChI Keys (given for either of the isomers, e.g. CAS 136-85-6 and 29385-43-1), in the end the comparison was performed by sorting by molecular formula $C_7H_7N_3$ because no other substances were reported with this formula. The limitation in representing structural uncertainties in InChI Keys is currently the subject of a new working group within IUPAC to improve data comparison of such cases. Other cases included the example of simazine, terbutylazine, and related compounds shown in Fig. 2 and the (often) co-eluting substances tramadol and *O*-desmethylvenlafaxine. Tramadol was reported much more frequently; two participants reported both substances for the same peak, and at least one other participant ruled out the presence of *O*-desmethylvenlafaxine manually as the diagnostic fragments m/z 133.0649 and 107.0492 were absent in the measured spectrum. Another interesting case is dibutylphthalate, for which seven participants reported at least one of three isomers (di-*n*-butyl phthalate, diisobutylphthalate, *p*-dibutylterephthalate) for 1–3 peaks per participant, using different strategies (reporting one isomer multiple times versus trying to “guess” which isomer was which, through to reporting with a generic name). Again, this reveals a great need for reporting structural uncertainties in a way compatible with high-throughput screening. This also reveals the need for a retention time index in LC–MS–MS; the question of which phthalate isomer was which was solved elegantly by the GC–MS participants with a combination of chemical ionisation and retention index information (see below).

Dependence on target and suspect lists

The results of the target and suspect screening in the collaborative trial revealed a clear dependence on the target and suspect lists used, which were in turn also highly dependent on the main focus of each laboratory. This is revealed by the large range of target compounds reported: 354 of 622 targets were reported once only, and the maximum reported by one institute was 167, which is a very large number of compounds considering the low concentrations observed in the sample. Several participants reported multiple suspect identities for one peak, such that the number of suspect substances and non-target peaks in Table 4 may be an overestimation in some cases, especially where suspect matching was performed on the exact mass and isotope pattern only. Many participants took the suspect screening the necessary steps further, performing selection or rejection of the matching suspects using retention time and fragmentation information (measured versus library or versus predicted fragments where no library spectrum was available) before reporting the (tentative) identification. The danger of screening based on mass only was revealed by a few cases where the suspect list contained the exact mass of the compound represented as a salt, although

this can certainly not be detected as such in the instrument. For example, tris[4-(diethylamino) phenyl]methylm acetate, $C_{31}H_{42}N_3 \cdot C_2H_3O_2$, CCN(CC)c1cccc(cc1)[C+](c2ccc(cc2)N(CC)CC)c3ccc(cc3)N(CC)CC.CC(=O)[O-] was reported at m/z 516.3565 by one participant, whereas the monoisotopic mass of the substance without the acetate salt is 456.3379 Da. Another example is a substance with the formula $C_7H_8N_4O_2$, which was reported once by three participants, but with a different identity each time (theophylline, 1,7-dimethylxanthine, or theobromine). Whereas the foremost was reported as a target confirmed by reference standard, the other two compounds were detected additionally by one GC participant with distinct EI MS spectra and good NIST library match values: similarity 700, probability 87 % for 1,7-dimethylxanthine, and similarity 823, probability 80 % for theobromine, although only the latter was reported officially because a cut-off of 800 was used. This again reveals the need for spectra and other additional information to support suspect screening. The concept of performing smarter suspect screening (“screen smart”, carefully selecting the compounds to be screened to suit the investigation) was discussed as often being more efficient than “screen big” with large, unselective suspect lists. Although the latter has the advantage of large coverage, this is at the expense of a high rate of false positives and thus often a larger burden of proof. Participants also agreed that it would be useful in the future to exchange target and suspect compound lists between institutes using an open platform, for example STOFF-IDENT [41]; the NORMAN Association is investigating this as a planned activity for 2015.

Molecular formula assignment, adduct detection, and homologous series

Moving on from target and suspect screening, an important aspect of non-target identification is the detection of the adduct state and assignment of molecular formulas to components of interest, and this is illustrated using homologous series identified by participants in the trial sample. Whereas the GC participants could use library searching to identify these common substances, the LC participants performed suspect screening using (i) lists of surfactants published previously [3], (ii) tentative identification through recognition of diagnostic fragments (e.g. 89.0593 and 133.0857 corresponding to $C_4H_8O_2$ and $C_6H_{12}O_3$, respectively), through to (iii) retrospective screening for surfactants after the discovery of diagnostic fragments of surfactants during non-target identification efforts. The polyethylene glycol (PEG) homologues provide an interesting example of adduct detection and molecular formula assignment. These were detected as the $[M+H]^+$, $[M+NH_4]^+$, and/or $[M+Na]^+$ species by several participants; see ESM Fig. S2 for the distribution of the $[M+H]^+$ and $[M+NH_4]^+$ species measured by one participant and Figs. S3 and S4 in the ESM for the adducts and isotopes detected for some

of these species. Where at least two adduct species are present, the neutral mass of the compound can be determined unequivocally; if only one species is present the adduct state must be assumed when calculating the molecular formula. Table 5 shows the calculation of the molecular formulas for three PEG species detected as different adducts by one participant.

This example highlights the importance of considering adducts and applying a comprehensive peak grouping before non-target identification, which is a feature now offered by many open source and vendor software packages to different degrees of comprehensiveness. Once the correct neutral mass was identified, molecular formulas were calculated with the elements C, H, N, O, and S. From five to seven formulas with a high score (>90) were proposed; the top four are shown in Table 5. The score considers the mass, isotope match, and isotope spacing; however, the last three columns of Table 5 show that these four formulas could be reduced to only one or two possible formulas if fragmentation information was also considered in the selection of the molecular formula. Although several approaches now include fragment information in formula selection, including SIRIUS [67], MOLGEN-MS/MS [68], and mzMine [69], these were not used in the trial, whereas of the vendor software used by different participants, only SmartFormula3D from Bruker associated fragmentation information with the precursor ions for formula calculation.

Retention time information in LC–HRMS

An important objective of the trial was to assess the use of retention time information in the LC screening approaches. Most participants measured the substances provided (given in Methods) and used the results to calculate a retention time index (RTI) for the LC data. Seven participants reported these RTI for some or all of their reported values in the DCTs. The results of 26 target compounds that were detected at least four times were used to investigate the RTI further. The calculated RTI was found to have below 3 % root mean square deviation (RSD) for eleven of these targets, in the range 3–6 % RSD for another eleven targets, and above 6 % RSD for the remaining four of the 26 target compounds. The RTIs of target substances reported by two participants were used to calculate log D values that were in turn used to search the STOFF-IDENT database [41], where empirical matching criteria are included. Above 80 % (50 of 60, 81 of 96) of the target compounds that were in STOFF-IDENT were also within the calculated log D range using the RTI standards. Although the majority of substances matched, the non-match of 10 and 15 manually validated target substances (per participant, respectively) reveals that a consensus approach should be considered when using retention time prediction approaches [70].

A different approach was used by another participant (14) to reject suspects on the basis of retention time (and other

information. Their initial suspect screening yielded 1060 and 538 suspect hits in positive and negative mode, respectively, which was reduced to 81 and 75 suspect matches, respectively, using additional mass accuracy, isotope, and peak score criteria. Of these, a further 27 and 40 suspects were rejected using retention time models developed using a SVM kNN QSPR approach (Aalizadeh et al., in prep), and a further four and five were rejected because they were not ESI-amenable. Finally 78 suspects were reported, as given in Table 4 (48 positive, 30 negative).

Although all participants agreed that the retention information is useful and essential for identification efforts, many participants expressed a need to have this information better integrated into their current software. Additionally, the non-targets found by some participants were clearly compounds such as sulfonated surfactants and polyfluorinated compounds, which often perform poorly in retention time prediction and other QSPR calculations [27, 71]. A further example of how useful retention information can be to confirm the identification of suspect compounds is given below for GC–MS.

Non-target identification example

Several tentatively-identified non-target compounds were proposed by participants, despite the limited time (Table 4), and these compounds could be found retrospectively in the raw files from other participants. One participant reported a double peak at m/z $[M-H]^-$ 199.0431 at 19.48 min (small intensity) and 19.98 min (large intensity) of a 40 min run, tentatively identified as mesitylene sulfonic acid (2,4,6-trimethylbenzene sulfonic acid, ChemSpider ID 69438) or isomer, molecular formula $C_9H_{12}SO_3$, with the fragments m/z 184.0196 and 135.0815. Another participant found a signal at m/z 199.0428 with a double peak at 5.91 and 6.15 min of a 25 min run (small and large intensity, respectively) with peaks at m/z 79.9566 (SO_3^-), 80.9644 (SO_4H^-), 119.0500 ($C_9H_{11}^-$), 135.0811 ($C_9H_{12}O^-$, or $-SO_2$), 183.0115 ($-CH_4$), and 184.0194 ($-CH_3$) with HCD 90 (see Fig. 3). A total of 135 candidates with this formula are in ChemSpider, but using MetFrag to rank on the basis of the in silico fragmentation of the candidates, trimethyl benzene sulfonic acid isomers and a few ethyl isomers are all ranked equally, with a score of one and four explained peaks. Although the larger peak is also present in the blank and seems to come from the sorbents used in the large volume sampling, this is two orders of magnitude lower than the intensities observed in the samples for both participants. However, the standard for mesitylene sulfonic acid was available at the second institute and although the retention time matched well (5.96 min), the fragments did not match; only two fragments (m/z 79.9573 and 135.0824) were seen in the standard, at varying collision energies including HCD 90. Thus the tentative identification cannot be

