Published in final edited form as:

Environ Sci Process Impacts. 2019 June 19; 21(6): 988–998. doi:10.1039/c9em00112c.

Surface water and groundwater analysis using aryl hydrocarbon and endocrine receptor biological assays and liquid chromatography-high resolution mass spectrometry in Susquehanna County, PA†

Michelle Bamberger^a, Marika Nell^b, Ahmed H. Ahmed^c, Renee Santoro^d, Anthony R. Ingraffea^b, Rana F. Kennedy^e, Susan C. Nagel^e, Damian E. Helbling^b, Robert E. Oswald^c

^aVet Behavior Consults, Ithaca, NY, USA

^bSchool of Civil & Environmental Engineering, Cornell University, Ithaca, NY, USA

^cDepartment of Molecular Medicine, Cornell University, Ithaca, NY, USA.

^dPSE Healthy Energy, USA

^eDepartment of Obstetrics, Gynecology and Women's Health, University of Missouri, Columbia, MO, USA

Abstract

The contamination of surface water and ground water by human activities, such as fossil fuel extraction and agriculture, can be difficult to assess due to incomplete knowledge of the chemicals and chemistry involved. This is particularly true for the potential contamination of drinking water by nearby extraction of oil and/or gas from wells completed by hydraulic fracturing. A case that has attracted considerable attention is unconventional natural gas extraction in Susquehanna County, Pennsylvania, particularly around Dimock, Pennsylvania. We analyzed surface water and groundwater samples collected throughout Susquehanna County with complementary biological assays and high-resolution mass spectrometry. We found that Ah receptor activity was associated with proximity to impaired gas wells. We also identified certain chemicals, including disclosed hydraulic fracturing fluid additives, in samples that were either in close proximity to impaired gas wells or that exhibited a biological effect. In addition to correlations with drilling activity, the biological assays and high-resolution mass spectrometry detected substances that arose from other anthropogenic sources. Our complementary approach provides a more comprehensive picture of

†Electronic supplementary information (ESI) available. See DOI: [10.1039/c9em00112c](https://doi.org/10.1039/c9em00112c)

reol@cornell.edu.

Contributions

M. B.—initial design of study, obtained funding, recruited participants, follow-up communication with participants, collected samples, participated in writing manuscript. M. N.—designed, performed, and analysed LC-HRMS experiments, developed comparison of LC-HRMS data with biological assays and locations of impaired wells, participated in writing manuscript. A. H. A.—designed, performed and analysed Ah receptor assays, participated in writing manuscript. R. F. K.—designed, performed and analysed endocrine receptor assays. R. S. and A. R. I.—updated and provided information on impaired wells. S. C. N.—Designed and analysed endocrine receptor assays, participated in writing manuscript. D. E. H.—designed and analysed LC-HRMS experiments, developed comparison of LC-HRMS data with biological assays and locations of impaired wells, participated in writing manuscript. R. E. O.—initial design of study, obtained funding, collected samples, analysed data, prepared figures, participated in writing manuscript.

Conflicts of interest

There are no conflicts to declare.

water quality by considering both biological effects and a broad screening for chemical contaminants.

1. Introduction

High volume hydraulic fracturing (HVHF) activities have boomed in Susquehanna County, PA, which overlies the organic-rich Marcellus Shale.¹ This boom has raised concerns about water contamination from these activities, which use a mix of water, sand, and chemicals known as hydraulic fracturing fluid (HFF) to enable the extraction of natural gas. The chemical additives used in HFF include biocides, surfactants, and corrosion inhibitors, all of which raise concerns about environmental and health impacts in the event of spills or other means of water contamination.^{2–6} The environmental, animal, and human health impacts of HFF additives are of particular concern in Susquehanna County, where studies have documented evidence of methane, inorganic compounds, and gasoline or diesel range organic compounds in groundwater samples collected in the area.^{1,7–10}

Water quality monitoring for HFF additives is complicated by the large number and chemical diversity of the additives used for well completion and maintenance. Over one thousand unique chemicals have been disclosed as HFF additives or detected in HVHF wastewaters.^{3,11} In addition, all monitoring is done in the context of other possible deleterious anthropogenic contributions to water quality, such as agricultural activity and household wastewater. The identification of potential pollutants in water samples is often done by screening for target compounds that may or may not be relevant choices and may exclude important contaminants. Two very different but complementary approaches have been employed recently to cast a wider net and possibly provide a more comprehensive picture of the biological consequences of those pollutants and the presence of the pollutants that may be responsible. The first employs biological assays to assess the effects of water samples on specific target receptor proteins expressed in mammalian or yeast cells. Biological activity was assessed by measuring agonist or antagonist activity of the extracted water sample to the aryl hydrocarbon (Ah) receptor expressed in yeast cells and the estrogen (ER), androgen (AR), progesterone (PR), and glucocorticoid (GR) receptors expressed in mammalian cells. The Ah receptor is a transcription factor that regulates gene expression in response to a variety of compounds, in particular aromatic hydrocarbons. The ER, AR, PR, and GR activities were monitored to assess endocrine disruption. The second employs liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) to screen for a wide range of suspect chemicals in water samples. This approach has some similarity to effect-directed analysis where samples that exhibit a biological effect are tested further to identify the causal chemical agent.¹² In this case, we assess potential toxicity and endocrine activity with biological assays and chemical composition in bulk, and then screen those bulk characterizations for associations with anthropogenic activities.

Concerns over the presence of contaminants in water sources are ultimately a question of whether health effects are produced by exposure to water of humans or animals (food, companion, and wildlife). The EPA and others have developed a series of high throughput assays based on an adverse outcome pathway approach to test for endocrine activity.¹³ This

approach can be used to determine whether water sample extracts can activate or inhibit endocrine receptors expressed in mammalian or yeast cells and, coupled with reporter systems, to provide a rapid and relatively inexpensive test for bioactivity. In particular, this has been accomplished for endocrine receptors^{14–16} and the aryl hydrocarbon receptor (Ah).^{17,18}

Considering the physicochemical properties of disclosed HFF additives, a majority (>900) are expected to be amenable to analysis by liquid chromatography mass spectrometry (LC-MS).¹⁹ Further, the large number of potential analytes and the limited availability of authentic reference standards suggest that suspect screening by means of high-resolution mass spectrometry (HRMS) may be an ideal approach to at least tentatively identify specific chemicals and assess water quality impacts of HVHF activities.²⁰ Suspect screening leverages the high mass accuracy and resolving power of HRMS to screen full-scan mass spectral acquisitions for MS features that match the expected exact mass of lists of suspect chemicals.²¹ The presence of resulting suspect hits can be further confirmed through data prioritization²¹ and careful inspection of MS (isotopic patterns) and MS2 spectra.²² Suspect screening of LC-HRMS acquisitions has been previously used to identify suspect chemicals in surface water,²³ lake sediments,²⁴ and a variety of other complex matrices.^{25–27} Only a few studies have used LC-HRMS and suspect screening to assess water quality in samples collected in areas surrounded by HVHF activities.²⁸

The objective of this study was to assess the water quality impacts of hydraulic fracturing operations on surface water and groundwater in Susquehanna County in 2015 by exploring the biological activity and organic chemical composition. Fifty-three water samples were collected from surface water and groundwater sites in locations surrounded by HVHF activities, agriculture, and mixed-use residential and commercial areas. The proximity of sampling sites to HVHF wells with some evidence of impairment (reported compromised structural integrity of casing and cement)²⁹ was determined using the GPS coordinates of the sampling site and the impaired gas well. To identify organic chemicals, we measured extracts of the water samples by means of LC-HRMS and used a conventional approach to screen for 218 anthropogenic organic chemicals including agricultural chemicals, wastewater-derived chemicals, and industrial chemicals. We also used suspect screening to search for 732 chemicals that have been reported as HFF additives in previous publications.

