

Elevated levels of diesel range organic compounds in groundwater near Marcellus gas operations are derived from surface activities

Brian D. Drollette^a, Kathrin Hoelzer^b, Nathaniel R. Warner^c, Thomas H. Darrah^d, Osman Karatum^e, Megan P. O'Connor^e, Robert K. Nelson^f, Loretta A. Fernandez^g, Christopher M. Reddy^f, Avner Vengosh^h, Robert B. Jacksonⁱ, Martin Elsner^b, and Desiree L. Plata^{a,1}

^aDepartment of Chemical and Environmental Engineering, Yale University, New Haven, CT 06511; ^bInstitute of Groundwater Ecology, Helmholtz Zentrum München, 85764 Oberschleissheim, Germany; ^cDepartment of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA 16802; ^dSchool of Earth Sciences, The Ohio State University, Columbus, OH 43210; ^eDepartment of Civil and Environmental Engineering, Duke University, Durham, NC 27708; ^fDepartment of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; ^gDepartments of Civil and Environmental Engineering and Marine and Environmental Sciences, Northeastern University, Boston, MA 02115; ^hDivision of Earth and Ocean Sciences, Duke University, Durham, NC 27708; and ⁱSchool of Earth, Energy, and Environmental Sciences, Woods Institute for the Environment and Precourt Institute for Energy, Stanford University, Stanford, CA 94305

Edited by Andrea Rinaldo, Ecole Polytechnique Federale Lausanne, Lausanne, Switzerland, and approved September 16, 2015 (received for review June 11, 2015)

Hundreds of organic chemicals are used during natural gas extraction via high-volume hydraulic fracturing (HVHF). However, it is unclear whether these chemicals, injected into deep shale horizons, reach shallow groundwater aquifers and affect local water quality, either from those deep HVHF injection sites or from the surface or shallow subsurface. Here, we report detectable levels of organic compounds in shallow groundwater samples from private residential wells overlying the Marcellus Shale in northeastern Pennsylvania. Analyses of purgeable and extractable organic compounds from 64 groundwater samples revealed trace levels of volatile organic compounds, well below the Environmental Protection Agency's maximum contaminant levels, and low levels of both gasoline range (0–8 ppb) and diesel range organic compounds (DRO; 0–157 ppb). A compound-specific analysis revealed the presence of bis(2-ethylhexyl) phthalate, which is a disclosed HVHF additive, that was notably absent in a representative geogenic water sample and field blanks. Pairing these analyses with (i) inorganic chemical fingerprinting of deep saline groundwater, (ii) characteristic noble gas isotopes, and (iii) spatial relationships between active shale gas extraction wells and wells with disclosed environmental health and safety violations, we differentiate between a chemical signature associated with naturally occurring saline groundwater and one associated with alternative anthropogenic routes from the surface (e.g., accidental spills or leaks). The data support a transport mechanism of DRO to groundwater via accidental release of fracturing fluid chemicals derived from the surface rather than subsurface flow of these fluids from the underlying shale formation.

natural gas extraction | high-volume hydraulic fracturing | groundwater | hydrophobic organic contaminants | transport mechanisms

Technological advances in high-volume hydraulic fracturing (HVHF) have led to the expansion of unconventional fossil fuel extraction in the United States over the past decade (1–3). Despite the clear economic and national security benefits associated with domestic fuel production, the collocation of industrial practices with residential areas raises concerns for public and environmental health (4–6). In particular, it is unclear whether the organic chemicals that are used in relatively small proportions (but potentially large volumes) and injected into deep shale formations can contaminate shallow drinking water aquifers. Several questions emerge. If organic chemicals are detected in groundwater, did they arrive via surface discharges, shallow subsurface pathways (e.g., leaking gas wells), or deep transport routes? Furthermore, are organic compounds present in groundwater derived from naturally occurring, geogenic sources or associated with industrial activities, such as HVHF? Finally, what are the chemical fingerprints that enable one to make this distinction?

Although few studies have examined the occurrence and origin of organic contaminants in groundwater in HVHF regions (7), the presence of light hydrocarbon gases (i.e., methane and ethane) and inorganic constituents has been investigated frequently. Osborn et al. (8) and Jackson et al. (9) demonstrated elevated methane levels within 1 km of unconventional gas wells over the Marcellus Shale. Further, Darrah et al. (10) showed that stray gas contamination in a subset of groundwater wells likely resulted from poor well integrity (i.e., casing and cementing issues). In contrast, inorganic chemical constituents (e.g., Cl^- and Br^-) in groundwater over the Marcellus Shale seem to reflect geogenic sources and provide evidence of hydraulic connectivity between shallow groundwaters and deeper formation brines on geological timescales in some areas in northeastern Pennsylvania (11–13). This deep-origin, saline groundwater has a chemical and isotopic fingerprint similar to the Marcellus brines (11) but distinct from Marcellus flowback water (14). However, it is unknown whether this deep saline water carries a unique organic chemical fingerprint of either geogenic or anthropogenic origin.

Significance

Organic compounds found in drinking water aquifers above the Marcellus Shale and other shale plays could reflect natural geologic transport processes or contamination from anthropogenic activities, including enhanced natural gas production. Using analyses of organic compounds coupled with inorganic geochemical fingerprinting, estimates of groundwater residence time, and geospatial analyses of shale gas wells and disclosed safety violations, we determined that the dominant source of organic compounds to shallow aquifers was consistent with surface spills of disclosed chemical additives. There was no evidence of association with deeper brines or long-range migration of these compounds to the shallow aquifers. Encouragingly, drinking water sources affected by disclosed surface spills could be targeted for treatment and monitoring to protect public health.

Author contributions: B.D.D., K.H., N.R.W., T.H.D., A.V., R.B.J., M.E., and D.L.P. designed research; B.D.D., K.H., N.R.W., T.H.D., O.K., M.P.O., R.K.N., L.A.F., and D.L.P. performed research; C.M.R., A.V., R.B.J., and D.L.P. contributed new reagents/analytic tools; B.D.D., K.H., N.R.W., T.H.D., and D.L.P. analyzed data; and B.D.D., T.H.D., L.A.F., and D.L.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. Email: desiree.plata@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1511474112/-DCSupplemental.

The same mechanistic approaches taken to source apportion methane and inorganic compounds have not yet been applied to organic compounds in groundwater. Although Gross et al. (15) reported surface spills of hydraulic fracturing fluids and wastes that could affect groundwater with organic chemicals in Colorado, and the Pennsylvania Department of Environmental Protection (PA DEP) has assessed multiple instances of local groundwater pollution by gas drilling operations (16), these studies relied on voluntary industry accident disclosure and did not probe for possible alternative exposure paths (i.e., through a broad geospatial sample set). In addition, they did not provide a detailed characterization of the water that could identify the dominant transport processes associated with the contamination (i.e., organic and inorganic markers, along with hydrocarbon composition, noble gas isotopes, and spatial distribution analysis). Recently, Llewellyn et al. (7) investigated a localized incident of stray gas groundwater pollution in Pennsylvania that exposed the potential for groundwater contamination from natural gas extraction practices, where the authors attributed the contamination to flow from HVHF wells through shallow subsurface pathways. Although critical and detailed, the targeted nature of this case study (i.e., sampling wells with documented contamination or close to contamination sites) precludes the identification of geologic transport mechanisms that may be occurring in the region. Thus, the mechanisms of organic chemical transport in groundwater associated with HVHF regions remain unclear.

To address this research gap, we sampled 64 private residential groundwater wells, ranging from 9–213 m deep, over a 3-y period (2012–2014) in northeastern Pennsylvania ($n = 62$) and in southern New York ($n = 2$) for analyses of GC-amenable organic compounds (Fig. 1). Fifty-nine samples were analyzed for volatile organic compounds (VOCs) and gasoline range organic compounds (GRO; defined as the hydrocarbons eluting between 2-methylpentane and 1,2,4-trimethylbenzene; approximately between nC_6 and nC_{10}), and 41 were also analyzed for diesel range organic compounds (DRO; defined as the hydrocarbons eluting between nC_{10} and nC_{28}) (17). Analytical details are provided in *SI Appendix*. [Note that compounds included in the Environmental Protection Agency (EPA)-designated definitions of GRO and DRO are not necessarily gasoline or diesel derived]. A subset of these samples was analyzed using comprehensive two-dimensional gas chromatography (GC $\times$ GC) to evaluate whether compound-specific organic chemical fingerprints were associated with either HVHF activities or natural geologic processes. Complementary analyses of the inorganic chemical, methane stable isotope, and

helium composition (i.e., [^4He] and $^3\text{He}/^4\text{He}$) were conducted to evaluate potential transport mechanisms for organic compounds into the shallow groundwater (i.e., from surface spills, leaky well casings, or communication with deep shale formations). Finally, we investigated the spatial distribution of disclosed surface spills, active shale gas wells, and groundwater samples with elevated GRO and DRO to determine whether there is an increased risk associated with the collocation of natural gas extraction activities with drinking water supplies over the Marcellus Shale.

Results and Discussion

Trace levels of GRO and DRO compounds were detected in 9 of 59 (0–8.8 ppb total GRO) and 23 of 41 (0–157.6 ppb total DRO) groundwater samples, respectively. Although the highest concentrations of GRO and DRO were always detected within 1 km of active shale gas operations, this difference in concentration within 1 km ($n = 21$) and beyond 1 km ($n = 20$) from shale gas wells was only significantly higher in the case of DRO (Fig. 2; $P = 0.01$, Mann–Whitney U test; *SI Appendix*, Fig. S1). A sensitivity analysis indicated that the statistically significant difference in DRO levels persisted at 0.75–3.0 km (at 0.25-km intervals) from a gas well, as well as at 0.76 km away (the PA DEP's suggested distance of evaluation; $P < 0.01$) (18). Notably, although the highest GRO- and highest DRO-containing samples all occurred within 1 km of a shale gas well, the elevated GRO and DRO were not cooccurring (Fig. 2, *Inset*; discussed below). Finally, trace levels (<1 ppb) of VOCs, including BTEX compounds (benzene, toluene, ethylbenzene, and xylenes), were detected in six samples (10%) at levels well below the EPA's drinking water maximum contaminant level recommendations (19).