Table 5 Calculation of molecular formulas for non-target masses associated with different adducts and having diagnostic fragments for polyethylene glycol species, m/z 89.0593 and 133.0857; columns 1–6 provided by one participant, columns 7–9 (Yes, No) indicate whether the formula would have also matched given fragment information

Detected m/z Adduct	Neutral mass	Molecular formula	Score	Mass difference (ppm)	Matches fragment information		
					89, 133 (C ₄ H ₉ O ₂ , C ₆ H ₁₃ O ₃)	177, 221 ^a (C ₈ H ₁₇ O ₄ , C ₁₀ H ₂₁ O ₅)	265, 309 ^a (C ₁₂ H ₂₅ O ₆ , C ₁₄ H ₂₉ O ₇)
388.2550	NH ₄ ⁺	370.2212	98	1.2	Yes	Yes	No
			98	-2.2	Yes	Yes	Yes
			97	-1.0	Yes	No	No
432.2813	NH ₄ ⁺		91	-0.7	No	No	No
			98	1.5	Yes	Yes	No
			98	-2.0	Yes	Yes	No
			98	-1.5	Yes	Yes	Yes
			97	1.4	No	No	No
481.2629	Na ⁺	458.2734	98	-1.4	Yes	Yes	Yes
			97	-1.7	Yes	Yes	No
			97	1.4	Yes	Yes	Yes
			95	1.0	No	No	No

^a Fragment information for m/z 177.1122, 221.1385, 265.1647, and 309.1909 (shown as unit mass above for space reasons) taken from the NIST 2014 MS-MS library spectrum of octaethylene glycol (NIST# 1229846). Molecular formulas were calculated with the elements C, H, N, O, and S, the score (calculated using MassHunter Workstation Software B.06, Agilent Technologies) considers mass, isotope match, and isotope spacing

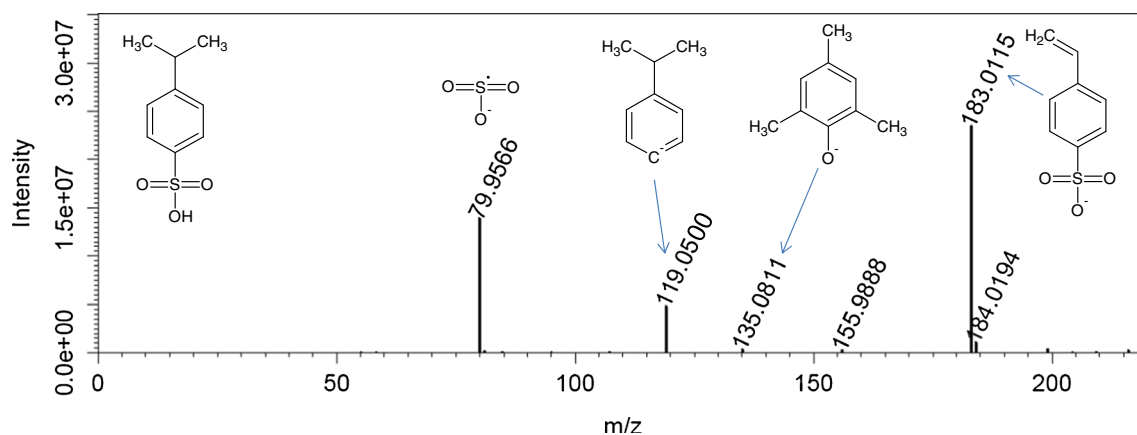


Fig. 3 Spectrum of the non-target found at 6.15 min by one participant, a tentatively identified candidate (4-isopropyl-benzenesulfonic acid, *insert*), and the resulting fragments. The original tentative identification

of mesitylene sulfonic acid can be rejected on the basis of mismatching fragments: see ESM Figs. S5 and S6

confirmed. However, the peak seen by the second participant at m/z 183.0115 is diagnostic for linear alkylbenzene sulfonic acid (LAS) species [3] and the structure below may be able to explain the spectra. Although not yet confirmed, this example reveals how measurements of the same sample by two different laboratories can be used to help in the identification.

GC-MS

For the GC-MS results, four institutes reported target compounds (confirmed with reference standards), with 64 results reported for 55 unique compounds. The target substances detected multiple times were all polycyclic aromatic hydrocarbons (PAHs): phenanthrene (four), anthracene and naphthalene (three each), and fluorene, fluoranthene, and pyrene (twice each). The remaining identifications were performed using library information, usually the NIST MS library [11] versions 02 (175,214 spectra of 147,350 compounds), 08 (220,460 EI MS spectra of 192,108 compounds), and 11 (243,893 EI MS spectra of 212,961 compounds), although one participant also used the Wiley library, 7th Edition [12]. Match factors in the range 700–800 (of a maximum of 999) were used as the lower limit for library matches (some also in combination with the probability estimates) and most participants also confirmed these matches manually. An additional 295 compounds were reported from library matches (identified but not confirmed with reference standards), such that in all, 255 unique compounds were reported with a single identity as a target substance or library match. A further 47 peaks were reported as library matches but with the isomer unidentified, and 35 reported substances remained unknown (many participants did not report unknown substances). Merging the target and library match compound lists revealed that phenanthrene was detected by six of seven participants, and dibutylphthalate, diethylphthalate, tris(1-chloro-2-propanol)phosphate (TCPP), and caffeine (five times each) and 4-oxoisophorone, benzothiazole,

cyclohexylisocyanate and triphenylphosphine oxide, and 1, 1'-(1,4'phenylene)bis-ethanone (four each) were also detected frequently. Several substances overlapped with the LC-MS targets and suspects, notably the several phthalate and alkyl phosphate species, caffeine, and carbamazepine; in total 46 substances were reported as identified in both LC and GC, including the suspects and library matches, of the 858 unique compounds reported in total (~5.4 %).

The sample preparation was not ideal for GC analysis and the participants also used different solvents. Therefore, the relatively few substances detected and limited comparability between GC participants is not completely unreasonable. A separate sample for GC-MS analysis is recommended for future trials. Although some participants also quantified their target substances, this data is not presented here because the dataset was too small for a meaningful evaluation.

High resolution GC-MS

One participant used high resolution GC-APCI QTOFMS during the trial, combining the results with EI MS in an interesting approach. As mentioned above, the workflow to identify substances with the GC-APCI QTOFMS approach effectively followed the target, suspect, non-target approaches used for LC-HRMS screening, rather than the standard target screening plus NIST database search used generally in EI MS. Further participants used GC×GC, while another participant used the combination of chemical ionisation and EI MS to identify the molecular mass corresponding with the EI MS spectrum, shown in Fig. 4. The base peak m/z 149 (indicative of phthalates) was observed three times but was only associated with the m/z $[M+H]^+$ 279 twice, supporting the proposed NIST matches of dibutylphthalate isomers. However, consideration of the experimental KRI with the literature values in NIST prompted an alteration of the tentatively-proposed identities in Fig. 4 (shown by the dashed boxes and blue arrows),

because diisobutylphthalate (shown left) had a lower retention index than dibutylphthalate (shown right). Final confirmation using a reference standard, however, is still necessary. The HRMS–MS spectrum obtained via APCI QTOFMS–MS is shown in Fig. 4b.

As for the LC participants, the GC participants mentioned that the data processing is a limiting factor for using GC×GC and/or soft ionisation in combination with EI MS and the need for exact mass spectra, especially MS–MS spectra, was expressed. This reinforces the observation that the workflow used for identification is determined by the ionisation technique used (EI versus softer ionisation methods) and the size of the spectral library available rather than the chromatography.

Discussion and perspectives

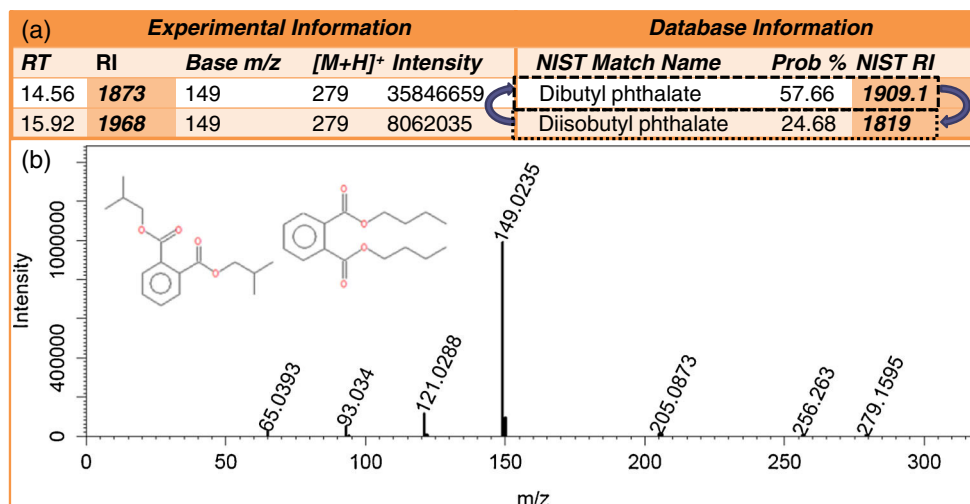
Data processing aspects, HRMS–MS

The participation of many institutes in analysing the one sample reveals that non-target screening methods have come a long way in a short time and that a comprehensive target screening complemented with suspect screening forms an essential prerequisite to non-target identification. Most substances reported were identified through target or suspect screening of data acquired with non-targeted acquisition. Tentatively identified non-target compounds were reported by a few participants, even though non-target identification is still a very time-intensive process. Many participants included MS–MS and retention time information where possible to support the proposed identity of suspect compounds, addressing the disadvantages of performing suspect screening on the basis of exact mass only. Suspect screening proved a useful step to propose tentative identifications for several compounds and ease the burden of identification. The large overlap between target and suspect compounds proves that this was successful

in many cases. The selection of the peaks of interest for non-target identification remains very subjective; it is clear that not all of the up to several thousand unknown peaks can be identified. One option is the prioritisation of masses occurring in several samples (e.g. Refs. [3, 9], and although only one sample was available for the trial, several participants only considered peaks when they occurred in at least two repeated analytical runs, to reduce the effect of noise or other interferences on the results. In terms of non-target identification, whereas some participants determined the molecular formula first, followed by elucidation on the basis of the one or several possible molecular formulas, other participants went straight to identification with an exact mass. The latter strategy, although resulting in more candidates, has been revealed recently to be more effective from a data processing perspective [72] (it is for example easier to select candidates with one or some formula(s) from an exact mass search than to merge the results of several separate calculations).