2. Methods

2.1 Sample information

Fifty-three surface water and groundwater samples were collected in one liter amber glass bottles from Susquehanna County, Pennsylvania in 2015. One liter samples were prepared by filtration to remove any insoluble particles followed by solid-phase extraction using Oasis HLB cartridges (5 cm³, 200 mg, Waters, Milford MA) within two weeks of collection. HLB cartridges are designed to extract both hydrophilic and lipophilic compounds. After loading the sample, the column was washed with 6 ml of 5% v/v methanol (HPLC grade) followed by elution with 6 ml of methanol into amber glass vials. Eluents were evaporated and reconstituted in 500 µl of methanol. Sample extracts or pure methanol (control) were used to assess biological activity and approximately 50 µl of each sample was provided for analysis

by means of LC-HRMS. No activity was observed in the methanol control for any of the biological assays. Sample extracts were stored at -20°C until analysis. All assays and chemical analyses were performed blinded of the sample location and identity. Because the biological assays did not target specific compounds and the LC-HRMS screening included hundreds of compounds, it was impractical to obtain extraction efficiencies for all relevant compounds. Thus, the subsequent analytical methods used would have to be considered qualitative or semi-quantitative.

2.2 Ah receptor activity

The *Saccharomyces cerevisiae* YCM3 strain of yeast (ATCC strain MYA-3637 (ref. 18)) was used for the AhR-signaling assay. Cells were grown in a synthetic minimal yeast nitrogen base (YNB) medium (without amino acids and with ammonium sulfate and glucose; adenine sulfate, histidine, leucine, and uracil were added to correct for auxotrophic deficiencies) overnight to saturation (OD_{600} 1.5). Cells were diluted in $10\text{--}40\ \mu\text{l ml}^{-1}$ YNB medium containing galactose instead of glucose to induce protein expression. In a 96 well plate, $95\ \mu\text{l}$ of the culture was combined with $5\ \mu\text{l}$ of the test sample and incubated for 17 h at 25°C with shaking at 300 rpm. The Promega Beta-Glo Assay System kit was used to detect Ah receptor activity at 25°C . To each well, $100\ \mu\text{l}$ of the beta galactosidase (GAL) reagent was added and then incubated and shaken for 30 min at 25°C . The luminescence was then read with a BioTek Synergy 2 Multi-Mode Microplate Reader.

Assays were repeated three times with three replicates in each. The response was divided into four response categories based on the fold-increase above background. The response categories were defined based on the most concentrated dilution as follows: category 1: 1.8 to 2-fold above background; category 2: 2- to 3-fold above background; category 3: 3- to 4-fold above background; category 4: more than 4-fold above background. In each case, the response increased monotonically from the highest to lowest dilution. Background was defined as the response of cells treated with the carrier solvent ($5\ \mu\text{l}$ of methanol; average of 9 replicates). Overall, four Fisher Exact Probability Tests were performed so that for an α of 0.05, the level used to judge statistical significance was 0.0125 (Bonferroni correction).

2.3 Endocrine receptor activity

Reconstituted samples were stored at -20°C , and protected from light until tested. To be applied to cells, stock samples were diluted 100-fold or 400-fold in tissue culture medium, creating final concentrations, in contact with mammalian cells, of 20 times ($20\times$) or five times ($5\times$) the original water concentration for mammalian assays. Ishikawa human endometrial cancer cells (Sigma cat # 99040201, Sigma-Aldrich Co., St. Louis, MO) were maintained and transiently transfected with plasmids: 3XEREtk-luciferase for anti-estrogenic assays (this assay utilized an endogenous estrogen receptor), the androgen receptor (CMVAR1) and PSA-Enh E4TATA-luc for anti-androgenic activity, progesterone receptor B (pcDNA3 PRB) and the progesterone/glucocorticoid response element linked to the firefly luciferase gene (MMTV-luc) for anti-progestogenic activity, or glucocorticoid receptor (pRST7-GR) and MMTV-luc as described previously.^{15,16} Cells were incubated with a dilution series of 17β -estradiol and ICI 182780 (the positive agonist and antagonist controls used for anti-estrogenic assays), dihydrotestosterone and flutamide (the positive

agonist and antagonist controls used for anti-androgenic assays), progesterone and mifepristone (RU-486) (the positive agonist and antagonist controls used for anti-progestogenic assays), and dexamethasone and mifepristone (RU-486) (the positive agonist and antagonist controls used for anti-glucocorticogenic assays), vehicle (1% methanol), agonist or the water sample extracts as described previously.^{15,16} Each sample was tested in triplicate within each assay; androgen, estrogen, and glucocorticoid receptor assays were repeated two times and the progesterone assay was repeated three times. Activity of test samples was calculated as a ratio of relative light units (RLU) of test sample and added agonist activity: 20 pM 17 β -estradiol for anti-estrogenic activity, 500 pM dihydrotestosterone for anti-androgenic activity, 70 pM dexamethasone for anti-glucocorticogenic activity and 50 pM progesterone for anti-progestogenic activity and compared with the responses to control agonists and/or antagonists. The CellTiter 96 nonradioactive cell proliferation assay (Promega; #G4000, Promega Corporation, Madison, Wisconsin) was used to assess cell toxicity in Ishikawa cells as previously described.¹⁵ No samples showed toxicity at 5 $\times$ the original water sample concentration. Three samples from surface ponds showed toxicity at 20 $\times$ and all subsequent tests for those three samples were conducted at 5 $\times$.

2.4 LC-HRMS

2.4.1 Sample preparation.—Sample extracts were fully evaporated in a laminar flow hood and reconstituted in LC-MS grade water at three-times the original volume provided. The sample extracts were vortexed, centrifuged, and transferred to 180 μ l glass inserts with 1.8 ml glass LC-MS vials. The sample extracts were then spiked with a mixture of 48 isotope labelled internal standards (ILISs) as detailed in previous work.³⁰ The volume used for each spike was adjusted so that each sample extract would contain 100 μ g l⁻¹ of each ILIS. Three blind samples were prepared by evaporating 50 μ l aliquots of LC-MS grade methanol in a laminar flow hood, reconstituting them with 150 μ l of LC-MS grade water, and spiking with the same mix and concentration of ILISs. No phthalates were measured in the blank/blind samples.

