Environmental Pollution 234 (2018) 297–306

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Suspect screening and non-targeted analysis of drinking water using point-of-use filters☆

Seth R. Newton [a, *], Rebecca L. McMahen [a, b], Jon R. Sobus [a], Kamel Mansouri [b, c, 1], Antony J. Williams [c], Andrew D. McEachran [b, c], Mark J. Strynar [a]

[a] *United States Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC 27709, United States*
[b] *Oak Ridge Institute for Science and Education Research Participant, 109 T.W. Alexander Drive, Research Triangle Park, NC 27709, United States*
[c] *United States Environmental Protection Agency, National Center for Computational Toxicology, Research Triangle Park, NC 27709, United States*

ARTICLE INFO

Article history:
Received 27 June 2017
Received in revised form
7 November 2017
Accepted 8 November 2017

Keywords:
Drinking water
Exposome
Suspect screening
Non-target analysis
High resolution mass spectrometry

ABSTRACT

Monitored contaminants in drinking water represent a small portion of the total compounds present, many of which may be relevant to human health. To understand the totality of human exposure to compounds in drinking water, broader monitoring methods are imperative. In an effort to more fully characterize the drinking water exposome, point-of-use water filtration devices (Brita® filters) were employed to collect time-integrated drinking water samples in a pilot study of nine North Carolina homes. A suspect screening analysis was performed by matching high resolution mass spectra of unknown features to molecular formulas from EPA's DSSTox database. Candidate compounds with those formulas were retrieved from the EPA's CompTox Chemistry Dashboard, a recently developed data hub for approximately 720,000 compounds. To prioritize compounds into those most relevant for human health, toxicity data from the US federal collaborative Tox21 program and the EPA ToxCast program, as well as exposure estimates from EPA's ExpoCast program, were used in conjunction with sample detection frequency and abundance to calculate a "ToxPi" score for each candidate compound. From ~15,000 molecular features in the raw data, 91 candidate compounds were ultimately grouped into the highest priority class for follow up study. Fifteen of these compounds were confirmed using analytical standards including the highest priority compound, 1,2-Benzisothiazolin-3-one, which appeared in 7 out of 9 samples. The majority of the other high priority compounds are not targets of routine monitoring, highlighting major gaps in our understanding of drinking water exposures. General product-use categories from EPA's CPCat database revealed that several of the high priority chemicals are used in industrial processes, indicating the drinking water in central North Carolina may be impacted by local industries.

1. Introduction

Safe drinking water supplies are critical for public health and it has been estimated by the World Health Organization (WHO) that a 10% reduction in worldwide disease could be achieved by improvements related to drinking water alone, including sanitation, hygiene, and water resource management (Prüss-Üstün et al., 2008). Furthermore, it is estimated that 70–90% of disease risks are due to differences in environments (Rappaport and Smith, 2010), which includes direct exposures via consumption of drinking water. Chemicals that are present in water supplies can increase risk for disease and adverse health outcomes over long-term exposure periods (WHO, 2013). It has been demonstrated for various chemical classes, including perfluorinated chemicals, that drinking water can be one of the most important pathways for human exposure (Egeghy and Lorber, 2011; Lorber and Egeghy, 2011). Even so, it has been estimated that only 40% of US consumers used any kind of water purification device in 2014 (Anumol et al., 2015). Certain chemicals are regulated under the Safe Drinking Water Act, but these chemicals constitute only a small fraction of the number of chemicals present in drinking water (US EPA, 2016). New compounds can be added to this list if they are

---

☆ This paper has been recommended for acceptance by Maria Cristina Fossi.
* Corresponding author. 109 TW Alexander Dr., Durham, NC 27709, United States.
E-mail address: newton.seth@epa.gov (S.R. Newton).
[1] Current Affiliation: Scitovation LLC, Research Triangle Park, NC 27709, United States.

https://doi.org/10.1016/j.envpol.2017.11.033
0269-7491/Published by Elsevier Ltd.

discovered and deemed to pose a threat to human health. These additions, however, require developing and validating "targeted" methods, which is a slow and expensive process. Furthermore, this process requires some *a priori* knowledge of the compounds for which methods should be developed. As of yet, there is no reliable mechanism to identify and prioritize novel compounds. There are needs, then, for: 1) a more complete picture of chemical exposures via drinking water consumption; 2) methods of rapidly identifying emerging chemicals that may be of importance to human health; and 3) means with which to properly assess exposure-disease relationships and risks to human health (Villanueva et al., 2014).