A new trend revealed during the trial is the move towards finding related substances, with several strategies used. These included searching for common fragments or neutral losses, suspect screening for homologues, and even homologue series detection. Other strategies that have been used elsewhere included selection by mass defect to avoid matrix interferences in sediments [23], and although the Kendrick mass defect is often used for natural organic matter or petroleum analysis (e.g. Refs. [73, 74]), this was not mentioned by any of the participants and it remains to be seen whether this may have an application to environmental samples. As discussed above, challenges still remain with regard to the reporting of substances that cannot be fully characterised structurally using the information available (e.g. long chain surfactants, lipid species with undefined branching and/or double bond locations) and also the level of confirmation required before being able to report these substances in investigations. As discussed in the Introduction, a combination of a revised version of the

Fig. 4 (a) Use of complementary ionisation techniques (EI, PCI), library match, and retention information (experimental KRI versus measured values in NIST) by one participant to identify two dibutylphthalate isomers and reverse the order of proposed database matches on the basis of retention information. (b) The corresponding APCI QTOFMS–MS spectrum of the suspect dibutylphthalate reported by another participant, which could be either isomer (both structures: see *insert*)



EU IP system (to, e.g., better reflect the availability of instruments with much higher resolution than the 10,000 in the original guideline [6]) with the level system shown in Fig. 1 could be a way to move forward. A further improvement would be better processing of data-independent acquisition data, such that fragments are available for all substances, not just selected precursors. Another interesting approach could be to speed up non-target identification using alternative databases, for example those including natural products, to assist in finding background natural compounds that are not necessarily of interest for non-target identification of potentially hazardous anthropogenic substances.

The number of targets and suspects detected only once indicates that the results are highly dependent on the institute knowledge and focus area and on the target and suspect lists used. This reinforces the need for additional evidence before reporting suspect identification and for an improved representation of identification confidence, for example updated IPs. It is clear from the number of different software packages and approaches used that the data processing techniques are very different, with vendor versus open source approaches and many different sources of information used by the participants. The need to harmonise information sources was expressed and, as a result of the recognition that one participant's target is another participant's suspect or unknown, an agreement was made to exchange target and suspect lists in a centralised platform, where the development of search criteria and prioritisation methods for these lists will be essential to enable application of the "screen smart" approach. Additionally, the results of this study have been provided as supporting information for use in future investigations. Enhancing this with the upload of mass spectra of target compounds to an open access database, for example the NORMAN MassBank [43], would help improve the success of target, suspect, and even non-target screening immensely. Another need clearly expressed by the participants was for better data processing of complementary ionisation modes and methods (e.g. positive or negative, APCI, and APPI) and for data-independent acquisition techniques, such that fragmentation can be obtained for all substances rather than selected precursor masses, and a better integration of retention information.

Lessons learned from the trial

The first collaborative non-target screening trial run by the NORMAN Association was a very ambitious endeavour coinciding with the Third Joint Danube Survey in response to the need expressed by members in late 2012. The response from participants was very positive and the objective of discussing ways to harmonise non-target screening of environmental samples within Europe was fulfilled with a series of three well-attended workshops [75–77]. The analytical methods are already reasonably well harmonised, with the

main differences resulting from instrument-specific settings that participants cannot reasonably be expected to change. Although the data processing workflows were not as harmonised, with many participants expressing the need for better integration of information, workflows, and software, a great deal can be learnt by reflecting on the results of the other participants. Participants were given a lot of freedom for this collaborative trial, and there was also interest in a more tightly-defined interlaboratory comparison in the future, with an additional spiked sample and a more comprehensive suspect list from greater comparability.

All participants provided the raw data of their measurements for this trial, such that a retrospective analysis of the results will be possible, and finding a suitable environmental repository for this data will be the subject of future efforts. Many participants were keen to investigate their results further, using the information gained during the trial, to perform retrospective identification. The trend for retrospective analysis, used already during the evaluation of these results, is increasing and the importance of this in the future should not be underestimated.

Acknowledgments The authors gratefully acknowledge the contributions of many others who helped during the analysis, reporting and discussion of the results presented here: Philipp Longree, Martin Loos, Matthias Ruff and Jennifer Schollee (Eawag), Margit Petre, Anett Kloß and Werner Brack (UFZ), Reza Aalizadeh, Nikiforos Alygizakis and Pablo Gago Ferrero (University of Athens), Sylvain Merel (University of Tübingen), Alexander van Nuijs, Alin Ionas and Adrian Covaci (University of Antwerp), Martijn Pijnappels (Rijkswaterstaat), Christelle Margoum (Irstea), Claudio Bortolini (University of Padua), Marie-Hélène Devier and Hélène Budzinski (University of Bordeaux, France), Christophe Tondelier and Mathilde Chachignon (VERI), Tobias Bader and Thomas Lucke (Zweckverband Landeswasserversorgung), Michael Schlüsener (BfG), Mar Esperanza (Suez Environnement), Siniša Repc (Croatian Waters), Peter Tarabek (Slovak National Water Reference Laboratory), Christoph Ruttkies (IPB, Halle, Germany), the ChemSpider support team, and Thierry Faye from Agilent. The water sample of the Danube was kindly provided by the International Commission for the Protection of the Danube River (ICPDR) as part of the Joint Danube Survey 3. Author contributions: ELS, HPS, JS, IMI, PO, MK, TS, PH, TL, SG, NST, CZ, and JH were involved in the core team preparing the trial, evaluation, workshop, and publication; MI, TP, RdB, MR, MO, UK, WS, AG, NN, GL, PB, SB, DS, and PR participated in the trial and publication efforts.

This trial was initiated by the NORMAN Association with in-kind contributions from participants; no fee was charged for participation. This work was supported in part by the SOLUTIONS project, which received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under Grant Agreement No. 603437.

Compliance with ethical standards

Conflict of Interest The authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent No violation of human or animal rights occurred during this investigation.