2.4.2 Analytical method.—A previously reported²¹ analytical method using high-performance LC coupled to a Q Exactive (quadrupole-orbitrap) high-resolution mass spectrometer was used with some modifications. The mobile phase consisted of LC-MS grade water and LC-MS grade methanol, each amended with 0.1% formic acid (v/v). Samples were injected at 20 μ l on an XBridge C-18 analytical column (2.1 $\times$ 50 mm, particle size 3.5 μ m) at 25 $^{\circ}$ C. The samples were divided into two separate sequences during the data acquisition phase. Prior to each run and in the middle of each run, mass calibrations and accuracy checks were performed; mass accuracy was always within $\pm$ 1 ppm. Each sample was analyzed in triplicate using positive mode heated electrospray ionization. Full-scan MS and data dependent top N = 3 MS2 acquisitions were obtained for each sample extract using a retention time dependent exclusion list containing peaks previously measured in the three blank samples. After suspect screening, samples were re-measured with a data dependent inclusion list containing the exact masses of all suspect hits if the original acquisition failed to collect relevant MS2 data.

2.4.3 Target screening.—The MS data was processed using a conventional target screening method for 218 anthropogenic organic chemicals including agricultural chemicals (*e.g.*, herbicides, insecticides, fungicides), wastewater-derived chemicals (*e.g.*, pharmaceuticals, personal care products, food additives, illicit drugs), and other industrial chemicals (including some used in HVHF operations).³¹ Each of the 218 target chemicals was assigned a surrogate ILIS with a similar retention time. Details of the analytes included can be found in previous studies.^{30,32}

2.4.4 Extraction of MS features.—A data processing pipeline established with the *enviMass*³³ package in the R Statistical Software was then used to extract MS features from the full-scan MS acquisitions.³⁴ Briefly, .RAW files were converted to the nonproprietary .mzXML file type³⁵ and fully-resolved chromatographic peaks were identified. The mass-to-charge ratio (m/z) of each of the picked peaks was then recalibrated based on the measured m/z of each of the ILISs in each of the samples. Picked chromatographic peaks had a minimum intensity of 104, a minimum signal-to-noise ratio of 5, and a minimum signal-to-base ratio of 2. Full peak picking settings are provided elsewhere.³⁴ Then, a replicate filter was applied to remove all picked peaks that were not measured ubiquitously across triplicate measurements of each sample. A blind filter was applied to remove all picked peaks that were present in the blank samples at a greater intensity than in the field samples. Isotopologues, adducts, and in-source generated fragments attributed to a single parent peak were grouped into components and lower-order isotopologues, adducts, and in-source fragments were removed. Then, the remaining picked peaks that had the same m/z and retention time across the field samples were grouped together as non-target MS features. Finally, the intensities of each peak contained in each MS feature were normalized based on the intensities of the ILISs in each of the samples.

2.4.5 Prioritization of MS features.—There were 130 chemicals tentatively identified in the target screening and 67 114 non-target MS features identified in at least one of the sample extracts. As a means to prioritize target chemicals and non-target MS features for further exploration and possible structural elucidation, eleven prioritization criteria were defined based on the proximity of the sample location to impaired gas wells or previously measured biological effects. These criteria include: (1) sample sites located within 1 mile of any impaired gas well; (2) sample sites located within 1 mile of at least six impaired gas wells; (3) samples exhibiting any Ah activity; (4) samples with Ah activity defined in categories 2–4; (5) samples with Ah activity defined in categories 3–4; (6) samples with Ah activity defined in category 4; (7) samples exhibiting AR antagonist activity; (8) samples exhibiting GR antagonist activity; (9) samples exhibiting ER agonist activity; (10) samples exhibiting ER antagonist activity; and (11) samples exhibiting PR antagonist activity. For each criterion, the matrix of MS features was partitioned into a group consisting of samples meeting the criterion and a group consisting of samples not meeting the criterion. The average peak area ratio of the target chemicals and the average normalized intensity of the non-target MS features in each group were calculated. These averages were compared between the groups of samples meeting the criterion and the group of samples not meeting the criterion. If the average peak area ratio or normalized intensity was at least 1.5 times higher in the group meeting the criterion, the chemical or feature was selected for further

exploration and possible structural elucidation. For example, there were 210 non-target MS features that were present in sites with Ah activity defined in category 4 and had an average normalized intensity greater than 1.5 times the normalized intensity of all other sample sites. A Student's *t*-test was used as a separate metric to evaluate significant differences ($p < 0.05$) in average peak ratios or normalized intensities between target chemicals and MS features in the groups meeting and not meeting each criterion.

2.4.6 Suspect screening.—The exact masses of the prioritized non-target MS features were then compared to the exact masses of 732 chemicals that have been reported as HFF additives in previous publications.^{11,19} The accurate masses of each of the prioritized non-target MS features were compared to the exact masses of each chemical on the suspect lists. The $[M]^+$, $[M + H]^+$, $[M + NH_4]^+$, and $[M + Na]^+$ adducts of each suspect chemical were included in the screening because these adducts were previously observed with our analytical method.^{21,32}

2.4.7 Communicating confidence.—Prioritized target chemicals and suspect hits were assigned confidence levels based on criteria established elsewhere.³⁶ Briefly, suspect hits with fully-resolved chromatographic peaks and MS spectra that match the theoretical isotope pattern of the suspect chemical are assigned a confidence of level 4 (unequivocal molecular formula). Suspect hits with sufficient evidence to support a structure based on MS spectra and MS2 spectra are assigned a confidence of level 3 (probable structure). Suspect hits with MS2 spectra that match MS2 spectra reported in the mzCloud database or that of fragments generated *in silico* for the suspect chemical using the MetFrag³⁶ software are assigned a confidence of level 2 (likely structure). When available, authentic standards of suspect hits were acquired, and if the retention time, MS spectra, and MS2 spectra of the suspect hit match with that of the authentic standard, the suspect hit is assigned a confidence of level 1 (confirmed structure).

3. Results and discussion

3.1 Sample collection and natural gas infrastructure

Water samples were collected at a variety of sites (Fig. 1A) throughout Susquehanna County, Pennsylvania from private drinking water wells (33), streams (6), ponds (9), springs (4) and a lake (1). The GPS coordinates were recorded and used to determine the Haversine distance (distance between two points on a sphere) from natural gas infrastructure. Fig. 1B displays the locations of sites of natural gas infrastructure (gas wells, compressor stations, dehydrators). Natural gas wells with casing and/or cement impairment as determined by Ingraffea *et al.*²⁹ are also highlighted in the figure. The average distance from a horizontal natural gas well to each sampling site was 0.55 ± 0.29 miles, with the distances ranging from 0.077 to 1.1 miles. Because of the density of natural gas infrastructure and the likelihood that gas wells vary as to the potential release of environmental contaminants, additional information was required to determine if extraction activity was associated with biological activity and chemical detection. Therefore, the list of natural gas wells with some indication of impairment compiled by Ingraffea *et al.*²⁹ was used for further analysis of the potential impact of these gas wells on surface water and groundwater. This list was extracted from

data obtained from the Pennsylvania Department of Environmental Protection (PADEP; January 1, 2000–December 31, 2012 (ref. 37)). Twenty-seven of the 53 water samples that were collected were within one mile of an impaired well.