In this region, there are multiple potential sources of elevated DRO in groundwater, including (i) upward migration of naturally occurring, formation-derived organic compounds over geologic time; (ii) lateral transport of drilling muds, flowback, or produced fluids from faulty wells; (iii) leaking oil and gas waste containment ponds; (iv) input of organic contaminants from surface spills of either raw chemicals or residual fracturing fluids; and (v) leaking underground storage containers or local traffic. To evaluate these sources systematically, we used geochemical fingerprinting of inorganic constituents (i.e., Br/Cl ratios) (11), groundwater residence times (i.e., ^4He concentration) (10, 13), and dissolved methane concentrations (8, 9), coupled with our GRO, DRO, and geospatial analysis.

Upward Migration from Deep Formation. Warner et al. (11) demonstrated that shallow groundwater in some areas in northeastern Pennsylvania is saline with molar Br/Cl ratios similar to deeper Marcellus Formation water (designated “Type D” water), suggesting natural upward migration of deep saline water over geologic timescales. Using this inorganic fingerprinting approach, we found no statistical difference in GRO or DRO contents based on water type (*SI Appendix*, Fig. S2; GRO, $P > 0.05$; DRO, $P > 0.05$; Kruskal–Wallis test), suggesting that the increased GRO and DRO signals were not a result of upward migration of deeper, naturally occurring formation fluids. Furthermore, samples with elevated GRO (>5 ppb), but lower DRO (<50 ppb), which might be considered geogenic and shale derived, considering the distinct transport rates of each (i.e., retarded transport of higher-molecular-weight compounds due to slower diffusion through and higher sorptivity to porous media), were not found in Type D waters uniformly (i.e., two of three were not Type D).

Water migration from the Marcellus formation to shallow groundwater would also lead to enrichments in ^4He and fractionation of air-saturated noble gases (i.e., $^{20}\text{Ne}/^{36}\text{Ar}$) (10, 13). In contrast, we find that the highest concentration DRO and GRO samples occur in tritium-active groundwater (i.e., relatively young) and have the lowest ^4He abundance (an integrated proxy for residence time; *SI Appendix*, Fig. S3) and, therefore, the lowest apparent crustal residence times (10). This suggests that contamination occurs in the younger groundwater systems. Consequently, our results indicate that connectivity with deep

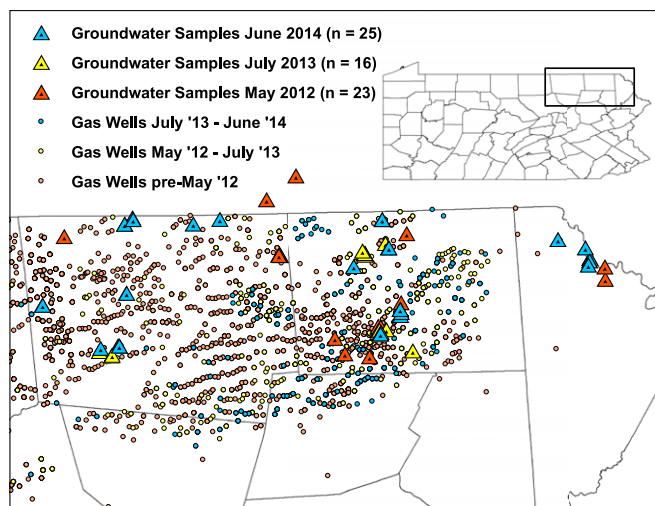


Fig. 1. Shallow groundwater sample locations and the existing active shale gas wells at those times. Five samples were collected in December 2014 and included in the June 2014 data points. Shale gas well locations were obtained from the Pennsylvania Spatial Data Access.

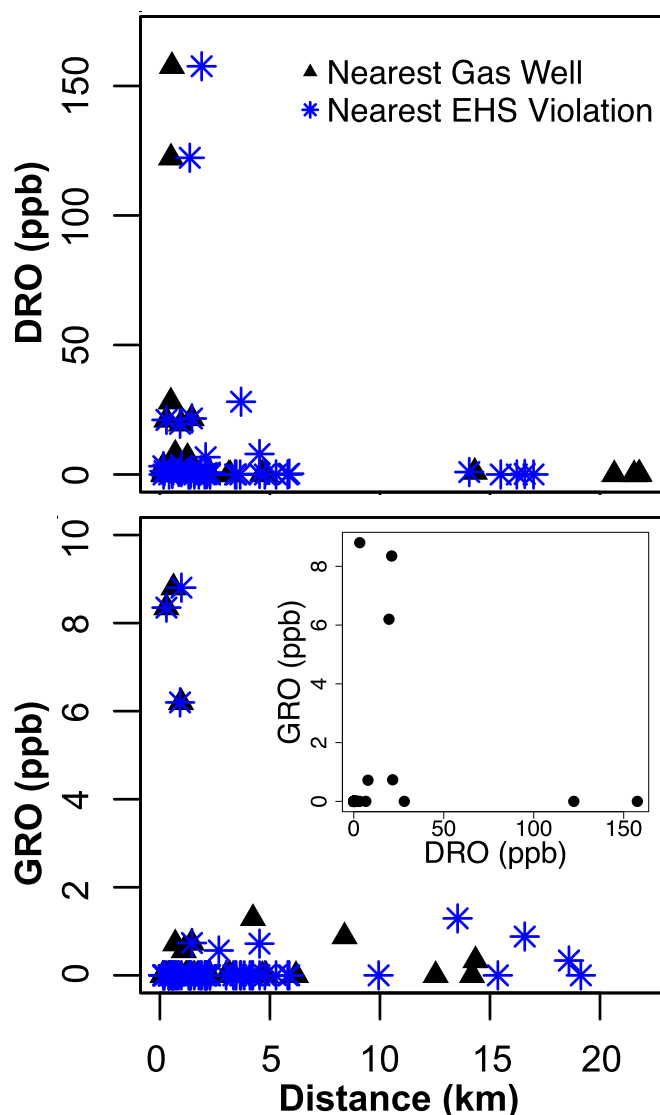


Fig. 2. DRO (Top) and GRO (Bottom) concentrations in shallow groundwater with respect to the distance from the nearest active shale gas well (black triangles) or gas well with an EHS violation (blue stars). DRO was significantly correlated ($P = 0.01$, Spearman correlation) with the distance to the nearest shale gas well and with the distance to the nearest EHS violation ($P = 0.03$). GRO was not correlated with distance to the nearest gas well ($P = 0.42$) or with the distance to the nearest EHS violation ($P = 0.36$). There was no correlation between GRO and DRO (Inset).

subsurface brines is not a dominant source of organic compounds in the shallow groundwater.

Lateral Transport from Faulty Wells. An alternate source of organic compounds to shallow aquifers could be faulty gas well casings, as poor well integrity has been documented in gas wells targeting the Marcellus Shale (6, 9, 10, 20, 21). Llewellyn et al. (7) recently reported groundwater contamination of 2-*n*-butoxyethanol, whereas others documented stray gas contamination of light hydrocarbons (e.g., nC_1 – nC_3) from poor well integrity (8–10). Also, previous studies suggested that hydraulic fracturing fluids and denser formation brines could migrate laterally, but on timescales longer than those typical of methane transport (5, 22, 23). Such migration through porous media from a well casing would result in elevated GRO with lower levels of DRO (due to the higher diffusivities and generally lower sorptivities of the lower-molecular-weight compounds), along with higher levels of methane and a salinity signature similar to that of flowback or produced waters. Methane

abundance from paired samples or previous sampling campaigns showed no correlation with GRO or DRO (*SI Appendix*, Fig. S4), and the noble gas analysis provided no evidence for fugitive gas contamination in the elevated GRO and DRO samples [e.g., low air-saturated water abundances ($[^{36}\text{Ar}]$, $[\text{N}_2]$), or $^4\text{He}/\text{CH}_4$ (10)]. Furthermore, samples with elevated GRO (>5 ppb) had relatively low methane and Br^- (<1 ppm for both). Thus, leaky well casings are an unlikely source of GRO compounds.

It is possible that a leaky well casing during slickwater injection could be a source of elevated GRO or DRO without commensurate brine or methane inputs in a relatively young well. If so, then one might expect some relationship between GRO or DRO occurrence in groundwater and the age of the nearest HVHF well if and only if chemical or bulk fluid transport times were fast relative to or on the same order as the well ages. To entertain the possibility of a well-age effect, we calculated time from the “spud date” (the drilling date of the nearest HVHF well) to sample collection. Well ages ranged broadly from 10 d to over 5 y with a fairly even temporal distribution, and levels of GRO and DRO were not correlated with the age of the well (DRO, $P > 0.05$; GRO, $P > 0.05$; *SI Appendix*, Fig. S5).