Recent advances in analytical techniques have led to the detection of various contaminants in water which would have otherwise gone undetected using traditional targeted methods (Schymanski et al., 2015; Strynar et al., 2015). These advanced techniques often employ high resolution mass spectrometry (HRMS), or tandem HRMS, to either match unknown sample features to compounds within spectral and/or spectra-less databases (a technique known as suspect screening analysis [SSA]), or elucidate structures of unknowns that may not be contained in a database (a technique known as non-targeted analysis [NTA]). While these two techniques differ, they are often discussed together as they are complimentary to each other. SSA/NTA workflows are rapidly evolving, and are becoming more frequently used to detect differences (or similarities) between two or more groups of samples in case-control style experiments. Example applications include: detecting a chemical spill in a river after a baseline chemical signature has been established (Bader et al., 2016); evaluating the contribution of various tributaries to a river (Ruff et al., 2015); or singling out unknown features that appear in landfill leachate and in downstream drinking water (Müller et al., 2011).

SSA/NTA approaches may also be applied to environmental samples in support of general monitoring – that is, to broadly screen for the occurrence of chemicals in a selected medium. The ability to rapidly identify unknown compounds during routine monitoring is essential to fully explore the exposome, defined as the sum of all exposures (exogenous and endogenous) for an individual over a lifetime (Wild, 2005). In order to sequence the exposome, it is useful and necessary, from an analytical standpoint, to compartmentalize exposures by matrix. Examples of monitoring studies that focus on a specific matrix can be found for dust (Rager et al., 2016), river water (Schymanski et al., 2015), waste water (Schymanski et al., 2014b), etc. but drinking water remains relatively unexplored with regards to SSA/NTA. This is somewhat surprising, as drinking water is a fairly simple matrix to which humans are exposed in similar amounts, in contrast to dust or waste water, which require clean-up steps after extraction, and for which exposure amounts are not well known.

When applied to environmental and biological samples, SSA/NTA methods have the potential to allow rapid chemical characterization without the need for standards or *a priori* knowledge of sample constituents. Confidence in the identification of unknowns can be communicated in terms of levels outlined by Schymanski et al. (2014a), where the highest level of confidence (level 1) requires confirmation by an analytical standard, and the next level of confidence (level 2) requires evidence for a probable structure. A goal for researchers using SSA/NTA methods should be to confidently classify as many unknowns as possible into level 2, and not necessarily level 1, as it is not practical, or even possible, to confirm all unknowns with analytical standards. Chemicals of highest concern can then be confirmed with standards, if possible, and categorized into level 1. Confidence in level 2 identifications will most likely come about through the development of several different tools that build increasing confidence of positive detection. As we are in the early years of a burgeoning exposomics field, researchers must find ways to prioritize unknowns into those that they believe are most likely to be relevant to human and environmental health (Sobus et al., 2017). Recently, a method to prioritize the vast number of unknowns in a sample by incorporating toxicity and exposure information was presented by Rager et al. (2016). We have sought to apply this method to drinking water in the Raleigh/Durham/Chapel Hill area of North Carolina, United States, and improve upon it using tools and data available from EPA's CompTox Chemistry Dashboard (hereafter referred to as "the Dashboard", https://comptox.epa.gov/dashboard), a newly developed web application that supports SSA/NTA workflows (McEachran et al., 2017b). We have also sought to demonstrate that SSA/NTA methods can rapidly identify contaminants in drinking water that are not routinely monitored and would likely go undetected without these methods.

2. Materials and methods

2.1. Materials

Information about the materials used in this study can be found in the Supporting Information (SI).

2.2. Sample collection

Samples were collected in a pilot scale study by installing a Brita® Basic Faucet Filter in the homes of nine North Carolina residents. Provided in the SI is a list of chemicals that Brita® Basic Faucet Filters are known to remove from drinking water (SI, Table S1), as well as a table of organic chemicals included in the Safe Drinking Water Act (Table S2). Some residents received drinking water from their local municipalities, while other residents received their drinking water from a private well. Information about the water source and municipality can be found in Table 1. Although the samples are labeled by location, many of the drinking water treatment facilities report purchasing water from other facilities so it is possible the sampling location is not fully indicative of the original drinking water source. The study participants were asked to use the filter for cold water during everyday use until the indicator light on the filter turned red, signaling that the filter was at its maximum capacity. This process took between 1 and 4 months for each sample with an average sampling time of 68 days. The participants were asked to return their filters for analysis upon seeing the red indicator light.

2.3. Sample extraction and processing

The filter was removed from the plastic casing using a band saw with a clean blade and placed into a plastic bag for storage until extraction. The filters were individually lyophilized for three days to remove any water which remained in the filter pores. The filters

Table 1
Sample information.

Sample #	Location	Source Type	Population Served
1	Durham	Municipal	265,472
2	Durham	Municipal	265,472
3	Apex	Municipal	46,831
4	Cary	Municipal	182,088
5	Chapel Hill	Municipal	83,300
6	Chapel Hill	Private Well	–
7	Raleigh	Municipal	540,000
8	Pittsboro	Municipal	4,401
9	Pittsboro	Private Well	–