References

- 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
- Hernandez F, Pozo OJ, Sancho JV, Lopez FJ, Marin JM, Ibanez M (2005) Strategies for quantification and confirmation of multi-class polar pesticides and transformation products in water by LC-MS2 using triple quadrupole and hybrid quadrupole time-of-flight analyzers. *TraC Trends Anal Chem* 24(7):596–612
- Schymanski EL, Singer HP, Longree P, Loos M, Ruff M, Stravs MA, Ripolles Vidal C, Hollender J (2014) Strategies to Characterize Polar Organic Contamination in Wastewater: Exploring the Capability of High Resolution Mass Spectrometry. *Environ Sci Technol* 48(3):1811–1819
- Zedda M, Zwiener C (2012) Is nontarget screening of emerging contaminants by LC-HRMS successful? A plea for compound libraries and computer tools. *Anal Bioanal Chem* 403(9):2493–2502
- 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
- European Commission (2002) Commission Decision of 12 August 2002 Implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. Official Journal of the European Communities L221:29 <http://old.eu-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:221:0008:0036:EN:PDF>. Accessed 30 Jan 2015
- Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J (2014) Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ Sci Technol* 48(4):2097–2098
- Creek DJ, Dunn WB, Fiehn O, Griffin JL, Hall RD, Lei ZT, Mistrik R, Neumann S, Schymanski EL, Sumner LW, Trengove R, Wolfender JL (2014) Metabolite identification: are you sure? And how do your peers gauge your confidence? *Metabolomics* 10(3):350–353
- Letzel T, Lucke T, Schulz W, Sengl M, Letzel M (2014) OMI (Organic Molecule Identification) in water using LC-MS/MS): Steps from “unknown” to “identified”: a contribution to the discussion. *Lab More Int* 4:24–28
- Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reilly MD, Thaden JJ, Viant MR (2007) Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) *Metabolomics Standards Initiative (MSI)*. *Metabolomics* 3(3):211–221
- NIST/EPA/NIH (2011) NIST Mass Spectral Library, various editions. National Institute of Standards and Technology. US Secretary of Commerce, USA
- McLafferty FW (2000) Wiley Registry of Mass Spectral Data, 7th edn, ISBN-10: 0471440981
- Stein S (2012) Mass Spectral Reference Libraries: An Ever-Expanding Resource for Chemical Identification. *Anal Chem* 84(17):7274–7282. doi:10.1021/ac301205z
- Kind T, Fiehn O (2010) Advances in structure elucidation of small molecules using mass spectrometry. *Bioanal Rev* 2(1–4):23–60
- Scheubert K, Hufsky F, Bocker S (2013) Computational mass spectrometry for small molecules. *J Cheminform* 5(12). doi:10.1186/1758-2946-5-12
- Portoles T, Mol JGJ, Sancho JV, Hernandez F (2014) Use of electron ionization and atmospheric pressure chemical ionization in gas chromatography coupled to time-of-flight mass spectrometry for screening and identification of organic pollutants in waters. *J Chromatogr A* 1339:145–153
- Hernandez F, Ibanez M, Portoles T, Cervera MI, Sancho JV, Lopez FJ (2015) Advancing towards universal screening for organic pollutants in waters. *J Hazard Mater* 282:86–95
- Helbling DE, Hollender J, Kohler HPE, Singer H, Fenner K (2010) High-Throughput Identification of Microbial Transformation Products of Organic Micropollutants. *Environ Sci Technol* 44(17):6621–6627
- Kern S, Fenner K, Singer HP, Schwarzenbach RP, Hollender J (2009) Identification of Transformation Products of Organic Contaminants in Natural Waters by Computer-Aided Prediction and High-Resolution Mass Spectrometry. *Environ Sci Technol* 43(18):7039–7046
- Huntscha S, Hofstetter TB, Schymanski EL, Spahr S, Hollender J (2014) Biotransformation of benzotriazoles: insights from transformation product identification and compound-specific isotope analysis. *Environ Sci Technol* 48(8):4435–4443
- Moschet C, Piazzoli A, Singer H, Hollender J (2013) Alleviating the Reference Standard Dilemma Using a Systematic Exact Mass Suspect Screening Approach with Liquid Chromatography-High Resolution Mass Spectrometry. *Anal Chem* 85(21):10312–10320
- Thurman EM, Ferrer I, Blotvogel J, Borch T (2014) Analysis of Hydraulic Fracturing Flowback and Produced Waters Using Accurate Mass: Identification of Ethoxylated Surfactants. *Anal Chem* 86(19):9653–9661
- Chiaia-Hernandez AC, Schymanski EL, Kumar P, Singer HP, Hollender J (2014) Suspect and nontarget screening approaches to identify organic contaminant records in lake sediments. *Anal Bioanal Chem* 406(28):7323–7335
- Howard PH, Muir DCG (2010) Identifying New Persistent and Bioaccumulative Organics Among Chemicals in Commerce. *Environ Sci Technol* 44(7):2277–2285
- Howard PH, Muir DCG (2011) Identifying New Persistent and Bioaccumulative Organics Among Chemicals in Commerce II: Pharmaceuticals. *Environ Sci Technol* 45(16):6938–6946
- Howard PH, Muir DCG (2013) Identifying New Persistent and Bioaccumulative Organics Among Chemicals in Commerce. III: Byproducts, Impurities, and Transformation Products. *Environ Sci Technol* 47(10):5259–5266
- Ulrich N, Schuurmann G, Brack W (2011) Linear Solvation Energy Relationships as classifiers in non-target analysis-A capillary liquid chromatography approach. *J Chromatogr A* 1218(45):8192–8196
- Wolf S, Schmidt S, Muller-Hannemann M, Neumann S (2010) In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinf* 11:148
- HighChem (2014) Mass Frontier, 7th edn. HighChem Ltd./Thermo Scientific, Bratislava
- Little JL, Cleven CD, Brown SD (2011) Identification of “Known Unknowns” Utilizing Accurate Mass Data and Chemical Abstracts Service Databases. *J Am Soc Mass Spectrom* 22(2):348–359
- Little JL, Williams AJ, Pshenichnov A, Tkachenko V (2012) Identification of “Known Unknowns” Utilizing Accurate Mass Data and ChemSpider. *J Am Soc Mass Spectrom* 23(1):179–185
- RSC (2014) ChemSpider www.chemspider.com. Royal Society of Chemistry. Accessed 17 Dec 2014
- ICPDR (2015) Joint Danube Survey 3 www.danubesurvey.org. ICPDR - International Commission for the Protection of the Danube River. Accessed 23 Jan 2015
- Liska I, Wagner F, Deutsch K, Sengl M, Slobodnik J (2015) Joint Danube Survey 3 Final Scientific Report (in print). Vienna, Austria
- Schulze T, Krauss M, Bahlmann A, Hug C, Walz K-H, Brack W (2014) Onsite large volume solid phase extraction – how to get 1000 litres of water into the laboratory? Society for Environmental Toxicology and Chemistry (SETAC) Europe 24th Annual Meeting, Basel, Switzerland, 11–15 May, 2014
- ChemAxon (2015) MarvinSketch Calculator Plugins 15.1.12.0 <http://www.chemaxon.com>. Accessed 30 Jan 2015

37. Greco G, Grosse S, Letzel T (2013) Serial coupling of reversed-phase and zwitterionic hydrophilic interaction LC/MS for the analysis of polar and nonpolar phenols in wine. *J Sep Sci* 36(8):1379–1388
38. LW/HSWT (2014) DAIOS Substance Database <http://www.daios-online.de/daios/>. Zweckverband Landeswasserversorgung, Langenau, Germany. Accessed 18 Dec 2014 (login only)
39. Muller A, Schulz W, Ruck WKL, Weber WH (2011) A new approach to data evaluation in the non-target screening of organic trace substances in water analysis. *Chemosphere* 85(8):1211–1219
40. NCBI (2014) PubChem (<https://pubchem.ncbi.nlm.nih.gov/>). National Center for Biotechnology Information. Accessed 17/12/2014
41. LFU/LW/HSWT/TUM (2014) STOFF-IDENT Substance Database <http://bb-x-stoffident.hswt.de/stoffidentjpa/app> Environmental Agency of Bayern, Germany (Bayerisches Landesamt für Umwelt). Accessed 1 Apr 2015 (login only)
42. MassBank (2014) MassBank www.massbank.jp. Accessed 17 Dec 2014
43. MassBank (2014) NORMAN MassBank www.massbank.eu. Accessed 17 Dec 2014
44. 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 MY, Nakanishi H, Ikeda K, Akimoto N, Maoka T, Takahashi H, Ara T, Sakurai N, Suzuki H, Shibata D, Neumann S, Iida T, 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 Spectrom* 45(7):703–714
45. HighChem (2014) mzCloud <https://www.mzcloud.org/>. HighChem Ltd., Bratislava, Slovakia. Accessed 17 Dec 2014
46. SIS (2014) The NIST 14 Mass Spectral Library <http://www.sisweb.com/software/ms/nist.htm>. Scientific Instrument Services. Accessed 17 Dec 2014
47. Wiley (2013) Wiley Registry of Mass Spectral Data, 10th Edn. Wiley, ISBN: 978-0-470-52037-6
48. Broecker S, Herre S, Wust B, Zweigenbaum J, Pragst F (2011) Development and practical application of a library of CID accurate mass spectra of more than 2,500 toxic compounds for systematic toxicological analysis by LC-QTOF-MS with data-dependent acquisition. *Anal Bioanal Chem* 400(1):101–117
49. Agilent (2015) Broecker, Herre, and Pragst Accurate Mass Personal Compound Database and Library (PCDL) for Forensics and Toxicology. Agilent Technologies, Santa Clara
50. Agilent (2015) Accurate Mass Personal Compound Database and Library (PCDL) for Pesticides. Agilent Technologies, Santa Clara
51. Scripps (2015) METLIN: Metabolite and Tandem MS Database <http://metlin.scripps.edu/>. Scripps Center for Metabolomics. Accessed 23/3/2015
52. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G (2005) METLIN: a metabolite mass spectral database. *Ther Drug Monit* 27(6):747–751
53. Bruker (2014) Bruker ToxScreener™ - A Comprehensive Screening Solution for Forensic Toxicology. Bruker Daltronics, Inc., Bremen, Germany
54. Sciex (2015) iMethod Application - LC/MS/MS Meta Library Version 1.0 for Cliquant Software <http://sciex.myshopify.com/products/imethod-application-lcmsms-meta-library-version-10-for-cliquant-software>. AB Sciex. Accessed 23/3/2015
55. Gerlich M, Neumann S (2013) MetFusion: integration of compound identification strategies. *J Mass Spectrom* 48:291–298
56. Waters (2014) MassFragment (version 1.3) http://www.waters.com/webassets/cms/library/docs/720004823en.pdf?locale=en_US. Accessed 27 Jan 2015
57. Tellstroem V, Dunsbach R (2014) Technical Note TN-26: SmartFormula 3D – the new Dimension in Substance Identification – From Mass Spectrum to Chemical Formula. Bruker Daltronics, Inc., Bremen, Germany
58. Hill AW, Mortishire-Smith RJ (2005) Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. *Rapid Commun Mass Spectrom* 19:3111–3118
59. Agilent (2011) Agilent MassHunter Molecular Structure Correlator (MSC) Software, Revision A. Agilent Technologies, Santa Clara
60. NIST (2005) Automated Mass Spectral Deconvolution and Identification System (AMDIS), 26th edn. National Institute of Standards and Technology (NIST), US Department of Defense, USA
61. Daylight (2012) SMILES- A Simplified Chemical Language <http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html>. Daylight Chemical Information Systems Inc. Accessed 30 Jan 2015
62. IUPAC (2012) The IUPAC International Chemical Identifier <http://www.iupac.org/inchi/>. International Union of Pure and Applied Chemistry. Accessed 30 Jan 2015
63. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open Babel: An open chemical toolbox. *J Cheminform* 3:33
64. Liska I, Wagner F, Slobodnik J (2008) Joint Danube Survey 2: Final Scientific Report. Editors: Liska I, Wagner F, Slobodnik J ICPDR – International Commission for the Protection of the Danube River, Vienna, Austria
65. Kind T, Fiehn O (2007) Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinf* 8:105
66. Weiss S, Jakobs J, Reemtsma T (2006) Discharge of three benzotriazole corrosion inhibitors with municipal wastewater and improvements by membrane bioreactor treatment and ozonation. *Environ Sci Technol* 40(23):7193–7199
67. Rasche F, Svatos A, Maddula RK, Bottcher C, Bocker S (2011) Computing Fragmentation Trees from Tandem Mass Spectrometry Data. *Anal Chem* 83(4):1243–1251
68. Meringer M, Reinker S, Zhang JA, Muller A (2011) MS/MS Data Improves Automated Determination of Molecular Formulas by Mass Spectrometry. *Match-Commun Math Co* 65(2):259–290
69. Pluskal T, Uehara T, Yanagida M (2012) Highly Accurate Chemical Formula Prediction Tool Utilizing High-Resolution Mass Spectra, MS/MS Fragmentation, Heuristic Rules, and Isotope Pattern Matching. *Anal Chem* 84(10):4396–4403
70. Schymanski EL, Gallampois CMJ, Krauss M, Meringer M, Neumann S, Schulze T, Wolf S, Brack W (2012) Consensus Structure Elucidation Combining GC/EI-MS, Structure Generation, and Calculated Properties. *Anal Chem* 84(7):3287–3295
71. Eckel WP, Kind T (2003) Use of boiling point-Lee retention index correlation for rapid review of gas chromatography-mass spectrometry data. *Anal Chim Acta* 494(1–2):235–243
72. Schymanski EL, Gerlich M, Ruttkies C, Neumann S (2014) Solving CASMI 2013 with MetFrag, MetFusion and MOLGEN-MS/MS. *Mass Spectrometry* 3 (Special Issue 2):S0036
73. Reemtsma T (2009) Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry Status and needs. *J Chromatogr A* 1216(18):3687–3701
74. Reemtsma T (2010) The carbon versus mass diagram to visualize and exploit FTICR-MS data of natural organic matter. *J Mass Spectrom* 45(4):382–390
75. NORMAN Association (2015) Non-target screening techniques for environmental monitoring <http://www.norman-network.net/?q=node/190>. NORMAN Association. Accessed 27 Jan 2015
76. NORMAN Association (2015) Workshop on Non-Target Screening: <http://www.norman-network.net/?q=node/162>. NORMAN Association. Accessed 27 Jan 2015
77. NORMAN Association (2015) NORMAN MassBank Workshop <http://www.norman-network.net/?q=node/163>. NORMAN Association. Accessed 27 Jan 2015