Since the preceding well age argument relies on rapid fluid transport relative to the well ages, we note that typical bulk groundwater velocities are highly variable in the sampled aquifers and on the order of 0.1 – $8.2 \text{ km} \cdot \text{y}^{-1}$ (spanning velocities in alluvium to fractured bedrock aquifers) (24), and sorption-retarded transport velocities of the chemicals we detected (described below) would be on the order of 0.02 – $7.53 \text{ km} \cdot \text{y}^{-1}$ (~ 2 mo–50 y to migrate a 1-km distance; *SI Appendix*). However, depending on the topography, hydraulic connectivity, and large pressure gradients experienced during injection (5), transport times could be faster than predicted by simple porous media transport models. For example, Llewellyn et al. (7) argue that fracturing fluids could be driven 1–3 km in a 2- to 3-y timeframe, which is reasonable for the fractured bedrock case. Although it is not possible to put an exact timeframe on the fluid transport under the hypothetical condition of leaky casing during slickwater injection, two conditions emerge: Either (i) chemical transport is slow and could not give rise to the elevated DRO compounds observed here (i.e., within 1 km and less than 2 y), or (ii) transport is faster and a relationship between DRO and well age could have been observed, which it was not. In either case, our data suggest that leaky wells are not a source of DRO to nearby groundwater wells.

Leaking Oil and Gas Waste Containment Ponds. Following hydraulic fracturing, the flowback and produced waters are often stored in polymer-lined, open, waste-containment pits, which are demonstrated sources of contamination to surface water and groundwater in cases where the liner integrity was compromised (e.g., torn, ripped, folded, or other failure due to a physical breach that allowed fluid to pass unrestricted) (25). Although many of these pits have been phased out voluntarily, many were still in use at the time of our study. Unfortunately, Pennsylvania does not maintain a publically available database of the location of the polymer-lined containment pits, and no spatial analysis between elevated DRO or GRO levels and containment pits is possible.

Nonetheless, these pits were designed to allow volatile compounds to outgas and particles to settle (26), and the residual wastes are often highly saline with a high organic content (25, 27). Thus, leakage into groundwater from such containment basins would result in low GRO levels (due to volatile outgassing) and elevated DRO, such as observed in our samples (see *SI Appendix* for discussion of potential GRO/DRO fingerprints in groundwater and flowback water; *SI Appendix*, Fig. S8). Leaking from pits with compromised liners would give rise to elevated chloride and bromide in the high-DRO samples, which was not observed (*SI Appendix*, Fig. S6). Therefore, diffusive transport of DRO through uncompromised liners could give rise to the observed chemical composition of the groundwater. However, the types of compounds revealed in our compound-specific analysis (detailed below) have very long transport times through model polymers characteristic of

Fig. 3. Locations of EHS violations associated with unconventional gas well operations as reported by PA DEP Oil and Gas Reporting website (30).

such liners. Considering the fastest-possible transport, the compound would not migrate through a 4-mm liner to the soil interface after 4 y (only $2 \times 10^{-27}\%$ of the water-side content would migrate to 1 mm depth after 4 y; model details provided in [SI Appendix, Fig. S7](#)). Over these long timescales, transport through intact pit liners could not have given rise to the DRO observed in our samples. This implies that organic chemical transport through the liners was not the primary source of material in our samples.

Surface Spills of Hydraulic Fracturing Chemicals. State databases are maintained for disclosed releases of hydraulic fracturing components at the surface, because these present a direct route for surface water and groundwater contamination. Surface releases could result in low GRO due to volatilization within weeks (28, 29), with elevated levels of DRO leaching to groundwater. Such elevated DRO with low GRO was observed in our samples, with higher DRO close to active shale gas wells (<1 km). To further evaluate the possible link between elevated DRO and proximity to the nearest well, we analyzed PA DEP violation reports. According to the PA DEP Oil and Gas Compliance Report, 5,791 violations were reported associated with 1,729 unconventional gas wells throughout the state between January 1, 2007 and June 1, 2014 (30) and classified as either “Administrative” or “Environmental Health & Safety” (EHS) violations. DRO concentrations were elevated significantly in groundwater samples in close proximity to EHS violations ($P = 0.03$, Spearman correlation; Fig. 3), but GRO concentrations were not ($P = 0.36$). Furthermore, DRO occurrence in samples within 2 km of an unconventional well with an EHS violation ($n = 20$) was statistically higher than in samples further away than 2 km ($n = 21$; $P = 0.03$, Mann–Whitney U test), whereas GRO did not show the same relationship ($P = 0.92$). Neither DRO nor GRO levels were significant at the 1- or 0.76-km cutoff distances, perhaps due to the fact that the distribution of shale gas wells with an EHS violation is spatially diffuse compared with individual shale gas well locations.

Groundwater well depth could also provide information on the nature of the flow path of the compounds detected in our samples. For example, because vertical transport times are long, a deep or shallow source might give rise to a depth-dependent concentration gradient. There was no statistically significant difference between DRO or GRO concentrations in the shallowest (<100 m) or deepest (>100 m) sampled wells (DRO, $P = 0.57$; GRO, $P = 0.89$; Mann–Whitney U test), or at any other depth cutoff (50, 75, or 125 m), and neither GRO nor DRO was correlated with depth (*SI Appendix, Fig. S9*). This could be an artifact of the scale and spatial resolution of the sampling effort. Designed to cover a large area (~7,400 km²) and constrained by well access, the groundwater

samples were separated by widely varying lateral distances (7 ± 15 and 11 ± 18 km for groundwater wells containing detectable DRO and GRO, respectively). As a result, any point source or spatially constrained “plume” of organic material could conceivably affect only a small population of groundwater wells, obfuscating any effect of well depth on the GRO or DRO concentration. Nevertheless, the samples with the highest GRO and DRO are found in groundwater wells less than 100 m deep.

Leaking Underground Storage Tanks or Local Traffic. We also explored the hypothesis that leaking underground storage tanks that typically contain gasoline, diesel, or fuel oil for both domestic and industrial use could provide a significant source of GRO and/or DRO. Leaking tank incident data obtained from PA DEP (31) showed no spatial correlations with DRO ($P = 0.95$, Spearman correlation) or GRO ($P = 0.81$) in the groundwater samples (*SI Appendix, Fig. S10*). In addition, the chemicals identified in the compound-specific analysis are not commonly stored in underground storage containers and are distinct from the chromatographic fingerprints of gasoline, diesel, or hydraulic fluids (*SI Appendix, Fig. S11*). Indeed, were these materials present, they would be readily obvious via our analytical methods (detection limits near 100 pg-L^{-1} or parts per quadrillion). Their absence implies that leaking underground storage tanks were not a source of material to the groundwater. Similarly, if local truck traffic were a source, then one might expect a distinct chemical fingerprint and correlation with distance to the nearest road. No such fingerprints (*SI Appendix, Fig. S11*) or correlations existed (DRO, $P = 0.78$; GRO, $P = 0.63$), suggesting that traffic was not responsible for the DRO observed in the studied groundwater.

Organic Chemical Fingerprinting via GC×GC Time-of-Flight MS. To further evaluate the source of elevated DRO compounds, we conducted a compound-specific investigation using GC×GC with time-of-flight MS. In particular, a subset ($n = 12$) of groundwater

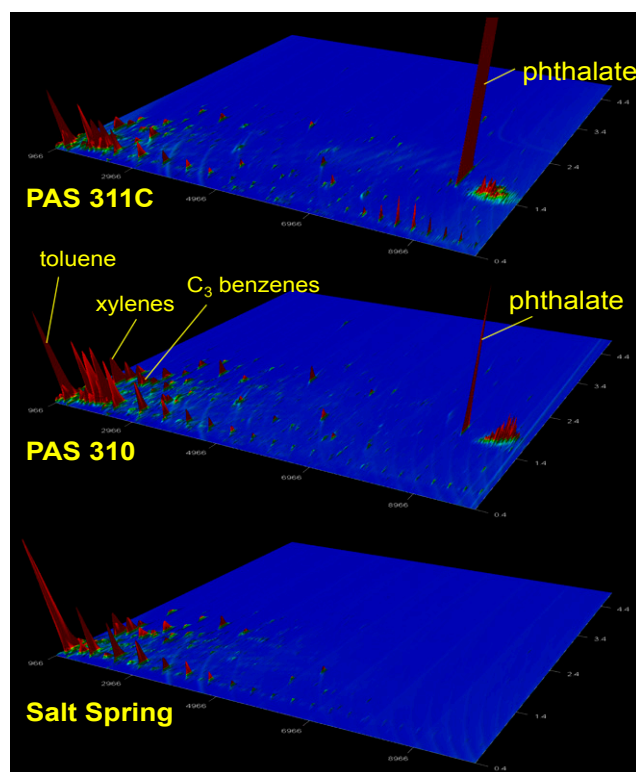


Fig. 4. GC×GC time-of-flight MS extracted ion (m/z 41) chromatograms of two shallow groundwater samples (PAS311C and PAS310) that contained bis(2-ethylhexyl) phthalate (labeled as phthalate) and the natural salt spring that did not.

liquid–liquid extracts was analyzed, including those with the highest DRO levels, along with a field blank and one natural salt spring in Susquehanna County, Pennsylvania, that is typically used to indicate the composition of natural gas and brine-rich fluids derived from shale rock sources in the northern Appalachian Basin (7–12). The samples with the highest DRO ($n = 2$) contained bis(2-ethylhexyl) phthalate (Fig. 4), whereas the salt spring, lower-DRO samples ($n = 8$) and field blank did not (*SI Appendix, Fig. S12*). Fatty acid phthalate esters [including bis(2-ethylhexyl) phthalate] are used in drilling and in hydraulic fracturing fluids (32), and bis(2-ethylhexyl) phthalate has been reported in Marcellus Shale, Barnett Shale, and Denver–Julesburg basin flowback waters (27, 32–34), in runoff and surface waters following an incident at a gas well site (35), and in multiple residential groundwater wells in Dimock, Pennsylvania, where the EPA has identified contamination directly from hydraulic fracturing operations (discerned by the onset timing of the contamination) (36).