3.2 Assays of biological activity

3.2.1 Ah receptor activity.—The Ah receptor is known to interact with a wide variety of natural and synthetic compounds.^{38,39} When expressed in yeast cells with appropriate response elements, Ah receptor activity becomes an attractive and low-cost sensor for potential environmental contamination. Ah receptor activity (binned into four levels of activity, see Methods) was detected in 14 groundwater samples and 8 surface water samples. Ah activity is shown in Fig. 2A along with the number of gas wells within one mile showing some degree of impairment.²⁹ Of the 27 samples taken within one mile of an impaired natural gas well, 18 showed Ah activity, while only 4 out of 26 samples that were taken outside of one mile showed Ah activity. This demonstrates a significant association between Ah receptor activity and proximity to impaired natural gas wells ($p = 0.00015$, Fisher Exact Probability Test). Of the four not within a mile of an impaired well, one sampling site, a drinking water well, was 0.13 mile from a horizontal well. The striking degree of overlap is shown spatially in Fig. 2B, with sampling sites showing Ah activity largely occurring near impaired gas wells. When samples were taken from water wells and ponds on the same property (eight instances), in all but one case, samples had either no Ah receptor activity in both samples (three instances) or Ah activity was detected in both samples (four instances).

3.2.2 Endocrine receptor activity.—Potential endocrine activity was assessed using four endocrine receptors (androgen [AR], estrogen [ER], progesterone [PR], and glucocorticoid [GR]) expressed in mammalian cells. In samples with significant activity, the magnitude of the effects seen for both agonism and antagonism are similar to those observed previously in areas of intensive drilling activity.¹⁶

3.2.3 Endocrine receptor antagonism.—Inhibitory activity (antagonism) was determined by the decrease in response by co-application of the agonist for each receptor with an extract of the water sample. Unlike Ah receptor activity, which was found largely in groundwater samples (64% in groundwater, 36% in surface water), endocrine receptor inhibition was significantly more likely ($p = 0.0103$, Fisher Exact Probability Test) to be found in surface water samples (30% in groundwater, 70% in surface water). Of the 27 samples taken within one mile of an impaired natural gas well, 10 showed an inhibition of at least one endocrine receptor, while 7 out of 25 samples taken outside of one mile showed inhibition (Fig. 3). The association with proximity to impaired natural gas wells was not statistically significant ($p = 0.185$, Fisher Exact Probability Test).

The presence of inhibitory endocrine activity in water samples was not associated with proximity to an impaired natural gas well, but of the seven samples that were not near an impaired well, five were observed during sample collection to be in the midst of agricultural activity. Although we have not attempted to quantify the many potential contributions of agricultural activity to these results, some chemicals from agricultural runoff are known to be endocrine disrupters,^{40,41} and it is plausible or perhaps even likely that endocrine

inhibition could arise from multiple environmental sources that could include natural gas activity, agriculture, and the presence of pharmaceuticals.

3.2.4 Endocrine receptor agonism.—In some cases, the response for the four endocrine receptors to a water sample was greater than the control level, suggesting that an agonist of receptor activity was present. This was observed in 22 samples, and in each case, the agonist activity was associated with ER (Fig. 4). One sample (drinking water well) showed agonism toward ER, AR, and GR receptors, and another (sampled from a stream), showed agonism with both the ER and PR receptors. In stark contrast to endocrine antagonism, of the 22 samples showing agonist activity, all but 2 were from groundwater samples. No significant association with unimpaired or impaired wells was observed when all samples were analysed ($p = 0.036$, Fisher Exact Probability Test).

3.2.5 Discussion of biological activities.—Ah receptor activity exhibited a strong correlation with proximity to impaired natural gas wells. Neither endocrine receptor agonism nor antagonism showed such a correlation. It is not clear whether this lack of correlation is due to the absence of endocrine disrupting substances contributed by HVHF activity or simply the fact that other activities, such agriculture, also contributed to the results.

3.3 LC-HRMS screening

The Ah receptor activity assays suggested that a biological activity could be associated with water samples near impaired natural gas wells. We used a complementary approach to screen for chemical composition of each sample extracted on the Oasis HLB cartridge. The analytical method allowed for a broad detection of organic chemicals in the range of 100–1000 Da that can be ionized during positive mode electrospray ionization and retained on a C18 analytical column. We acknowledge that this decision limits our ability to detect known chemicals that exhibit endocrine or AhR activity (*i.e.*, phenolic compounds and PAHs), though the majority of the chemicals on our target and suspect lists are more efficiently ionized in positive polarity mode. Full-scan high-resolution mass spectra were acquired for each sample, and target and suspect screening techniques were used to identify chemicals of interest. Chemicals of interest were prioritized for structural elucidation based on enhanced abundance in one of eleven prioritization groups (see Methods) related to proximity to impaired wells or observed Ah or endocrine disruption activity.

3.3.1 Target anthropogenic organic chemicals.—The target chemicals that were identified using a compound library containing 218 anthropogenic organic chemicals and confirmed after prioritization are provided in Table 1, along with the reason for prioritization, our confidence in the annotation, and the significance level. A total of 13 target chemicals were identified, nine of which were confirmed with an authentic standard (level 1) and four of which were assigned a confidence of level 3 (because the MS2 data acquired was insufficient to confirm their structure unequivocally). Seven of the 13 target chemicals had significantly greater peak area ratios in at least one prioritization group (marked with ^a in Table 1). All normalized peak area ratios for identified target chemicals in every sample are provided in the ESI.†

The identified compounds include four chemicals that have been disclosed as HFF additives or wastewater constituents (coumarin, tributyl phosphate, bis(2-ethylhexyl phthalate), and diethyl phthalate).³ Notably, tributyl phosphate was significant with respect to the '>6 impaired wells' prioritization group. However, it also has many commercial and industrial uses and has been measured in surface water and groundwater sources at low levels for decades.^{21,42,43} The other three chemicals were significant with respect to antagonistic endocrine effects, but not proximity to impaired wells. This could be evidence that some of the toxicity and endocrine disrupting behavior measured in these water samples is related to anthropogenic chemicals sourced from non-HVHF activities.