Analytical and Bioanalytical Chemistry

Electronic Supplementary Material

Non-target screening with high resolution mass spectrometry: Critical review using a collaborative trial on water analysis

Emma L. Schymanski, Heinz P. Singer, Jaroslav Slobodnik, Ildiko M. Ipolyi, Peter Oswald, Martin Krauss, Tobias Schulze, Peter Haglund, Thomas Letzel, Sylvia Grosse, Nikolaos S. Thomaidis, Anna Bletsou, Christian Zwiener, María Ibáñez, Tania Portolés, Ronald de Boer, Malcolm J. Reid, Matthias Onghena, Uwe Kunkel, Wolfgang Schulz, Amélie Guillon, Naïke Noyon, Gaëla Leroy, Philippe Bados, Sara Bogialli, Draženka Stipaničev, Pawel Rostkowski, Juliane Hollender

Supplementary 2 - 216_2015_8681_MOESM2_ESM.tar

Contents:

Figures

Fig. S1	Peak distribution comparison: HILIC-RP and RP	2
Fig. S2	Chromatography of Polyethylene Glycol (PEG) Suspects	3
Fig. S3	Relative adduct intensities of selected PEG Suspects	3
Fig. S4	Adduct and Isotope Distribution for a selected PEG Suspect	3
Fig. S5	Comparison of extracted ion chromatograms: non-target and standard	4
Fig. S6	Comparison of MS/MS spectrum: non-target and standard	4

Tables

Table S1	Homogeneity and stability results for 20 GC substances	5
Table S2	Homogeneity and stability results for 50 LC substances	5
Table S3	Analytical methods summary for GC participants, part 1	7
Table S4	Analytical methods summary for GC participants, part 2, plus software	7
Table S5	List of suspect substances detected in previous surveys	8
Table S6	Summary of target, suspect and non-target substances reported for GC	8

Abbreviations:

GC, LC	Gas Chromatography, Liquid Chromatography
HILIC	Zwitterionic hydrophilic interaction chromatography
PEG	Polyethylene glycol homologues
RP	Reversed phase
RT	Retention time

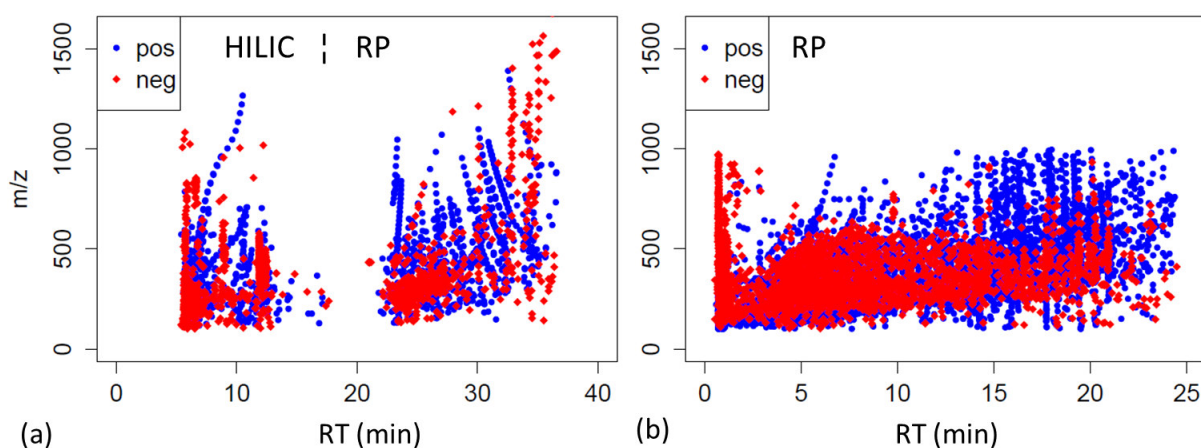


Fig. S1 Distribution of all peaks detected in positive (blue circle) and negative (red diamond) mode (a) in the coupled HILIC-RP system of participant L3 (655 positive, 633 negative on HILIC; 866 positive, 451 negative on RP) and (b) in the RP system of participant L2 (6029 and 2668 peaks, respectively)

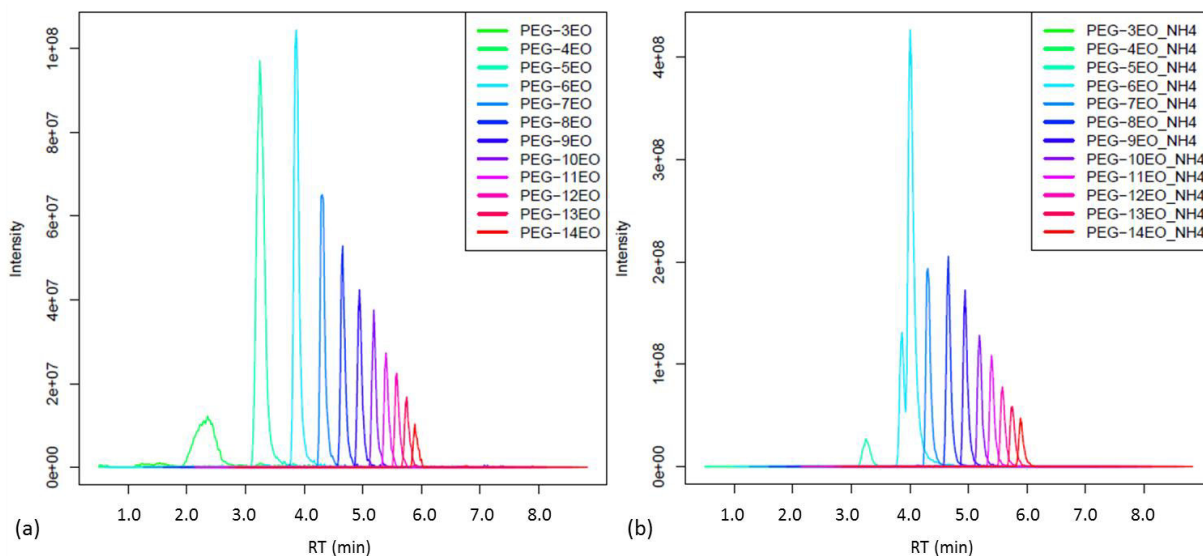


Fig. S2 Polyethylene Glycol (PEG) Series in the Collaborative Trial Sample. (a) $[M+H]^+$ species for triethyleneglycol (PEG-3EO, $C_6H_{14}O_4$) to tetradecaethylene glycol (PEG-14EO, $C_{28}H_{58}O_{15}$). (b) The corresponding $[M+NH_4]^+$ species