Bis(2-ethylhexyl) phthalate is a ubiquitous chemical that is used in many industrial practices and materials, and it is difficult to attribute its presence solely to hydraulic fracturing activities. However, we present several lines of evidence that this particular phthalate is likely to be derived from HVHF activities. First, only our highest DRO samples contained bis(2-ethylhexyl) phthalate, suggesting that the compound was not derived from any step of our own analytical procedure. Second, if PVC pipes (known to contain phthalates and to be pervasive in water distribution systems) were a source of the bis(2-ethylhexyl) phthalate, then one would expect a widespread presence in the analyzed samples ($n = 12$). In contrast, it was only detected in the highest DRO samples ($n = 2$). Third, compound-specific analysis of the natural salt spring did not contain the phthalate. Thus, the presence of bis(2-ethylhexyl) phthalate likely reflects its presence in the contamination source and is not an artifact of our sampling or preparation protocol.

Curiously, this particular phthalate has relatively low aqueous solubility. In chemical disclosure databases, bis(2-ethylhexyl) phthalate is reportedly used in “perfball” form (i.e., it is transported and injected as a solid). A solubilized form of a phthalate could be derived from surface spills of flowback or produced waters, or transported through containment pit liners. However, the former would carry a brine signature, which was not observed in the high-DRO groundwater samples (*SI Appendix, Fig. S6*), and the latter has prohibitive transport timescales and could not give rise to the phthalate observed here (*SI Appendix, Fig. S7*). Consequently, our data suggest that some solubilized form of the phthalate (e.g., perfballs placed in a liquid carrier) is responsible for their appearance in shallow aquifers sampled in this study. Indeed, in all cases where the phthalate was detected, toluene was present as a cocontaminant. Further, because bis(2-ethylhexyl) phthalate is a disclosed additive in fracturing fluids, it is both (*i*) plausible that its presence in these samples is due to accidental surface releases of the parent fluids in the Marcellus region and (*ii*) reasonable given our statistical spatial analysis using the disclosed EHS violation database, as well as the complementary inorganic, methane, and helium abundance measurements. Nevertheless, one cannot rule out the possibility that the bis(2-ethylhexyl) phthalate is derived from some non-HVHF source and just coincidentally correlated with proximity to disclosed HVHF EHS violations.

Bis(2-ethylhexyl) phthalate is “reasonably anticipated to be a human carcinogen” (37). Due to the analytical challenges of obtaining clean blanks and ubiquitous industrial use, the environmental fate of phthalates has been understudied since their presumable first appearances after the advent of plastics in the 1970s (38–41). Detection of phthalates in environmentally derived samples, as well as their source apportionment, may have been overlooked out of fear of cross-contamination from other sources. However, if HVHF practices are using phthalates (which are disclosed, but not with great frequency), the environmental geochemistry community is challenged to develop robust methods to track and source apportion these materials. Careful efforts to avoid contamination (i.e., the use of precombusted, all-glass or metal materials) and accountability for all other potential local industrial sources will be critical.

Implications. To our knowledge, this is the first study of its kind to evaluate, on a regional scale, different possible mechanistic sources of organic compounds detected in drinking water wells in the Marcellus region using complementary inorganic chemical analyses and residence time approximations. Based on the evaluation of different possible mechanisms, our data are consistent with a surface-derived source of organic compounds in the study area, possibly from releases of hydraulic fracturing materials near drill sites. The question arises: Is the spill rate associated with unconventional shale gas development worse than any other industrial chemical or energy extraction activity? Unfortunately, a quantitative comparison cannot be made due to the construct of the PA DEP disclosed violation reports, for which details are limited (42). Often, ambiguous language is used to describe the nature of the violation and volume estimates of reported releases are not provided. If volume data were available, an appropriate comparison of the environmental impacts of these releases could be calculated with the ratio of volume-of-spill to total-transported volume, as is done with crude oil [in 2014, 0.00007% of all oil transported by sea was spilled (43)]. Incorporating volumetric data on releases from natural gas operations would allow direct comparisons to other energy industries. Clearly, such a report relies on accurate self-reporting or more robust monitoring (44). Further, if the PA DEP required volume and chemical identity estimates, a more accurate assessment of the relative risks due to domestic energy extraction could be constructed.

Irrespective of the reporting nuances, it is clear that surface releases of fracturing fluids are usually accidental. Therefore, it is not necessarily the hydraulic fracturing process (i.e., the fluid injection) that can lead to groundwater contamination, but rather, the existence of the operation itself (i.e., the inherent risk associated with mechanical failure and human error in industrial practice). Domestic natural gas production necessitates collocation of residential areas with extraction facilities, and, like many industrial activities, the economic benefits come with some level of environmental and public health risk.

In summary, we show that some private residential groundwater wells contained trace concentrations of organic compounds (<200 ppb DRO) in close proximity to active shale gas wells and disclosed EHS violations. Surface sources are consistent with the presence of DRO compounds in groundwater with the lowest apparent groundwater residence times. We found no evidence for direct communication of deeper formation water or injected fracturing fluids with shallow drinking water wells due to upward migration from shale horizons. This result is encouraging, because it implies there is some degree of temporal and spatial separation between injected fluids and the drinking water supply. However, shallow groundwater should be monitored over longer timescales (45) in areas of enhanced fracturing activities [e.g., where preferential faults could enhance deep-to-surface communication (46)]. Future research should also focus on investigating chemical fingerprints of shale-derived organic matter via a careful comparison of raw fracturing fluids, flowback water, and geologic formation waters.

Materials and Methods

Shallow groundwater samples were collected in precombusted glass vials over three sampling campaigns from private residential groundwater wells. Wells were purged of stagnant water until stable readings of conductivity, pH, and temperature were recorded, upstream of any treatment system. The samples were fixed with acid then stored on ice until analysis within 14–28 d. For the organic compound analysis, light hydrocarbons were analyzed using standard purging and preconcentration techniques (see *SI Appendix* for details), whereas heavier hydrocarbons were concentrated via liquid–liquid extraction into organic solvents. Compounds and compound classes were quantified via gas chromatography with flame ionization detection and qualitatively identified with confirmed standards using GC–MS. A subset of liquid–liquid extracts was interrogated using comprehensive GC×GC with time-of-flight MS. Inorganic constituents were analyzed by methods detailed in Warner et al. (11). Methane was analyzed by methods detailed in Jackson et al. (9). Noble gases were analyzed by methods detailed in Darrah et al. (10, 13). Maps and spatial data analysis were prepared with ArcMap and all statistics were analyzed with the R statistical computing platform.

ACKNOWLEDGMENTS. We thank Alissa White and Emma D'Ambro for their assistance and valuable discussion. This work was supported by Duke

University's Pratt School of Engineering and National Science Foundation (NSF) CBET Grant 1336702 and NSF EAGER Grant EAR-1249255.