The remaining nine chemicals identified in this analysis are often associated with agricultural activities, wastewater influence, or industrial operations. For example, carbamazepine was identified and found to be significant with respect to the Ah activity defined in category 4. Carbamazepine is an antiepileptic drug most commonly associated with wastewater influence,⁴⁴ so its presence is not likely associated with HVHF activities in the sample area. Similarly, caffeine (wastewater influenced), paraxanthine/theophylline (caffeine metabolites), and benzophenone (personal care product, recreation and wastewater influenced) were identified and prioritized based on significant antagonistic endocrine activity. These are not chemicals expected to be associated with HVHF activities and all three are frequently detected in water resources. However, the presence of these compounds and other anthropogenic compounds could enhance the toxicity of other compounds, including those sourced from HVHF activities, due to synergistic effects.¹⁵

3.3.2 Suspect HFF additives.—The suspect hits identified in samples using a suspect database of 732 HFF additives and confirmed after prioritization are provided in Table 2. We identified a total of 13 prioritized suspect hits, one of which was confirmed with an authentic standard (level 1) and the remaining were assigned a confidence of level 4 (unequivocal molecular formula) because MS2 spectra could not be interpreted to definitively support the structure and no mzCloud or MetFrag data could be used to confirm the structures. Three of the 13 suspect hits also had significantly greater normalized intensities in at least one prioritization group (marked with an asterisk in Table 2). All normalized intensities for identified suspect hits are provided in the ESI.[†] Two of those came from suspect hits identified in the 'impaired wells' prioritization groups and the other was identified in the ER antagonist prioritization group. These include diethylene glycol (impaired wells), 2-(dibutylamino)ethanol (>6 impaired wells), and *tert*-butyl peroxybenzoate (ER antagonist). The former two suspect hits might be suggested as indicators of nearby HVHF activities, though both unequivocal molecular formulae are relatively common and the possibility of significant but non-causal associations must be considered. Targeted research aimed at addressing that possibility would be required to make any stronger conclusions from these data. The level 1 suspect hit was triethyl phosphate (TEP), which was prioritized based on its association with samples displaying AR antagonism. Details of the confirmation are shown in the ESI.[†] TEP, a disclosed HFF additive and organophosphate plasticizer, has previously

[†]Electronic supplementary information (ESI) available. See DOI: [10.1039/c9em00112c](https://doi.org/10.1039/c9em00112c)

been detected in European surface waters at concentrations up to 180 ng l⁻¹.^{45,46} Therefore, its presence in these samples cannot be unequivocally associated with HVHF activities.

3.3.3 Discussion of chemical occurrences.—Our chemical analyses revealed ten chemicals whose structures were confirmed with authentic standards, three probable chemical structures, and twelve chemicals with unequivocal molecular formulas that match those of known HFF additives. Each of the 26 identified chemicals (13 from Table 1 and 13 from Table 2) were identified exclusively or at significantly greater abundance in samples in close proximity to impaired wells or that exhibited a specific biological effect.

Unlike most analyses of water samples, this approach can detect a wide range of compounds that do not necessarily have to be defined in advance of testing. Although a large number of compounds can be detected, there are some limitations in the initial extraction step and intrinsic to this LC-MS analysis in that some degree of hydrophobicity and ability to carry a positive charge are required. Compounds that have been associated with oil and gas activity have been found in all of the samples tested here, but many of these compounds have been found in water samples collected outside of Susquehanna County and far from oil and gas activity.^{31,34,47,48} This is likely due to the fact that many compounds used in HFF have a variety of other uses. In order to try to focus the analysis, we developed eleven prioritization groups based on the results of the biological assays and proximity to impaired wells. Of the potential HFF additives, tributyl phosphate (level 1), diethylene glycol (level 4) and 2-(dibutylamino)ethanol (level 4) were significantly associated with impaired wells. Another compound associated with HFF additives is *tert*-butyl peroxybenzoate (level 4), which was associated with ER antagonist activity but not proximity to impaired wells. As noted above, we cannot rule out the existence of sources of these chemicals other than as HFF additives. Overall, seventeen MS features with unequivocal molecular formulas (level 4) or higher level of confidence matching identified HFF additives were found in at least one of the prioritization groups.

Other compounds identified in these screenings are notable. Two phthalates [bis(2-ethylhexyl) phthalate (level 1) and diethyl phthalate (level 1)] are significantly associated with samples displaying endocrine receptor inhibition and Ah activity. In particular, bis(2-ethylhexyl) phthalate has been identified previously at levels of concern in water samples from Susquehanna County.^{8,10} Although it has been used extensively as a component of plasticizers, bis(2-ethylhexyl) phthalate has been reported in flowback waters from several shale formations.^{49–51} In addition, a number of other compounds were prioritized: several pharmaceuticals or pharmaceutical precursors (coumarin, metoprolol, adrenalone, carbamazepine), a flame retardant (tris(2-chloroethyl) phosphate), and pesticides or pesticide degradation products (atrazine-2-hydroxy and DEET).

Although we can significantly associate Ah receptor activity with proximity to impaired natural gas wells, we have not definitively identified chemicals responsible for those activities. Nevertheless, benzophenone, tri(2-chloroethyl) phosphate, tributyl phosphate, and sorbitan monooleate have been used in drilling for hydrocarbons and are associated with Ah activity in our analyses. In addition, several potential HFF additives were identified preferentially in proximity to impaired gas wells (tributyl phosphate, diethylene glycol, ethyl

acetoacetate, 2-(dibutylamino)ethanol, and dimethyl adipate). Likewise, a number of HFF additives were found to be associated with the endocrine disruption prioritization groups. In addition, a number of other compounds not associated with HFF additives were found in the prioritization groups that include aspects of Ah activity, proximity to impaired wells, and endocrine disruption.

4 Conclusions

Our approach here was first to determine if an association exists between biological activity and a subset of impaired wells. This has the potential of missing gas wells that may have contributed to water contamination but without registering infractions on the PADEP database; however, a strong association with Ah activity was observed. The LC-HRMS results demonstrate that a large variety of compounds can be detected in surface water and groundwater samples including industrial chemicals (such as disclosed HFF additives and chemicals associated with HVHF wastewater), agricultural chemicals, and pharmaceuticals. This range of substances complicates any attempt to pinpoint the source of contamination. However, by grouping water samples by biological activity and proximity to impaired wells (prioritization groups), we detected 17 potential HFF additives or wastewater constituents that are associated with Ah activity, ER activity, and proximity to impaired wells. Although most of these compounds have other uses in addition to natural gas extraction, the association with biological activity and impaired wells suggests that anthropogenic activities, including hydraulic fracturing operations, have resulted in water contamination.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was supported by grants from the Claneil Foundation to R. E. O. and the Atkinson Center for a Sustainable Future to R. E. O. and D. E. H., NIH R21ES026395 (S. C. N.) and the University of Missouri (S. C. N. and R. F. K.). The authors would like to thank Drs Adam Law, Gregory A. Weiland, and John Dennis for helpful discussions.

References

1. Osborn SG, Vengosh A, Warner NR and Jackson RB, Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, 108, 8172–8176. [PubMed: 21555547]
2. Elliott EG, Ettinger AS, Leaderer BP, Bracken MB and Deziel NC, A systematic evaluation of chemicals in hydraulic-fracturing fluids and wastewater for reproductive and developmental toxicity, *J. Exposure Sci. Environ. Epidemiol.*, 2017, 27, 90–99.
3. EPA, Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States (Final Report - EPA/600/R-16/236F), <https://cfpub.epa.gov/ncea/hfstudy/recordisplay.cfm?deid=332990>.
4. Rogers JD, Burke TL, Osborn SG and Ryan JN, A Framework for Identifying Organic Compounds of Concern in Hydraulic Fracturing Fluids Based on Their Mobility and Persistence in Groundwater, *Environ. Sci. Technol. Lett.*, 2015, 2, 158–164.