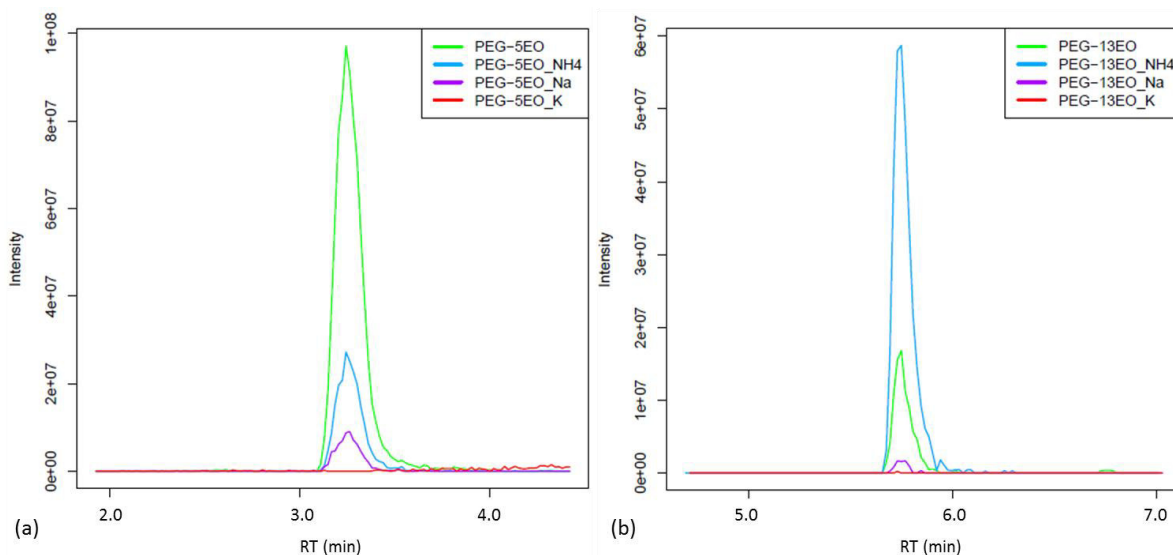


Fig. S3 (a) Adduct intensities for the PEG-5EO suspects, where $[M+H]^+$ is the most intense and (b) adduct intensities from PEG-13EO, where $[M+NH_4]^+$ is the most intense signal. For PEG-13EO even the $[M+K]^+$ is detected

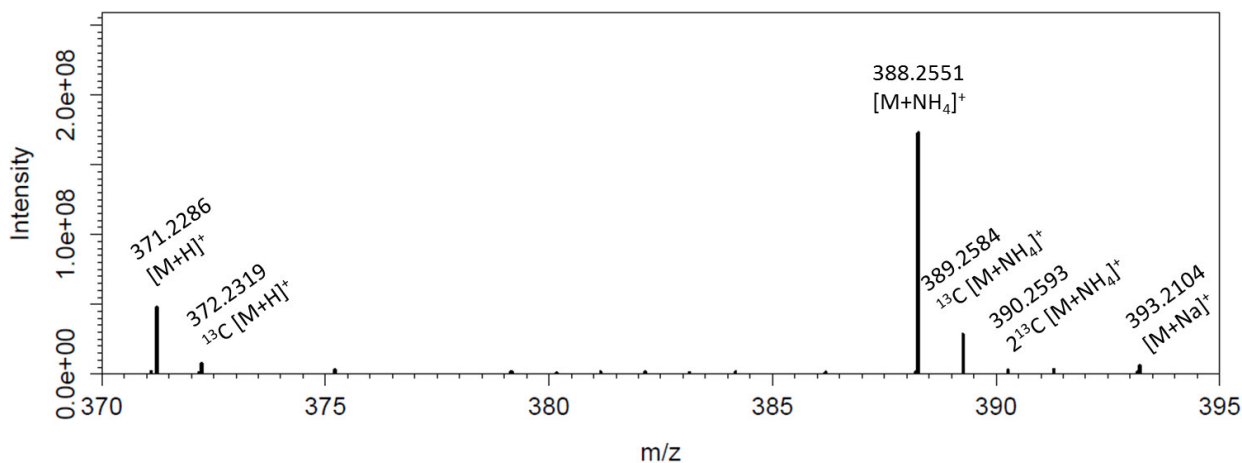


Fig. S4 Adduct and isotope signals observed for octaethylene glycol (PEG-8EO, $C_{16}H_{34}O_9$), measured by one participant

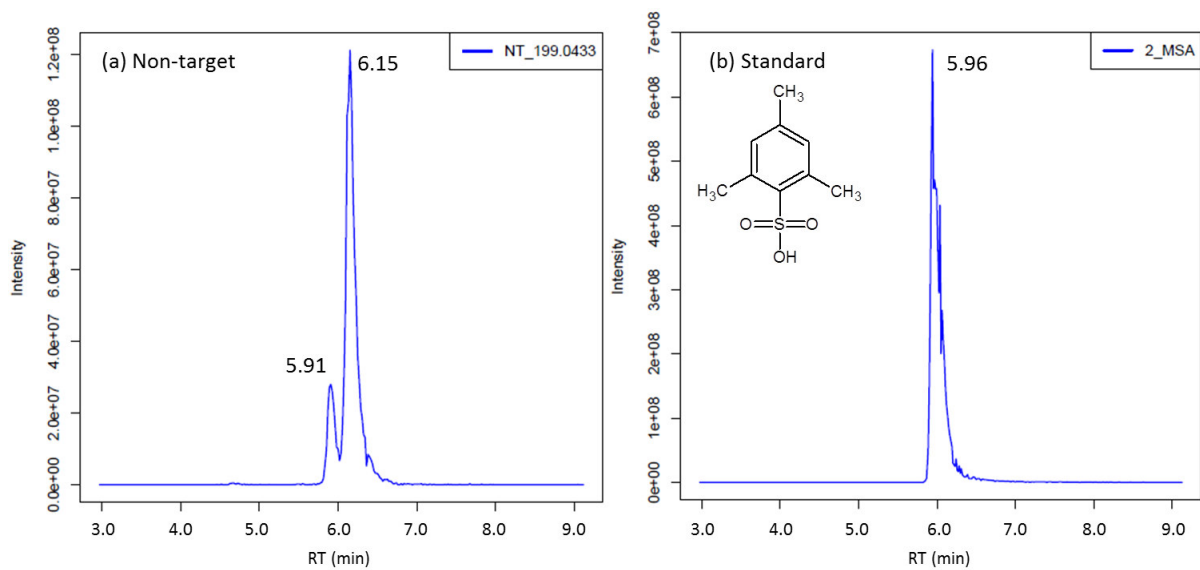


Fig. S5 Extracted ion chromatogram of (a) the non-target at m/z 199.0433 and (b) the standard of 2-mesitylene sulfonic acid (2-MSA, 2,4,6-trimethylbenzene sulfonic acid) measured under the same chromatographic conditions

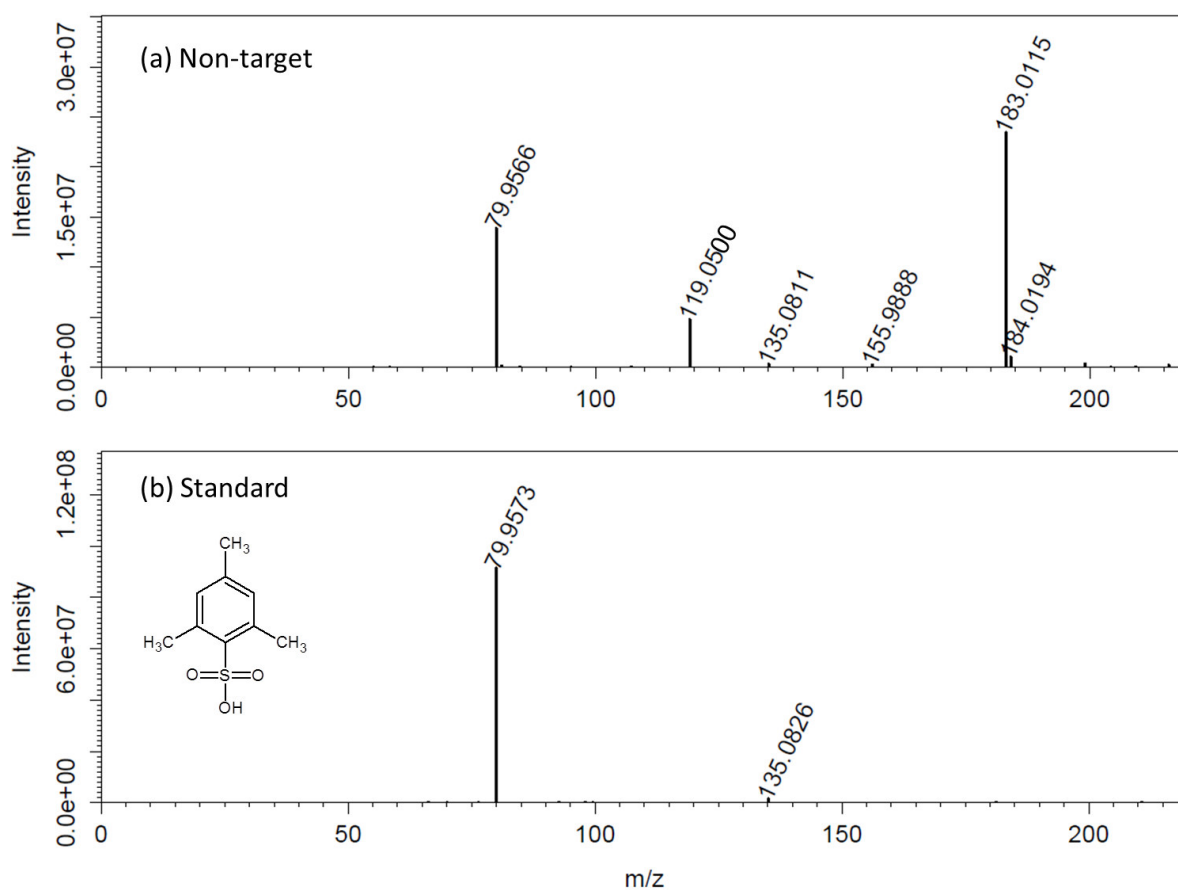


Fig. S6 MS/MS comparison (a) the measured non-target spectrum, HCD 90, RT 6.20 min (b) the reference standard 2-mesitylene sulfonic acid, HCD 90 at RT 5.97 min