- Kerr RA (2010) Energy. Natural gas from shale bursts onto the scene. *Science* 328(5986): 1624–1626.
- Malakoff D (2014) Gas revolution. The gas surge. Introduction. *Science* 344(6191): 1464–1467.
- US Energy Information Administration (2015) *Drilling Productivity Report, March 2015* (US Energy Information Administration, Washington, DC).
- Howarth RW, Ingraffea A, Engelder T (2011) Natural gas: Should fracking stop? *Nature* 477(7364):271–275.
- Vengosh A, Jackson RB, Warner N, Darrah TH, Kondash A (2014) A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environ Sci Technol* 48(15):8334–8348.
- Vidic RD, Brantley SL, Vandenbossche JM, Yoxtheimer D, Abad JD (2013) Impact of shale gas development on regional water quality. *Science* 340(6134):1235009.
- Llewellyn GT, et al. (2015) Evaluating a groundwater supply contamination incident attributed to Marcellus Shale gas development. *Proc Natl Acad Sci USA* 112(20): 6325–6330.
- Osborn SG, Vengosh A, Warner NR, Jackson RB (2011) Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing. *Proc Natl Acad Sci USA* 108(20):8172–8176.
- Jackson RB, et al. (2013) Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proc Natl Acad Sci USA* 110(28): 11250–11255.
- Darrah TH, Vengosh A, Jackson RB, Warner NR, Poreda RJ (2014) Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales. *Proc Natl Acad Sci USA* 111(39):14076–14081.
- Warner NR, et al. (2012) Geochemical evidence for possible natural migration of Marcellus Formation brine to shallow aquifers in Pennsylvania. *Proc Natl Acad Sci USA* 109(30):11961–11966.
- Llewellyn GT (2014) Evidence and mechanisms for Appalachian Basin brine migration into shallow aquifers in NE Pennsylvania, USA. *Hydrogeol J* 22(5):1055–1066.
- Darrah TH, et al. (2015) The evolution of Devonian hydrocarbon gases in shallow aquifers of the northern Appalachian Basin: Insights from integrating noble gas and hydrocarbon geochemistry. *Geochimica et Cosmochimica Acta*, 10.1016/j.gca.2015.09.006.
- Warner NR, et al. (2014) New tracers identify hydraulic fracturing fluids and accidental releases from oil and gas operations. *Environ Sci Technol* 48(21):12552–12560.
- Gross SA, et al. (2013) Analysis of BTEX groundwater concentrations from surface spills associated with hydraulic fracturing operations. *J Air Waste Manag Assoc* 63(4): 424–432.
- Office of Oil and Gas Management (2015) *Oil and Gas Reports: Water Supply Determination Letters* (Pennsylvania Department of Environmental Protection, Harrisburg PA).
- US Environmental Protection Agency (2003) *Method 8015D Nonhalogenated Organics Using GC/FID* (US Environmental Protection Agency, Washington, DC).
- Oil and gas wells, The Pennsylvania Code, Chapter 78 (1987). Available at www.pacode.com/secure/data/025/chapter78/chap78toc.html. Accessed September 24, 2015.
- US Environmental Protection Agency (2009) *National Primary Drinking Water Regulations* (US Environmental Protection Agency, Washington, DC).
- Davies RJ, et al. (2014) Oil and gas wells and their integrity: Implications for shale and unconventional resource exploration. *Mar Pet Geol* 56(0):239–254.
- Considine TJ, Watson RW, Considine NB, Martin JP (2013) Environmental regulation and compliance of Marcellus Shale gas drilling. *Environ Geosci* 20(1):1–16.
- Harrison SS (1983) Evaluating system for ground-water contamination hazards due to gas-well drilling on the Glaciated Appalachian Plateau. *Ground Water* 21(6):689–700.
- Harrison SS (1985) Contamination of aquifers by overpressuring the annulus of oil and gas wells. *Ground Water* 23(3):317–324.
- Galhar LW, Welty C, Rehfeldt KR (1992) A critical review of data on field-scale dispersion in aquifers. *Water Resour Res* 28(7):1955–1974.
- Ziemkiewicz PF, Quaranta JD, Darnell A, Wise R (2014) Exposure pathways related to shale gas development and procedures for reducing environmental and public risk. *J Nat Gas Sci Eng* 16(0):77–84.
- Rich AL, Crosby EC (2013) Analysis of reserve pit sludge from unconventional natural gas hydraulic fracturing and drilling operations for the presence of technologically enhanced naturally occurring radioactive material (TENORM). *New Solut* 23(1): 117–135.
- Hayes T (2009) *Sampling and Analysis of Water Streams Associated with the Development of Marcellus Shale Gas* (Gas Technology Institute, Des Plaines, IL).
- Serrano A, Gallego M, González JL, Tejada M (2008) Natural attenuation of diesel aliphatic hydrocarbons in contaminated agricultural soil. *Environ Pollut* 151(3): 494–502.
- Serrano A, Gallego M, González JL (2006) Assessment of natural attenuation of volatile aromatic hydrocarbons in agricultural soil contaminated with diesel fuel. *Environ Pollut* 144(1):203–209.
- Office of Oil and Gas Management (2015) *Oil and Gas Reports: Oil and Gas Compliance Report* (Pennsylvania Department of Environmental Protection, Harrisburg PA). Available at www.portal.state.pa.us/portal/server.pt/community/oil_and_gas_reports/20297. Accessed January 20, 2015.
- Department of Environmental Protection (2015) *Storage Tank Cleanup Locations* (Pennsylvania Department of Environmental Protection, Harrisburg PA). Available at www.portal.state.pa.us/portal/server.pt/community/storage_tank_cleanup_program/20605/storage_tank_cleanup_locations/1053538. Accessed January 22, 2015.
- Maguire-Boyle SJ, Barron AR (2014) Organic compounds in produced waters from shale gas wells. *Environ Sci Process Impacts* 16(10):2237–2248.
- Ferrer I, Thurman EM (2015) Chemical constituents and analytical approaches for hydraulic fracturing waters. *Trends Analyt Chem* 5:18–25.
- Lester Y, et al. (2015) Characterization of hydraulic fracturing flowback water in Colorado: Implications for water treatment. *Sci Total Environ* 512–513(0):637–644.
- Justice J (2014) *U.S. Environmental Protection Agency Pollution/Situation Report, Statoil Eisenbarth Well Response – Removal Polrep* (US Environmental Protection Agency, Washington, DC).
- US Environmental Protection Agency (2015) *Dimock Residential Groundwater Site*. (US Environmental Protection Agency, Washington DC). Available at www.epa.gov/reg3hwmd/npl/PAN000306785.htm. Accessed January 22, 2015.
- NTP (National Toxicology Program) (2014) *Report on Carcinogens, Thirteenth Edition* (US Department of Health and Human Services, Public Health Services, Research Triangle Park, NC). Available at ntp.niehs.nih.gov/pubhealth/roc/roc13/.
- Wade TL (1974) Measurements of hydrocarbons, phthalic acid, and phthalic acid ester concentrations in environmental samples from the north Atlantic. M.S. thesis (Univ of Rhode Island, Kingston, RI).
- Sanborn JR, Metcalf RL, Yu CC, Lu PY (1975) Plasticizers in the environment: The fate of di-n-octyl phthalate (DOP) in two model ecosystems and uptake and metabolism of DOP by aquatic organisms. *Arch Environ Contam Toxicol* 3(2):244–255.
- Hites RA, Biemann K (1972) Water pollution: Organic compounds in the Charles River, Boston. *Science* 178(4057):158–160.
- Mayer FL, Stalling DL, Johnson JL (1972) Phthalate esters as environmental contaminants. *Nature* 238(5364):411–413.
- Brantley SL, et al. (2014) Water resource impacts during unconventional shale gas development: The Pennsylvania experience. *Int J Coal Geol* 126(0):140–156.
- US Energy Information Administration (2014) *World Oil Transit Chokepoints* (US Energy Information Administration, Washington, DC).
- Manda AK, Heath JL, Klein WA, Griffin MT, Montz BE (2014) Evolution of multi-well pad development and influence of well pads on environmental violations and wastewater volumes in the Marcellus shale (USA). *J Environ Manage* 142(0):36–45.
- Myers T (2012) Potential contaminant pathways from hydraulically fractured shale to aquifers. *Ground Water* 50(6):872–882.
- Davies RJ, Mathias SA, Moss J, Hustoft S, Newport L (2012) Hydraulic fractures: How far can they go? *Mar Pet Geol* 37(1):1–6.

**Elevated levels of diesel range organic compounds in groundwater near
Marcellus gas operations are derived from surface activities**

Brian D. Drollette^a, Kathrin Hoelzer^b, Nathaniel R. Warner^c, Thomas H. Darrah^d, Osman
Karatum^e, Megan P. O'Connor^e, Robert K. Nelson^f, Loretta A. Fernandez^g, Christopher
M. Reddy^f, Avner Vengosh^h, Robert B. Jacksonⁱ, Martin Elsner^b, Desiree L. Plata^a

^a Department of Chemical and Environmental Engineering, Yale University, New Haven,
CT 06511

^b Institute of Groundwater Ecology, Helmholtz Zentrum München, Germany

^c Department of Civil and Environmental Engineering, Pennsylvania State University,
University Park, PA 16802

^d School of Earth Sciences, The Ohio State University, Columbus, OH 43210

^e Department of Civil and Environmental Engineering, Duke University, Durham, NC
27708

^f Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic
Institution, Woods Hole, MA 02543

^g Departments of Civil and Environmental Engineering and Marine and Environmental
Sciences, Northeastern University, Boston, MA 02115

^h Division of Earth and Ocean Sciences, Duke University, Durham, NC 27708

ⁱ School of Earth, Energy, and Environmental Sciences, Woods Institute for the
Environment, and Precourt Institute for Energy, Stanford University, Stanford, CA 94305

Corresponding Author: Desiree L. Plata; desiree.plata@yale.edu

Keywords: hydrophobic organic compounds, groundwater, high volume hydraulic
fracturing, natural gas extraction

Analytical Methods

Sample Collection: Groundwater wells were purged until temperature, pH, and electrical conductance readings were stable. Shallow groundwater samples were collected in pre-combusted 40-mL glass volatile organic analysis (VOA) vials containing 1-mL of 50% v/v hydrochloric acid (HCl), capped without headspace, and stored on ice or at 4°C until analysis. Samples were collected as close to the wellhead as possible and upstream of any water treatment systems.

Volatile Organic Compounds (VOCs): We used a modified version of EPA Method 624 and 8015D as described in Getzinger et al. (2015) [1] for quantification and qualitative identification of 50 unique VOCs via GC-FID and GC-MS (Table S1). Briefly, a 5-mL aliquot of groundwater was purged with helium on a purge and trap concentrator and subsequently transferred to a GC-FID (quantification with retention time identity confirmation) or GC-MS (confirming qualitative identification) using a Restek MegaMixTM standard mixture (Restek 502.2).

Gasoline Range Organic Compounds (GRO): We quantified GRO with the sample preparation techniques described for the VOC analysis and used a GRO mix standard for quantification (Restek #30065), following EPA Method 8015D. Chromatographic peaks eluting within the retention time window of 2-methylpentane and 1,2,4-trimethylbenzene were integrated and added together for a total signal representative of the organic compounds with boiling points within the defined gasoline range [2].

Diesel Range Organic Compounds (DRO): We used a liquid-liquid extraction technique to quantify DRO with EPA Method 8015D, as described in Getzinger et al. (2015) [1]. Approximately 500 mL of groundwater was extracted with 90:10 dichloromethane:methanol three times before rotary evaporation down to a final volume of 1-1.5 mL. Samples were quantified with a Restek DRO mix standard (Restek 31064) via GC-FID. Chromatographic peaks eluting within the retention time window of decane (nC_{10}) and octacosane (nC_{28}) were integrated and added together for a total signal representative of the organic compounds with boiling points within the defined diesel range [2].

Comprehensive two-dimensional gas chromatography: Two Leco GC×GC systems were used in this study: one coupled with a time-of-flight mass spectrometer (TOF-MS; Leco Pegasus 4D) and one coupled to a flame ionization detector (FID) detector. They were equipped with an Agilent 6890 GC (TOF-MS) and Agilent 7890 GC (FID system) and configured with split/splitless auto injectors (7683B series) and a dual stage cryogenic modulator (Leco, Saint Joseph, Michigan). Samples were injected in splitless mode. Two capillary GC columns were fitted in each GC×GC instrument, with a cryogenic modulator between the two. The first-dimension column was a non-polar Restek Rxi-1ms, (60 m length, 0.25 mm I.D., 0.25 µm film thickness) and the second-dimension separations were performed on a 50% phenyl polysilphenylene-siloxane column (SGE BPX50, 1.0 m length, 0.10 mm I.D., 0.1 µm film thickness). The modulator between the two columns operates with a liquid N₂ cold jet and dry N₂ hot jet operated at 10-15 °C above the temperature of the main GC oven. Hydrogen was used as carrier gas at a

constant flow rate of 0.90 mL min⁻¹ on the GC×GC-FID and helium at a flow rate of 1.00 mL min⁻¹ on the GC×GC-TOF.