5. Stringfellow WT, Domen JK, Camarillo MK, Sandelin WL and Borglin S, Physical, chemical, and biological characteristics of compounds used in hydraulic fracturing, *J. Hazard. Mater*, 2014, 275, 37–54. [PubMed: 24853136]
6. Yost EE, Stanek J, DeWoskin RS and Burgoon LD, Overview of Chronic Oral Toxicity Values for Chemicals Present in Hydraulic Fracturing Fluids, Flowback, and Produced Waters, *Environ. Sci. Technol*, 2016, 50, 4788–4797. [PubMed: 27050380]
7. Clara M, Windhofer G, Hartl W, Braun K, Simon M, Gans O, Scheffnecht C and Chovanec A, Occurrence of phthalates in surface runoff, untreated and treated wastewater and fate during wastewater treatment, *Chemosphere*, 2010, 78, 1078–1084. [PubMed: 20096917]
8. Drollette BD, Hoelzer K, Warner NR, Darrah TH, Karatum O, O'Connor MP, Nelson RK, Fernandez LA, Reddy CM, Vengosh A, Jackson RB, Elsner M and Plata DL, Elevated levels of diesel range organic compounds in groundwater near Marcellus gas operations are derived from surface activities, *Proc. Natl. Acad. Sci. U.S. A*, 2015, 112, 13184–13189. [PubMed: 26460018]
9. Fromme H, Kuchler T, Otto T, Pilz K, Muller J and Wenzel A, Occurrence of phthalates and bisphenol A and F in the environment, *Water Res*, 2002, 36, 1429–1438. [PubMed: 11996333]
10. ATSDR, Health consultation: Dimock groundwater site, https://www.atsdr.cdc.gov/HAC/pha/DimockGroundwaterSite/Dimock_Groundwater_Site_HC_05-24-2016_508.pdf.
11. Elsner M and Hoelzer K, Quantitative Survey and Structural Classification of Hydraulic Fracturing Chemicals Reported in Unconventional Gas Production, *Environ. Sci. Technol*, 2016, 50, 3290–3314. [PubMed: 26902161]
12. Brack W, Effect-directed analysis: a promising tool for the identification of organic toxicants in complex mixtures?, *Anal. Bioanal. Chem*, 2003, 377, 397–407. [PubMed: 12904950]
13. Browne P, Noyes PD, Casey WM and Dix DJ, Application of Adverse Outcome Pathways to U.S. EPA's Endocrine Disruptor Screening Program, *Environ. Health Perspect*, 2017, 125, 096001. [PubMed: 28934726]
14. Kassotis CD, Iwanowicz LR, Akob DM, Cozzarelli IM, Mumford AC, Orem WH and Nagel SC, Endocrine disrupting activities of surface water associated with a West Virginia oil and gas industry wastewater disposal site, *Sci. Total Environ*, 2016, 557–558, 901–910.
15. Kassotis CD, Klemp KC, Vu DC, Lin CH, Meng CX, Besch-Williford CL, Pinatti L, Zoeller RT, Drobnis EZ, Balise VD, Isiguzo CJ, Williams MA, Tillitt DE and Nagel SC, Endocrine-Disrupting Activity of Hydraulic Fracturing Chemicals and Adverse Health Outcomes After Prenatal Exposure in Male Mice, *Endocrinology*, 2015, 156, 4458–4473. [PubMed: 26465197]
16. Kassotis CD, Tillitt DE, Davis JW, Hormann AM and Nagel SC, Estrogen and androgen receptor activities of hydraulic fracturing chemicals and surface and ground water in a drilling-dense region, *Endocrinology*, 2014, 155, 897–907. [PubMed: 24424034]
17. Miller CA 3rd, Expression of the human aryl hydrocarbon receptor complex in yeast. Activation of transcription by indole compounds, *J. Biol. Chem*, 1997, 272, 32824–32829. [PubMed: 9407059]
18. Miller CA 3rd, A human aryl hydrocarbon receptor signaling pathway constructed in yeast displays additive responses to ligand mixtures, *Toxicol. Appl. Pharmacol*, 1999, 160, 297–303. [PubMed: 10544064]
19. Getzinger GJ, O'Connor MP, Hoelzer K, Drollette BD, Karatum O, Deshusses MA, Ferguson PL, Elsner M and Plata DL, Natural Gas Residual Fluids: Sources, Endpoints, and Organic Chemical Composition after Centralized Waste Treatment in Pennsylvania, *Environ. Sci. Technol*, 2015, 49, 8347–8355. [PubMed: 26147419]
20. Moschet C, Piazzoli A, Singer H and Hollender J, Alleviating the reference standard dilemma using a systematic exact mass suspect screening approach with liquid chromatography-high resolution mass spectrometry, *Anal. Chem*, 2013, 85, 10312–10320. [PubMed: 24161211]
21. Pochodylo AL and Helbling DE, Emerging investigators series: prioritization of suspect hits in a sensitive suspect screening workflow for comprehensive micropollutant characterization in environmental samples, *Environ. Sci.: Water Res. Technol*, 2017, 3, 54–65.
22. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP and Hollender J, Identifying small molecules via high resolution mass spectrometry: communicating confidence, *Environ. Sci. Technol*, 2014, 48, 2097–2098. [PubMed: 24476540]