Table S1 Homogeneity and stability testing results for 20 substances in GC-MS (selected randomly to cover the full retention time and polarity characteristics), calculated over 3 replicates. Samples were reconstituted in acetonitrile

		Homogeneity						Stability: Day X rel. to Day 1	
		Lot 15 (3 rep.)		Lot 33 (3 rep.)		Lot 50 (3 rep.)		10 : 1	30 : 1
Name	CAS	Avg [ug/L]	CV [%]	Avg [ug/L]	CV [%]	Avg [ug/L]	CV [%]	Rel. Conc.	Rel. Conc.
n-Hexadecanoic acid	57-10-3	0.1544	0.85	0.1395	2.38	0.1537	2.54	0.999	0.997
Phthalimide	85-41-6	0.1381	0.78	0.1261	0.84	0.1345	4.34	0.998	1.006
Dibutyl phthalate	84-74-2	0.1170	0.55	0.1061	4.22	0.1080	7.02	1.019	1.018
1-Chloroundecane	2473-03-2	0.1158	0.85	0.1046	1.30	0.1052	6.46	1.003	1.009
Cyclohexane, isocyanato-	3173-53-3	0.0981	0.81	0.0908	3.18	0.0916	2.24	1.009	1.000
Ethanone, 1,1'-(1,4-phenylene)bis-	1009-61-6	0.0945	0.92	0.0922	6.98	0.0900	2.88	0.995	0.995
Caprolactam	105-60-2	0.0819	0.85	0.0731	3.25	0.0757	3.35	1.001	1.002
1,4-Benzenediamine, N-(1-methylethyl)-N'-phenyl-	101-72-4	0.0772	0.84	0.0721	7.54	0.0698	5.05	1.000	1.003
Bisphenol A	80-05-7	0.0755	0.74	0.0675	1.76	0.0701	6.01	0.997	0.992
Phenol, 2,4-bis(1,1-dimethylethyl)-	96-76-4	0.0728	0.58	0.0703	6.07	0.0673	3.08	1.019	1.033
Tetrahydrofuran-2-one, 5-[1-hydroxyhexyl]-		0.0696	0.75	0.0627	2.57	0.0678	4.66	0.986	0.992
Benzoic acid	65-85-0	0.0677	0.90	0.0597	0.62	0.0632	3.36	1.019	1.013
Octadecane, 2-methyl-	1560-88-9	0.0609	0.89	0.0547	0.89	0.0565	2.74	1.004	1.005
Ethanone, 1,1'-(1,4-phenylene)bis-	1009-61-6	0.0606	0.38	0.0535	1.30	0.0542	6.08	1.004	1.001
Unknown (m/z 185)		0.0581	0.26	0.0565	5.18	0.0515	6.40	0.996	0.998
Cyclohexane, isothiocyanato-	1122-82-3	0.0547	0.90	0.0519	5.24	0.0495	4.63	0.993	0.997
Diethyl Phthalate	84-66-2	0.0535	0.74	0.0455	0.91	0.0508	2.28	1.001	1.013
Diisobutyl phthalate	84-69-5	0.0503	0.73	0.0482	5.31	0.0487	0.63	1.030	1.030
Unknown (m/z 133)		0.0493	0.42	0.0443	3.98	0.0468	3.33	1.028	1.021
Heptane, 1-(1-butenyloxy)-, (E)-	56052-80-3	0.0476	0.93	0.0399	1.66	0.0451	1.88	0.991	0.985

Table S2 Reproducibility and stability testing results for 50 substances in LC-MS. Stability test results were performed on Lot 27, stored dry at -20°C

		Reproducibility/Homogeneity				Stability: Day X / Day 1		
		Over 3 injections		Over 3 aliquots		X=7	X=15	X=48
Name	CAS	Avg. Peak Area	CV [%]	Avg. Peak Area	CV [%]	Rel. Conc.	Rel. Conc.	Rel. Conc.
Positive Compounds								
Desethylatrazine	6190-65-4	5.65E+5	10.9	5.67E+5	10.0	0.792	0.967	1.351
Terbutylazine	5915-41-3	1.10E+6	2.7	1.15E+6	1.9	0.750	0.929	1.104
Atrazine	1912-24-9	6.79E+5	3.2	6.72E+5	3.2	0.968	1.015	1.676
Prometryn	7287-19-6	5.12E+5	15.3	5.65E+5	7.3	1.138	1.098	1.503
Simazine	122-34-9	4.74E+5	10.0	4.66E+5	15.5	0.817	1.130	1.678
Tebuconazole	80443-41-0	7.09E+4	15.1	6.80E+4	18.5	1.346	0.777	2.100
Phenazone	60-80-0	1.33E+6	2.5	1.35E+6	5.9	1.064	1.093	1.445

		Reproducibility/Homogeneity				Stability: Day X / Day 1		
		Over 3 injections		Over 3 aliquots		X=7	X=15	X=48
Name	CAS	Avg. Peak Area	CV [%]	Avg. Peak Area	CV [%]	Rel. Conc.	Rel. Conc.	Rel. Conc.
Propyphenazone	479-92-5	3.65E+5	9.5	3.84E+5	3.6	0.872	1.075	1.617
Sulfamethoxazole	129378-89-8	4.40E+5	5.5	4.75E+5	2.8	0.841	0.606	1.177
Caffeine	58-08-2	4.01E+6	3.8	4.38E+6	19.2	1.060	1.002	1.637
Carbamazepine	298-46-4	2.31E+6	7.7	2.33E+6	4.8	0.934	1.254	1.279
Lidocaine	137-58-6	3.46E+5	7.9	3.43E+5	4.4	0.647	0.781	1.156
Tramadol	27203-92-5	2.64E+6	7.4	2.59E+6	3.0	0.934	1.047	1.228
10,11-Dihydro-10,11-dihydroxy-carbamazepine	298-46-4	1.06E+6	6.8	1.11E+6	2.7	0.738	0.870	1.179
4-Formylaminoantipyrine	1672-58-8	6.19E+6	3.7	6.12E+6	2.3	0.908	1.026	1.431
Phenylbenzimidazole SA	27503-81-7	6.07E+5	8.6	6.56E+5	10.0	0.498	0.707	0.920
Tris(2-chloroethyl) PO ₄	115-96-8	9.21E+4	14.5	9.23E+4	11.3	0.777	0.797	1.235
DEET	134-62-3	3.22E+6	4.7	3.34E+6	2.4	0.994	0.961	1.212
1H-Benzotriazole	27556-51-0	1.61E+7	3.1	1.66E+7	2.3	0.857	0.979	1.309
5-Methyl-1H-benzotriazole	136-85-6	2.07E+7	3.8	2.09E+7	2.7	0.842	0.967	1.276
p-Toluenesulfonamide	70-55-3	4.37E+5	9.1	4.28E+5	3.9	0.931	1.061	0.986
Tributylphosphate	126-71-6	2.33E+6	6.7	2.44E+6	12.2	0.784	0.952	1.169
Tris(1-chloro-2-propyl)PO ₄	26248-87-3	2.73E+7	5.3	2.71E+7	9.1	0.922	1.034	0.975
Hexa(methoxymethyl) melamine	3089-11-0	5.84E+6	8.1	5.94E+6	10.2	0.978	1.093	1.201
Triphenylphosphine oxide	791-28-6	1.66E+7	4.6	1.75E+7	3.5	0.925	1.083	1.324
2-(Methylthio) benzothiazole	615-22-5	4.44E+5	9.6	4.06E+5	13.6	0.557	0.492	0.622
Benzothiazole	95-16-9	2.44E+6	13.3	2.40E+6	16.3	0.866	0.984	1.098
Icaridin	119515-38-7	2.27E+6	1.0	2.15E+6	4.4	1.086	1.208	1.287
Metformin	1115-70-4	3.70E+6	13.6	3.60E+6	6.7	0.868	1.109	1.365
Benzyl dimethyl dodecylammonium	139-07-1	1.34E+6	7.2	1.37E+6	3.8	1.068	1.199	0.848
Lauryl diethanolamide	120-40-1	2.78E+5	2.7	3.20E+5	10.0	1.080	1.178	1.544
Lauramidopropylbetaine	4292-10-8	5.37E+6	3.1	5.59E+6	2.3	0.736	1.383	1.029
Negative compounds								
Bentazone	530-57-4	2.44E+5	4.0	2.77E+5	8.1	0.886	1.269	1.262
Metolachlor OA	152019-73-3	6.14E+4	12.6	7.04E+4	12.4	0.631	1.355	1.554
Metolachlor ESA	171118-09-5	3.06E+5	4.4	3.19E+5	9.4	0.878	1.083	1.569
Phenylbenzimidazole SA	27503-81-7	2.71E+5	3.6	2.83E+5	7.2	0.877	0.973	1.471
1H-Benzotriazole	27556-51-0	1.71E+5	14.6	1.63E+5	13.6	0.849	1.083	2.163
Benzophenone-4	582-33-2	2.40E+5	11.3	2.47E+5	7.8	0.969	0.971	1.558
2-Naphthalenesulfonic acid	120-18-3	3.26E+5	0.9	3.10E+5	5.6	0.881	1.264	1.485
2-Benzothiazole SA	941-57-1	2.29E+6	1.7	2.38E+6	3.7	0.909	0.917	1.309
p-Nitrophenol	100-02-7	2.21E+6	2.6	2.22E+6	3.7	0.957	0.913	1.059
2,4-Dinitrophenol	51-28-5	5.15E+5	4.0	5.00E+5	7.1	1.234	0.864	1.151
Acesulfame	33665-90-6	6.91E+5	4.8	6.84E+5	4.9	0.928	0.980	1.634
Sucralose	56038-13-2	1.94E+5	8.9	2.01E+5	8.4	0.920	0.830	1.163
Cyclamate	139-05-9	3.51E+5	3.1	3.53E+5	2.1	0.927	1.118	1.523
Aspartame	22839-47-0	8.83E+4	15.4	8.50E+4	10.4	1.166	0.678	1.431
Perfluorooctanoic acid	335-67-1	4.16E+5	10.7	4.04E+5	9.8	0.758	0.972	1.403
Perfluorohexanoic acid	307-24-4	2.30E+5	9.4	2.50E+5	4.2	0.876	0.952	1.571
Dodecylsulfate	151-21-3	7.52E+7	1.2	7.48E+7	1.5	0.942	1.054	1.374

SA= sulfonic acid. PO₄=phosphate.