GC×GC-TOFMS Method:

The temperature program of the main oven started isothermal at 50 °C (15 min) and was then ramped from 50 to 330 °C at 1.75 °C min⁻¹. The hot jet pulse width was 0.75 s and the modulation period was 10 s with a 4.25 s cooling period between stages. The second dimension oven was programmed from 55 °C (15 min) to 335 °C at 1.75 °C min⁻¹. The TOF-MS data were sampled at an acquisition rate of 50 spectra per second. The transfer line from the second oven to the TOFMS was deactivated fused silica (0.5 m length, 0.18 mm I.D.), constantly held at 315 °C. The TOF detector voltage was 1335 Volts and the source temperature 220 °C. The mass spectrometer employs 70 eV electron ionization and operates at a push pulse rate of 5 kHz allowing sufficient signal averaging time to ensure good signal-to-noise ratios while still operating at a high enough data acquisition rate to accurately process (signal average) spectra from the peaks eluting from the second dimension column in this high resolution separation technique (GC×GC-TOF second dimension peak widths range between 50 to 200 milliseconds).

GC×GC-FID Method:

For the GC×GC-FID analysis, 1 µL of each sample solution was injected into a 300 °C splitless injector with a purge time of 0.5 min. The first-dimension column and the dual stage cryogenic modulator resided in the main oven, whereas the second-dimension column was fitted in a separate oven, allowing for independent temperature control of all three. The temperature program of the main oven started isothermal at 40 °C (10 min) and was then ramped from 40 to 335 °C at 1.25 °C min⁻¹. The second dimension oven was programmed to remain isothermal at 45 °C for 10 minutes and then ramped from 45 to 340 °C at 1.25 °C min⁻¹. The hot jet pulse width was 0.50 seconds and the modulation period was 7.5 seconds with a 3.25 second cooling period between stages.

Inorganic constituents: Inorganic compounds were analyzed from samples collected either simultaneously with or prior to sample collection campaigns for organic compounds. Analyses were performed via methods described by Warner et al. (2012) [3].

Methane: Methane was analyzed from samples collected simultaneously with or prior to sample collection campaigns for organic compounds. Analyses were performed via methods described by Jackson et al. (2013) and Darrah et al. (2014) [4, 5].

Helium concentrations and noble gases: Helium and heavier noble gases were analyzed from samples collected simultaneously with or prior to sample collection campaigns for organic compounds. Analyses were performed via methods described by Jackson et al. (2013) and Darrah et al. (2014) [4, 5].

Field Setting

Samples were collected in northeastern Pennsylvania from three aquifers (Lock Haven, Catskill, and Alluvium) overlying the Marcellus Shale in the northern Appalachian Basin. Extensive descriptions of the geology and hydrogeologic settings can be found elsewhere [3-7].

Groundwater Contaminant Transport Calculations

Subsurface groundwater transport of bis(2-ethylhexyl) phthalate through an aquifer representative of northeastern Pennsylvania hydrogeology was estimated based on calculations and physical parameters (Table S2) from various sources [8-13].

The compound-specific, sorption-retarded transport velocity, v_c , is given by:

$$v_c = \frac{v_w}{R} \quad \text{Eq. S1}$$

where v_w is the bulk groundwater flow and R is the retardation factor given by:

$$R = 1 + \rho_s \frac{1 - \phi}{\phi} K_d \quad \text{Eq. S2}$$

and ρ_s is the sediment or soil bulk density, ϕ is the porosity, and K_d is the solid-water distribution coefficient. Here, we assume K_d is dominated by sorption to non-black carbon organic carbon phases [10], and so:

$$K_d = K_{oc} * f_{oc} \quad \text{Eq. S3}$$

where K_{oc} is the organic carbon partition coefficient and f_{oc} is the fraction of organic carbon in the porous medium.

Modeling chemical transport across a polyethylene liner

We deployed an analytical solution to calculate 1-D Fickian diffusion of bis(2-ethylhexyl) phthalate (DEHP; diethylhexyl phthalate) across a 4-mm polyethylene (PE) film, following Crank (1975) [14] and Schwarzenbach et al. (2002) [10]:

$$C_{PE}(x, t) = C_{PE}^* \operatorname{erfc} \left(\frac{x}{2(D_{PE}t)^{1/2}} \right) \quad \text{Eq. S4}$$

where $C_{PE}(x, t)$ (mass/cm³) is the concentration in the PE x cm from the side in contact with retention pond water at some time, t (s), C_{PE}^* is the concentration at the boundary of the PE equilibrated with water in the retention pond, and D_{PE} (cm²/s) is the diffusivity of DEHP within the PE. By neglecting the additional resistances to mass transport provided by a water-side diffusive boundary layer, diffusion within the soil porous media, and retardation within the PE caused by pigments (e.g., carbon black, commonly used to color the black PE liners used in containment ponds), this mathematical model simulates the furthest possible distance the concentrations front could reach within the PE film. D_{PE} for DEHP was estimated at $10^{-11.3}$ cm²/s, based on a relationship between molar volume (MV) and $\log D_{PE}$ (Eq. S5) obtained from measured D_{PE} for PAHs (polycyclic aromatic hydrocarbons) and PCBs (polychlorinated biphenyls) [15] and molar volume

estimated using the Sparc Performs Automated Reasoning in Chemistry (SPARC)
chemical property estimator.

$$\log D_{PE} (cm^2/s) = -0.014 MV (cm^3/mol) - 6.1 \quad Eq. S5$$

Using these values, DEHP will only diffuse 1 mm into the PE liner after 4 years (and only 2×10^{-27} % of the original concentration). Using an even more conservative calculation, employing 100-fold faster diffusion ($D_{PE}=10^{-9.3} cm^2/s$) indicates that only 2% of DEHP could cross a 4-mm thick PE liner by 12 months.

We confirmed the application of the analytical solution by running an explicit, finite-difference model to simulate transport across the PE using a previously prepared Matlab code modified to reflect boundary conditions relevant to a pond-PE-porous media system [16].

Table S1. Detection limits for DRO, GRO, and the targeted VOCs.

<i>Compound</i>	<i>LOD (ppb)</i>	<i>Compound</i>	<i>LOD (ppb)</i>
DRO	0.09	1,2-Dibromoethane	0.17
GRO	0.03	chlorobenzene	1.8
1,1-Dichloroethene	0.84	Ethylbenzene	2.9
Methylene chloride	1.0	m-Xylene+p-xylene	5.0
1,1-Dichloroethane	0.47	o-xylene	2.4
trans-1,2-Dichloroethene	1.2	styrene	0.26
2,2-Dichloropropane	2.4	bromoform	9.0
cis-1,2-dichloroethene	2.4	1,2,3-Trichloropropane	0.32
chloroform	2.7	Bromobenzene	0.54
Bromochloromethane	1.5	1,1,2,2-Tetrachloroethane	1.2
carbon tetrachloride	0.83	n-Propylbenzene	0.34
1,1,1-trichloroethane	0.83	2-Chlorotoluene	0.39
1,1-dichloropropene	0.83	1,3,5-Trimethylbenzene	0.19
Benzene	0.20	tert-Butylbenzene	0.37
1,2-dichloroethane	0.53	1,2,4-Trimethylbenzene	0.35
Trichloroethene	1.2	sec-Butylbenzene	0.42
1,2-Dichloropropane	0.70	1,3-Dichlorobenzene	0.55
bromodichloromethane	4.2	1,4-Dichlorobenzene	0.24
Dibromomethane	4.1	4-isopropyltoluene	0.24
cis-1,3-Dichloropropene	0.91	1,2-Dichlorobenzene	0.56
toluene	0.26	n-Butylbenzene	0.50
trans-1,3-Dichloropropene	1.2	1,2-Dibromo-3-chloropropane	7.2
1,1,2-Trichloroethane	1.5	1,2,4-Trichlorobenzene	1.0
Tetrachloroethene	0.41	Hexachlorobutadiene	2.6
1,3-Dichloropropane	0.66	Napthalene	0.88
dibromochloromethane	0.10	1,2,3-Trichlorobenzene	1.1

167

168

Table S2. Parameters used for subsurface transport calculations of bis(2-ethylhexyl) phthalate.

Variable	High velocity	Low velocity
K_{oc} bis(2-ethylhexyl) phthalate	4.998	5.078
f_{oc} ($\text{kg}_{oc} \text{ kg}_{sed}^{-1}$)	0.001	0.1
ρ_s (kg L^{-1})	1.6	2.4
ϕ	0.08	0.25
v_w (km yr^{-1})	8.219	0.109
v_c (km yr^{-1})	7.527	0.0234

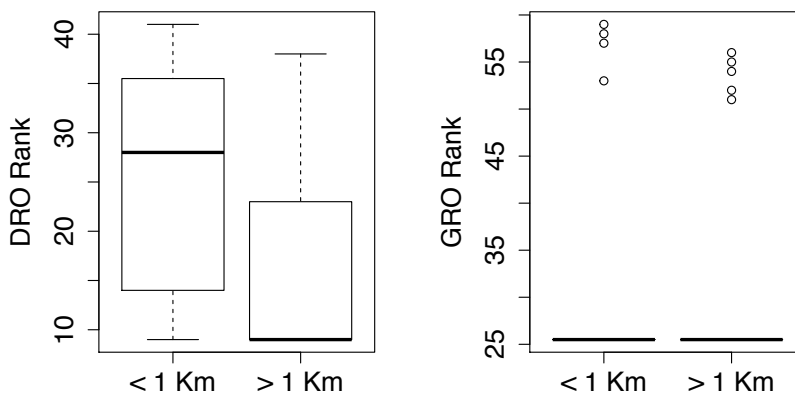


Figure S1. Ranked comparison of DRO (left) and GRO (right) between active and non-active zones (active zone defined as < 1 km from a shale gas well). There was a statistical difference in DRO between zones ($p = 0.01$, Mann-Whitney U test) and no difference in GRO between zones ($p = 0.90$).