23. Moschet C, Wittmer I, Simovic J, Junghans M, Piazzoli A, Singer H, Stamm C, Leu C and Hollender J, How a complete pesticide screening changes the assessment of surface water quality, *Environ. Sci. Technol*, 2014, 48, 5423–5432. [PubMed: 24821647]
24. Chiaia-Hernandez AC, Krauss M and Hollender J, Screening of lake sediments for emerging contaminants by liquid chromatography atmospheric pressure photoionization and electrospray ionization coupled to high resolution mass spectrometry, *Environ. Sci. Technol*, 2013, 47, 976–986. [PubMed: 23215447]
25. Gago-Ferrero P, Schymanski EL, Bletsou AA, Aalizadeh R, Hollender J and Thomaidis NS, Extended Suspect and Non-Target Strategies to Characterize Emerging Polar Organic Contaminants in Raw Wastewater with LC-HRMS/MS, *Environ. Sci. Technol*, 2015, 49, 12333–12341. [PubMed: 26418421]
26. Luo X, Zhang L, Niu Z, Ye X, Tang Z and Xia S, Liquid chromatography coupled to quadrupole-Orbitrap high resolution mass spectrometry based method for target analysis and suspect screening of non-ionic surfactants in textiles, *J. Chromatogr. A*, 2017, 1530, 80–89. [PubMed: 29153579]
27. Sollic M, Roy-Lachapelle A and Sauve S, Development of a suspect and non-target screening approach to detect veterinary antibiotic residues in a complex biological matrix using liquid chromatography/high-resolution mass spectrometry, *Rapid Commun. Mass Spectrom*, 2015, 29, 2361–2373. [PubMed: 26563707]
28. Oetjen K, Giddings CGS, McLaughlin M, Nell M, Blotevogel J, Helbling DE, Mueller D and Higgins CP, Emerging analytical methods for the characterization and quantification of organic contaminants in flowback and produced water, *Trends Environ. Anal. Chem*, 2017, 15, 12–23.
29. Ingraffea AR, Wells MT, Santoro RL and Shonkoff SB, Assessment and risk analysis of casing and cement impairment in oil and gas wells in Pennsylvania, 2000–2012, *Proc. Natl. Acad. Sci. U. S. A*, 2014, 111, 10955–10960. [PubMed: 24982144]
30. Li C, Klemes MJ, Dichtel WR and Helbling DE, Tetrafluoroterephthalonitrile-crosslinked beta-cyclodextrin polymers for efficient extraction and recovery of organic micropollutants from water, *J. Chromatogr. A*, 2018, 1541, 52–56. [PubMed: 29433822]
31. Carpenter CMG and Helbling DE, Widespread Micropollutant Monitoring in the Hudson River Estuary Reveals Spatiotemporal Micropollutant Clusters and Their Sources, *Environ. Sci. Technol*, 2018, 52, 6187–6196. [PubMed: 29742349]
32. Nell M and Helbling DE, Exploring matrix effects and quantifying organic additives in hydraulic fracturing associated fluids using liquid chromatography electrospray ionization mass spectrometry, *Environ. Sci.: Processes Impacts*, 2019, 21, 195–205.
33. Loos M, enviMass version 3.5 LC-HRMS trend detection workflow - R package. Zenodo, 2018.
34. Carpenter C, Wong LYJ, Johnson CA and Helbling DE, Fall Creek Monitoring Station: highly resolved temporal sampling to prioritize the identification of non-target micropollutants in a small stream, *Environ. Sci. Technol*, 2018, DOI: .
35. Kessner D, Chambers M, Burke R, Agus D and Mallick P, ProteoWizard: open source software for rapid proteomics tools development, *Bioinformatics*, 2008, 24, 2534–2536. [PubMed: 18606607]
36. Schymanski EL, Singer HP, Slobodnik J, Ipolyi IM, Oswald P, Krauss M, Schulze T, Haglund P, Letzel T, Grosse S, Thomaidis NS, Bletsou A, Zwiener C, Ibanez M, Portoles T, de Boer R, Reid MJ, Onghena M, Kunkel U, Schulz W, Guillon A, Noyon N, Leroy G, Bados P, Bogialli S, Stipanicev D, Rostkowski P and Hollender J, Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis, *Anal. Bioanal. Chem*, 2015, 407, 6237–6255. [PubMed: 25976391]
37. PADEP, www.depweb.state.pa.us/portal/server.pt/community/oil_and_gas_reports/20297.
38. Denison MS, Pandini A, Nagy SR, Baldwin EP and Bonati L, Ligand binding and activation of the Ah receptor, *Chem.-Biol. Interact*, 2002, 141, 3–24. [PubMed: 12213382]
39. Denison MS, Soshilov AA, He G, DeGroot DE and Zhao B, Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor, *Toxicol. Sci*, 2011, 124, 1–22. [PubMed: 21908767]
40. Benotti MJ, Trenholm RA, Vanderford BJ, Holady JC, Stanford BD and Snyder SA, Pharmaceuticals and endocrine disrupting compounds in U.S. drinking water, *Environ. Sci. Technol*, 2009, 43, 597–603. [PubMed: 19244989]

41. Combarous Y, Endocrine Disruptor Compounds (EDCs) and agriculture: The case of pesticides, *C. R. Biol*, 2017, 340, 406–409. [PubMed: 28826788]
42. Fries E and Puttmann W, Occurrence of organophosphate esters in surface water and ground water in Germany, *J. Environ. Monit*, 2001, 3, 621–626. [PubMed: 11785636]
43. Stackelberg PE, Furlong ET, Meyer MT, Zaugg SD, Henderson AK and Reissman DB, Persistence of pharmaceutical compounds and other organic wastewater contaminants in a conventional drinking-water-treatment plant, *Sci. Total Environ*, 2004, 329, 99–113. [PubMed: 15262161]
44. Wode F, van Baar P, Dunnbier U, Hecht F, Taute T, Jekel M and Reemtsma T, Search for over 2000 current and legacy micropollutants on a wastewater infiltration site with a UPLC-high resolution MS target screening method, *Water Res*, 2015, 69, 274–283. [PubMed: 25497426]
45. Bollmann UE, Moller A, Xie Z, Ebinghaus R and Einax JW, Occurrence and fate of organophosphorus flame retardants and plasticizers in coastal and marine surface waters, *Water Res*, 2012, 46, 531–538. [PubMed: 22142598]
46. Rodil R, Quintana JB, Concha-Grana E, Lopez-Mahia P, Muniategui-Lorenzo S and Prada-Rodriguez D, Emerging pollutants in sewage, surface and drinking water in Galicia (NW Spain), *Chemosphere*, 2012, 86, 1040–1049. [PubMed: 22189380]
47. Ferrer I and Furlong ET, Identification of alkyl dimethylbenzylammonium surfactants in water samples by solid-phase extraction followed by ion trap LC/MS and LC/MS/MS, *Environ. Sci. Technol*, 2001, 35, 2583–2588. [PubMed: 11432568]
48. Lara-Martin PA, Gonzalez-Mazo E and Brownawell BJ, Multi-residue method for the analysis of synthetic surfactants and their degradation metabolites in aquatic systems by liquid chromatography-time-of-flight-mass spectrometry, *J. Chromatogr*, 2011, 1218, 4799–4807.
49. Ferrer I and Thurman EM, Chemical constituents and analytical approaches for hydraulic fracturing waters, *Trends Environ. Anal. Chem*, 2015, 5, 18–25.
50. Lester Y, Ferrer I, Thurman EM, Sitterley KA, Korak JA, Aiken G and Linden KG, Characterization of hydraulic fracturing flowback water in Colorado: Implications for water treatment, *Sci. Total Environ*, 2015, 512, 637–644. [PubMed: 25658325]
51. Maguire-Boyle SJ and Barron AR, Organic compounds in produced waters from shale gas wells, *Environ. Sci.: Processes Impacts*, 2014, 16, 2237–2248.

Environmental Significance

Assessment of surface water and groundwater contamination by anthropogenic activities, such as extraction of oil and gas by hydraulic fracturing or agricultural activity, is challenging because of the large variety of chemical constituents, both known and unknown. We have applied both biological assays and liquid chromatography coupled to high-resolution mass spectrometry to analyze water quality and composition in an agricultural region of Pennsylvania (USA) that has experienced high levels of fossil fuel extraction activity. In particular, aryl hydrocarbon receptor activation was associated with proximity to gas wells showing some degree of impairment, and hydraulic fracturing fluid additives were detected in water samples in proximity to impaired gas wells and those with some biological activity.

Fig. 1.

(A) Location of water sampling sites with longitude and latitude scales. (B) Map (top) of natural gas infrastructure in Susquehanna County, PA. The smaller maps below indicate the position of the county within the northeast United States. Maps in this figure and subsequent figures were generated with Google Maps.

Fig. 2.