Table S3 Summary of GC Analytical Methods, Part 1 (Instrument, column, solvent, injection, flow)

Parti.	Instrument and Model	Column	Dimensions [m, mm, μ m]	Solvent	Inj. type (vol [μ L])	Flow [mL/min]
G1	Agilent 5795C	J&W/DB-5ms	30;0.25;0.25	ACN	LVI (150)	1.8
G2	Agilent 5975C	J&W/DB-1MS	60;0.25;0.25	DCM	LVI (50)	1.3
G3	Agilent 6890 LECO Pegasus 4D	Restek/Rxi-5-SIL MS, Rxi 17-SIL MS	30;0.25;0.25 2.5;0.15;0.15	TOL	SL (1)	1
G4a	Agilent 6890n	HP-5MS	30;0.25;0.25	TOL	PSL (1)	1.8
G4b	Thermo Trace GC Ultra- TSQ Quantum Access	TR-Pesticide II	30;0.25;0.25	TOL	LVI (50)	0.7-2.2
G5a	Agilent 6890A GCT	J&W/DB-5ms	30;0.25;0.25	HEX	SL (1)	1
G5b	Agilent 7890A Waters Xevo GC2 QTOF	J&W/DB-5ms	30;0.25;0.25	HEX	PSL (1)	2
G6	Agilent 6890 LECO Pegasus 4D	SGE BPX50 Agilent VF-1ms	30;0.25;0.25 1.5;0.1;0.1	DCM	SL (1)	1.0
G7	Agilent 7890 LECO Pegasus 4	Restek/Rxi-5Sil MS and Rxi 17	30;0.25;0.25 1.1;0.1;0.1	DCM	LVI (50)	1

ACN = acetonitrile, DCM = dichloromethane, TOL = toluene, HEX = hexane, LVI = large volume injection, SL = splitless injection, PSL = pulsed splitless injection. The carrier gas was helium for all, except EI with hydrogen.

Table S4 Summary of GC Analytical Methods Part 2(temperature, ionization, scan range) and Software

Parti.	Temperature Program	Ionization	Scan Range	RI	Target Software	Non-target Software, Evidence
G1	50°C(3)-30°C/min- 200°C(0)-5°C/min- 280°C(0)-30°C/min- 320°C(10)	EI, PCI, NCI	50-800	KRI	ChemStation	ChemStation, AMDIS, NIST, manual interp., retrospective information
G2	35°C(2.5)-11.4°C/min- 320°C(15)	EI	35-375	KRI	ChemStation	AMDIS, NIST. MS match
G3	40°C(2)-5°C/min- 320°C(10). 2 nd col +10°C	EI	40-600	None		NIST
G4a	65°C(1)-15°C/min- 320°C(3)	EI	45-450	KRI	ChemStation	Chemstation, XCMS, AMDIS, NIST 08. Library search, manual interpret.
G4b	90°C(7)-25°C/min- 180°C(0)-5°C/min- 275°C-10°C/min-290°C- 2.5°C/min-320°C(8.9)	EI	50-500	KRI	QualBrowser	Xcalibur, XCMS, AMDIS, NIST 08. Library search, manual interp.
G5a	90°C(1)-5°C/min- 275°C(0)-40°C/min- 320°C(2)	EI	50-650	KRI	Targetlynx, Chromalynx	Chromalynx, NIST. MS match, fragments and accurate mass
G5b	90°C(1)-5°C/min- 275°C(0)-40°C/min- 320°C(2)	APCI	50-650	KRI	Chromalynx	Chromalynx, NIST. MS match, fragments and accurate mass
G6	35°C(1)-5°C/min- 320°C(5), 2 nd col +20°C	EI	40-800	KRI	ChromaTof	NIST, Match factor >800, manual check; position on GCxGC
G7	50°C(2) - 25°C/min- 80°C(0)-4°C/min- 290°C(10). 2nd col +20°C	EI	50-700	KRI	ChromaTOF and Excel (in-house method)	NIST, Match factor >700, detected in $\geq 2/3$ replicates

EI=Electron impact, PCI/NCI = positive/negative chemical ionization, APCI = atmospheric pressure chemical ionization. MS = mass spectrum. (K)RI = (Kovats) retention index.

Table S5 List of suspected substances provided for the collaborative trial according to previous surveys (JDS2, JDS3) associated with the given sampling point

#	Compounds/Group	SMILES	CAS
Phosphates			
1.	Tributyl phosphate	<chem>CCCCOP(=O)(OCCCC)OCCCC</chem>	126-73-8
2.	Triphenyl phosphate	<chem>C1=CC=C(C=C1)OP(=O)(OC2=CC=CC=C2)OC3=CC=CC=C3</chem>	115-86-6
Phthalates			
3.	Diethyl Phthalate	<chem>CCOC(=O)C1=CC=CC=C1C(=O)OCC</chem>	84-66-2
4.	Dibutyl Phthalate	<chem>CCCCOC(=O)C1=CC=CC=C1C(=O)OCCCC</chem>	84-74-2
5.	DEHP	<chem>CCCCC(CC)COC(=O)C1=CC=CC=C1C(=O)OCC(CC)CCCC</chem>	117-81-7
Pharmaceuticals			
6.	Naproxen	<chem>C[C@@H](C1=CC2=C(C=C1)C=C(C=C2)OC)C(=O)O</chem>	22204-53-1
7.	Ibuprofen	<chem>CC(C)CC1=CC=C(C=C1)C(C)C(=O)O</chem>	15687-27-1
8.	Caffeine	<chem>CN1C=NC2=C1C(=O)N(C(=O)N2C)C</chem>	58-08-2
9.	Carbamazepine	<chem>C1=CC=C2C(=C1)C=CC3=CC=CC=C3N2C(=O)N</chem>	298-46-4
10.	Sulfamethoxazole	<chem>CC1=CC(=NO1)NS(=O)(=O)C2=CC=C(C=C2)N</chem>	723-46-6
Industrial Chemicals			
11.	Perfluoroheptanoate	<chem>FC(F)(C(F)(F)C([O-])=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F</chem>	375-85-9
12.	Perfluorooctanoate	<chem>FC(F)(C(F)(F)C([O-])=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F</chem>	335-67-1
13.	Perfluorooctan-sulfonate	<chem>C(S(=O)(=O)O)(F)(F)C(F)(F)C(C(C(C(C(F)(F)F)(F)F)(F)F)(F)F)(F)F</chem>	29081-56-9
14.	Bisphenol A	<chem>CC(C)(C1=CC=C(C=C1)O)C2=CC=C(C=C2)O</chem>	80-05-7
15.	1-Decanamine, N,N-dimethyl-	<chem>N(CCCCCCCCC)(C)C</chem>	1120-24-7
16.	Nonylphenol monocarboxylates	<chem>CCCCCCCCc1ccc(cc1)OCC(=O)O</chem>	3115-49-9
Pesticides			
17.	Bentazone	<chem>CC(C)N1C(=O)C2=CC=CC=C2NS1(=O)=O</chem>	25057-89-0
18.	2,4-D	<chem>C1=CC(=C(C=C1)Cl)OCC(=O)O</chem>	94-75-7
19.	Desethylatrazine	<chem>CC(C)NC1=NC(=NC(=N1)N)Cl</chem>	6190-65-4
20.	Atrazine	<chem>CC/N=C/1\N/C(=N/C(C)C)/NC(=N1)Cl</chem>	1912-24-9
21.	Isoproturon	<chem>CC(C)C1=CC=C(C=C1)NC(=O)N(C)C</chem>	34123-59-6
22.	Diuron	<chem>CN(C)C(=O)NC1=CC(=C(C=C1)Cl)Cl</chem>	330-54-1
23.	Terbutylazine	<chem>CCNC1=NC(=NC(=N1)Cl)NC(C)(C)C</chem>	5915-41-3
24.	Desethylterbutylazine	<chem>CC(C)(C)NC1=NC(=NC(=N1)N)Cl</chem>	30125-63-4

Table S6 Summary of GC-MS results by participant. ^qIndicates (semi-)quantified results

#	G1	G2	G3	G4	G5	G6	G7
Target	9 ^q	-	-	8 ^q	6	41	-
Suspect/Library Match	66 ^q	40	15 ^q	10	12	121	31
Isomeric mix	-	-	-	-	-	47	-
Unknown	31	-	-	-	5	-	-