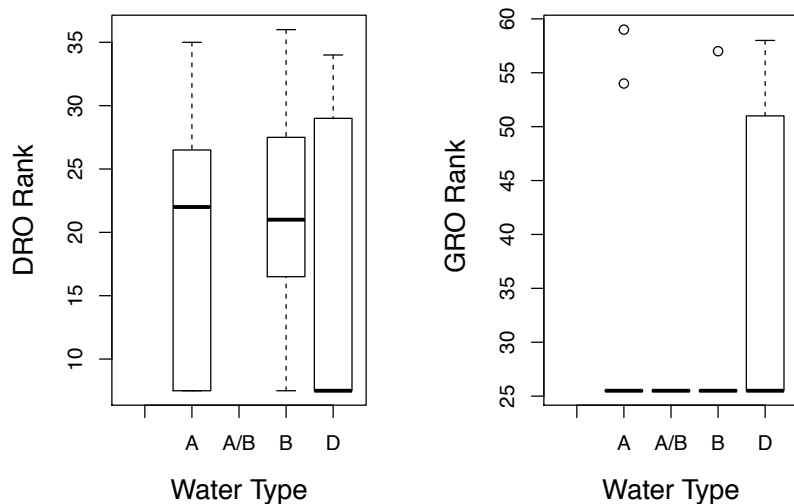


Figure S2. Ranked comparison of DRO (left) and GRO (right) with respect to different shallow groundwater types. Type D water has influence of Marcellus formation brine via natural connectivity [3]. There were no statistical differences in DRO or GRO between types (DRO, $p > 0.05$; GRO, $p > 0.05$; Kruskal-Wallis test).

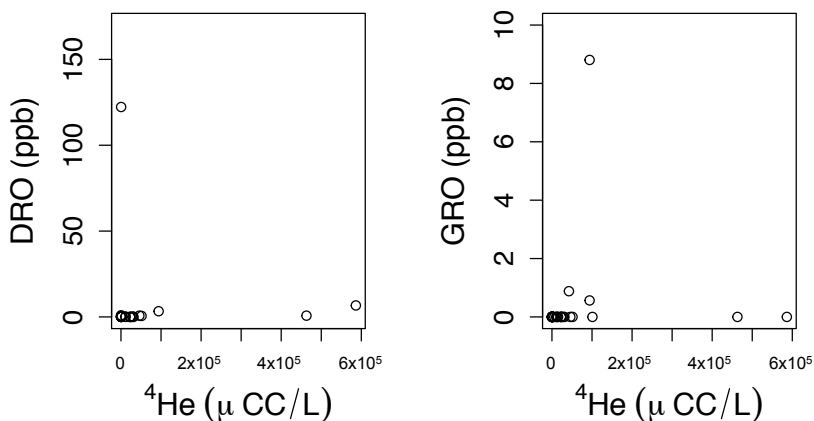


Figure S3. Relationship of DRO (left) and GRO (right) with respect to ^4He measurements taken either during the same sampling campaign or at a previous time. Statistical analysis showed no correlation (DRO, $p > 0.05$; GRO, $p > 0.05$, Spearman correlation).

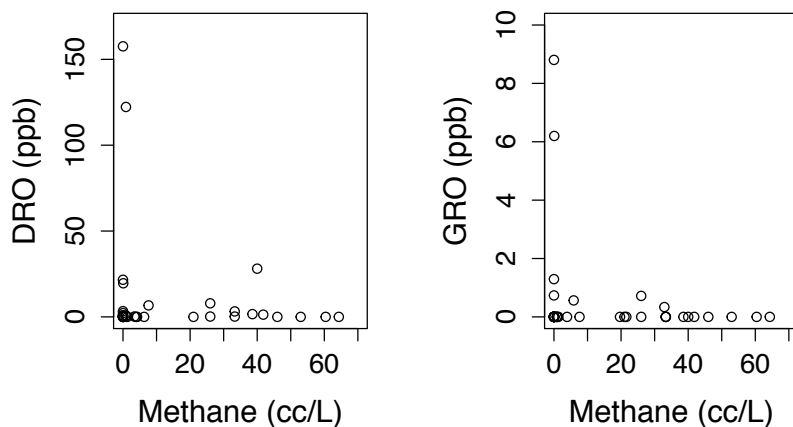


Figure S4. Relationship of DRO (left) and GRO (right) concentrations with respect to methane measurements taken either during the same sampling campaign or at a previous time. Statistical analyses showed no correlation (DRO, $p > 0.05$; GRO $p > 0.05$, Spearman correlation).

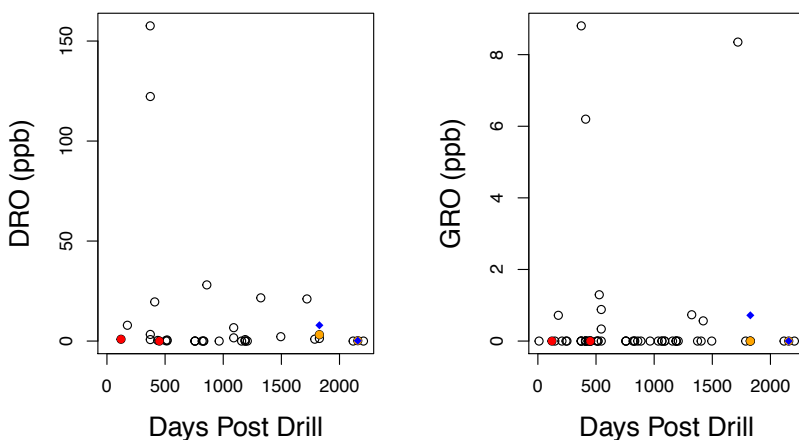


Figure S5. Relationship of DRO (left) and GRO (right) with the number of days since the nearest shale gas well was drilled. Statistical analyses showed no correlation (DRO, $p > 0.05$; GRO, $p > 0.05$, Spearman correlation). Note that only three groundwater wells were sampled repeatedly over a period of time (shown in color), and most of the data are individual wells that were sampled once during our three-year period.

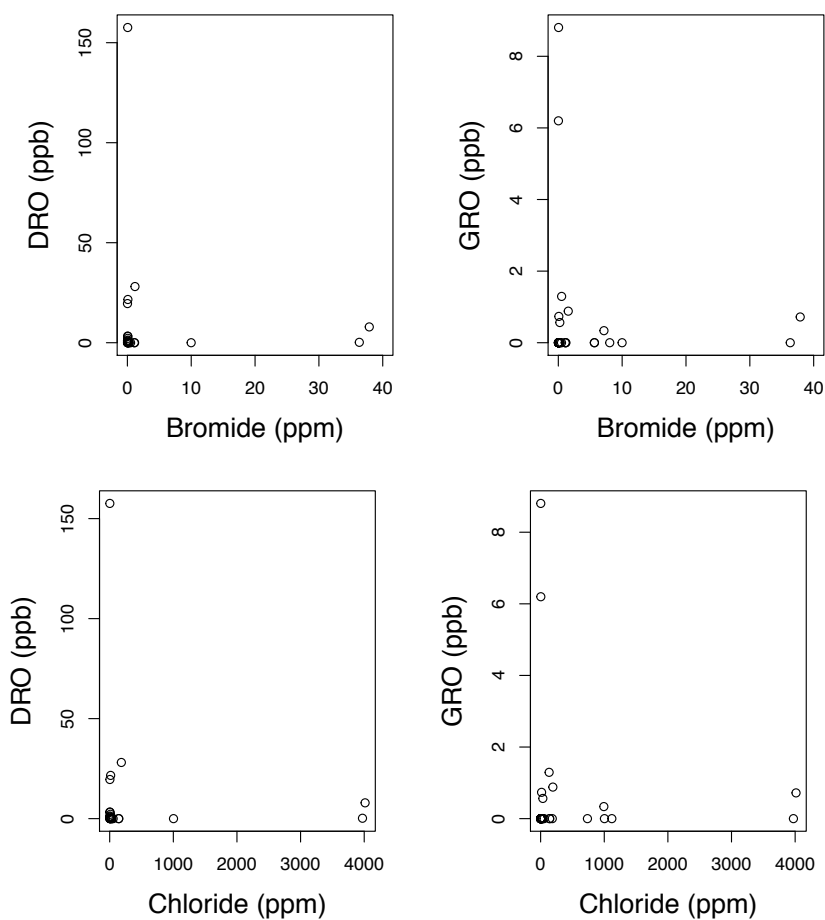


Figure S6. Relationship of DRO (left) and GRO (right) with bromide (top) and chloride (bottom). Statistical analyses showed no correlation (DRO and bromide, $p > 0.05$; GRO and bromide, $p > 0.05$; DRO and chloride, $p > 0.05$; GRO and chloride, $p > 0.05$).