(A) Each set of bars (blue and amber) represent a single sample (sorted by the number of impaired wells within one mile). The blue bars show the number of impaired wells within one mile for the site at which the sample was taken and the amber bars show the activity index for Ah activity (activity was binned into four categories, 1 to 4, with 4 being the highest activity; see Methods). The values on the y-axis represent both the number of impaired wells and the Ah activity index. All but four samples that tested positive for Ah activity were within one mile of an impaired well. (B) The spatial distribution of the sampling sites is shown on this map, with each symbol representing one sample. Pushpin icons indicate Ah activity and paddle icons indicate no Ah activity. Red icons are within one mile of an impaired well and light blue icons are not within one mile. Overlap between Ah activity and samples taken within one mile of an impaired well are indicated by the red pushpins.

Fig. 3.

(A) Blue bars are as described for Fig. 2. Significant endocrine receptor antagonism is indicated by a bar of unit size one in the color shown in the y-axis label. The y-axis scale refers to both the number of impaired wells and the antagonist activity. (B) The spatial distribution of the sampling sites is shown on the map, using the same conventions as described in the legend to Fig. 1.

Fig. 4.

(A) Blue bars are as described for Fig. 2. Significant endocrine receptor agonism is indicated by a bar of unit size one in the color shown in the y-axis label. The y-axis scale refers to both the number of impaired wells and the agonist activity. (B) The spatial distribution of the sampling sites is shown on the map, using the same conventions as described in the legend to Fig. 1.

Table 1

Chemicals identified from the target screening for 218 anthropogenic organic chemicals

<i>m/z</i> value	Retention time (min)	Match(es) in micropollutant database	CAS	Reason for prioritization	Confidence	MS-ready molecular formula	Adduct detected
147.0441	4.9	Coumarin ^c	91-64-5	>six impaired wells, AR antagonist, ^a ER antagonist, PR antagonist ^a	Level 1	C ₉ H ₆ O ₂	[M + H]
267.1717	12.9	Tributyl phosphate ^c	126-73-8	>six impaired wells, Ah activity defined in categories 3–4, Ah activity defined in category 4, ER agonist	Level 1	C ₁₂ H ₂₇ O ₄ P	[M + H]
198.1352	3.7	Atrazine-2-hydroxy	2163-68-0	All impaired well sites (up to 10 NA)	Level 1	C ₈ H ₁₅ N ₅ O	[M + H]
391.2848	18.2	Bis(2-ethylhexyl) phthalate ^c	117-81-7	Ah activity defined in category 4, AR antagonist, ^a GR antagonist, ^a ER antagonist, ^a PR antagonist ^a	Level 1	C ₂₄ H ₃₈ O ₄	[M + Na]
183.0804	9.0	Benzophenone	119-61-9	Ah activity defined in category 4, ER antagonist, PR antagonist ^a	Level 1	C ₁₃ H ₁₀ O	[M + H]
192.1382	7.4	DEET	134-62-3	All impaired well sites (up to 10 NA), Ah activity, GR antagonist	Level 1	C ₁₂ H ₁₇ NO	[M + H]
195.0877	3.4	Caffeine	58-08-2	AR antagonist, ^a GR antagonist, ER antagonist, ^a PR antagonis ^a	Level 1	C ₈ H ₁₀ N ₄ O ₂	[M + H]
223.0968	7.7	Diethyl phthalate ^c	84-66-2	AR antagonist, ^a GR antagonist, ER antagonist, ^a PR antagonis ^a	Level 1	C ₁₂ H ₁₄ O ₄	[M + H]
268.191	4.2	Metoprolol	51384-51-1	ER antagonist	Level 1	C ₁₅ H ₂₅ NO ₃	[M + H]
181.0722	2.7	Paraxanthine/theophylline	611-59-6/58-55-9	ER antagonist, ^a PR antagonist ^a	Level 3 ^b	C ₇ H ₈ N ₄ O ₂	[M + H]
182.0811	1.4	Adrenalone	99-45-6	Ah activity defined in categories 3–4, Ah activity defined in category 4	Level 3 ^b	C ₉ H ₁₁ NO ₃	[M + H]
237.1008	6.6	Carbamazepine	298-46-4	Ah activity defined in category 4 ^a	Level 3 ^b	C ₁₅ H ₁₂ N ₂ O	[M + H]
284.9611	6.2	Tris(2-chloroethyl)	115-96-8	Ah activity defined in category 4	Level 3 ^b	C ₆ H ₁₂ Cl ₃ O ₃ P	[M + H]

^aDenotes significance ($p < 0.05$) for Student's *t*-test described in prioritization of MS features.^bCompounds assigned a confidence of level 3 because MS2 fragments were unavailable or indistinguishable from another structure.^cHFF additives.

Table 2

Chemicals identified from the suspect list containing 732 HFF additives

<i>m/z</i> value	Retention time (min)	Match(es) in HFF database	CAS	Reason for prioritization	Confidence	MS-ready molecular formula	Adduct detected
183.0781	5.33	Triethyl phosphate	78-40-0	AR antagonist	Level 1	C ₆ H ₁₅ O ₄ P	[M + H]
129.0522	1.87	Diethylene glycol	111-46-6	Impaired wells ^a	Level 4	C ₄ H ₁₀ O ₃	[M + Na]
131.0703	2.18	Ethyl acetate	141-97-9	>6 impaired wells	Level 4	C ₆ H ₁₀ O ₃	[M + H]
147.0917	3.04	Quinoline and isoquinoline	91-22-5, 119-65-4	ER antagonist	Level 4	C ₉ H ₇ N	[M + NH ₄]
161.1073	4.16	Quinaldine	91-63-4	ER antagonist	Level 4	C ₁₀ H ₉ N	[M + NH ₄]
174.1853	6.02	2-(Diethylamino)ethanol	102-81-8	>6 impaired wells, ^a ER antagonist, PR antagonist	Level 4	C ₁₀ H ₂₃ NO	[M + H]
181.0719	2.72	Biopolymer	50-99-7	ER antagonist	Level 4	C ₆ H ₁₂ O ₆	[M + H]
181.0719	2.36	Biopolymer	50-99-7	ER antagonist	Level 4	C ₆ H ₁₂ O ₆	[M + H]
183.0628	4.16	Dimethyl glutarate or diethylene glycol monoacrylate	1119-40-0, 13533-05-6	ER agonist	Level 4	C ₇ H ₁₂ O ₄	[M + Na]
195.1016	4.33	<i>tert</i> -Butyl peroxybenzoate	614-45-9	ER antagonist ^a	Level 4	C ₁₁ H ₁₄ O ₃	[M + H]
197.0785	3.12	Dimethyl adipate	627-93-0	>6 impaired wells	Level 4	C ₈ H ₁₄ O ₄	[M + Na]
295.2266	9.93	Triton (octyl/nonylphenol ethoxylates)	9036-19-5, 68987-90-6, 9002-93-1	PR antagonist	Level 4	C ₁₈ H ₃₀ O ₃	[M + H]
446.3474	12.45	Sorbitan monooleate	1338-43-8	Ah activity defined in category 4	Level 4	C ₂₄ H ₄₄ O ₆	[M + NH ₄]

^aDenotes significance ($p < 0.05$) for Student's *t*-test described in prioritization of MS features.