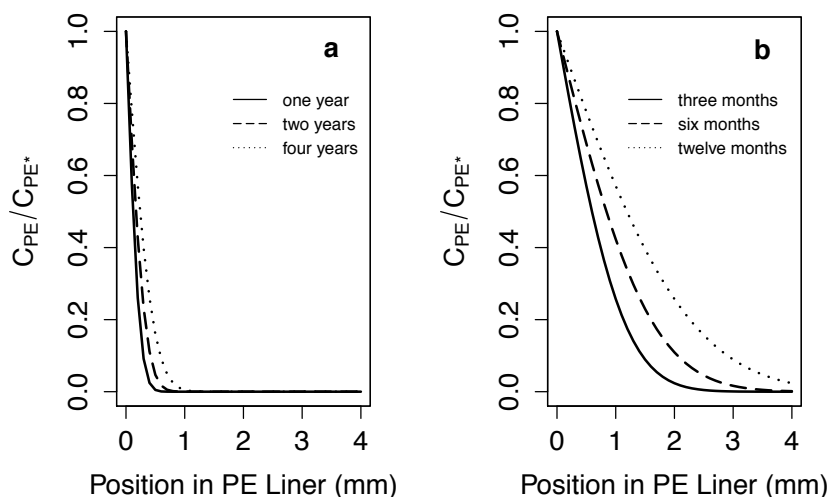


Figure S7. The 1-dimensional Fickian diffusion model for transport of bis(2-ethylhexyl)phthalate through a 4-mm thick polyethylene liner using a reasonable, estimated diffusivity of $10^{-11.3} \text{ cm}^2/\text{s}$ (a) and a 100-fold faster diffusion of $10^{-9.3} \text{ cm}^2/\text{s}$ (b). Note that this model represents the fastest-possible transport times, as we neglected resistances to mass transport provided by a water-side diffusive boundary layer, diffusion within the soil porous media, and retardation within the PE caused by pigments (e.g., carbon black, commonly used to color the black PE liners used in containment ponds).

GRO and DRO Fingerprinting

GRO/DRO ratios could be used as a tracer of these waters or potentially as an indicator of deep formation water migration. Noting that only a small number of samples ($n = 5$) had detectable levels of both GRO *and* DRO, it is clear that there is no discernable GRO/DRO “fingerprint” (Figure 2, inset, main text) in groundwater in from this region. Similarly, flowback water showed variable GRO/DRO ratios, with total GRO and DRO abundance much higher than that observed in groundwater (SI Fig. S8). Thus, GRO/DRO fingerprinting and enrichments are not viable candidates for tracing flowback water migration from either faulty well casings or leaking containment pits.

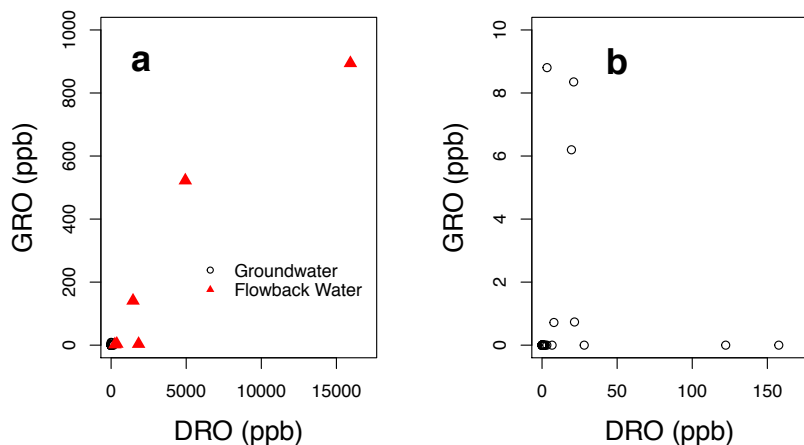


Figure S8. Comparison of GRO and DRO concentrations in shallow groundwater samples and flowback waters samples. Note that the flowback water samples (colored triangles) have much higher GRO and DRO (a), whereas groundwater samples (open circles) have comparatively low levels for GRO and DRO (b). There is no discernable GRO/DRO fingerprint in either groundwater or flowback water, as both exhibit a high degree of variability.

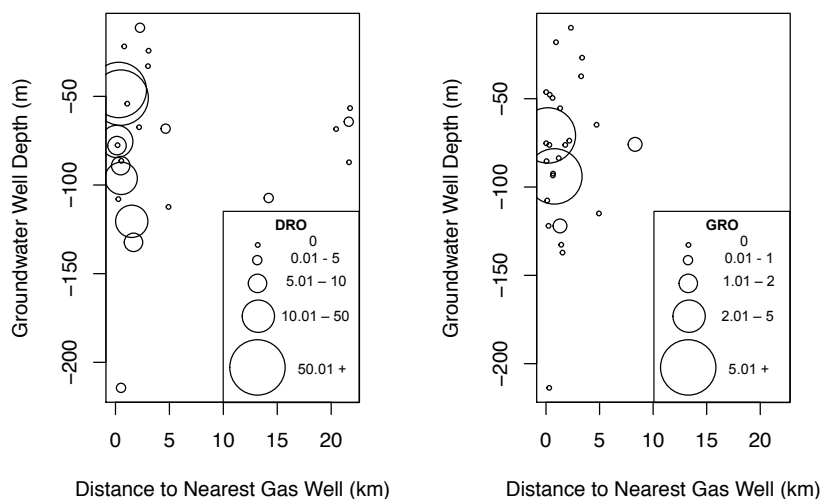


Figure S9. Relationship of DRO (left) and GRO (right) concentrations with sampled groundwater well depth and the distance to the nearest shale gas well. There was no statistical relationship between concentration and well depth (DRO $p > 0.05$, GRO $p > 0.05$). The samples with the highest concentrations were found in areas with relatively shallow well depths and close to shale gas wells.

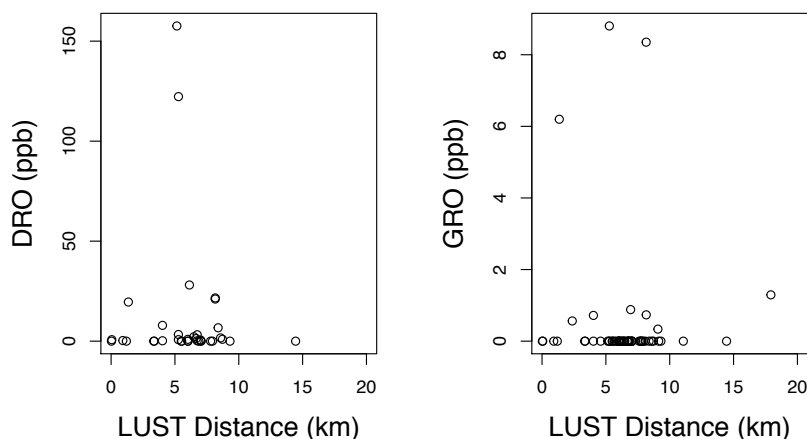


Figure S10. Spatial relationship of DRO (left) and GRO (right) concentrations with the distance to the nearest leaking underground storage tank (LUST) as reported by Pennsylvania Department of Environmental Protection (PA DEP). Statistical analyses showed no correlation (DRO, $p > 0.05$; GRO, $p > 0.05$). LUST location data available at http://www.depreportingservices.state.pa.us/ReportServer/Pages/ReportViewer.aspx?/Cleanup/Tank_Cleanup_Incidents. Accessed on November 2, 2014.

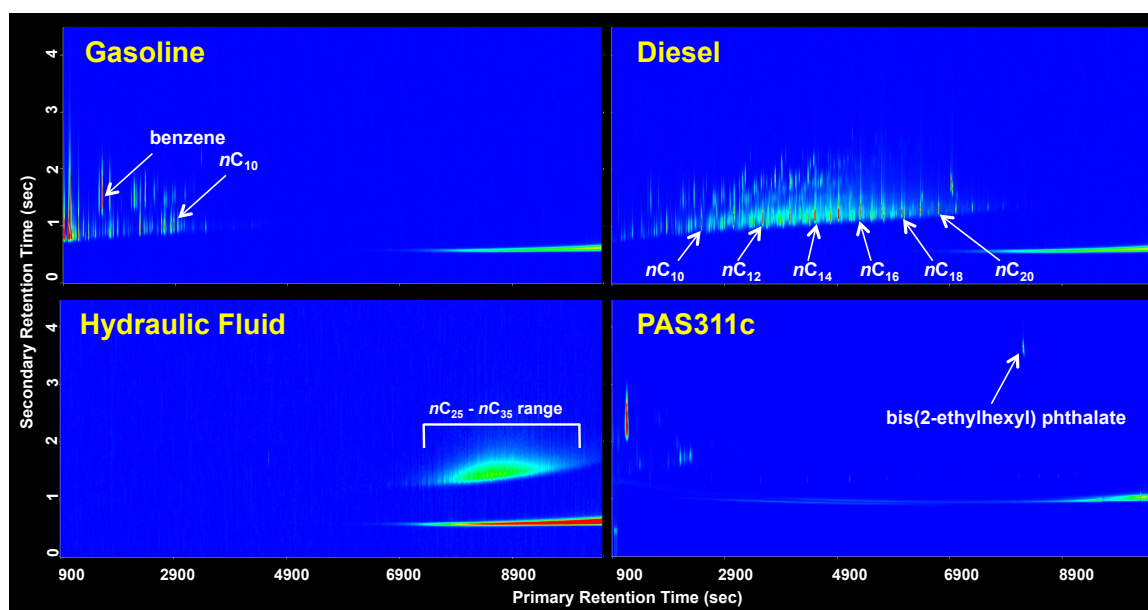


Figure S11. GCxGC-TOFMS analysis of gasoline, diesel fuel, and a hydraulic fluid common in industrial and agricultural applications, along with a groundwater samples (PAS 311C) that contained elevated levels of DRO. Note that the instrument response in the DRO region of the groundwater sample is dominated by the phthalate and unlike those of the fuels.

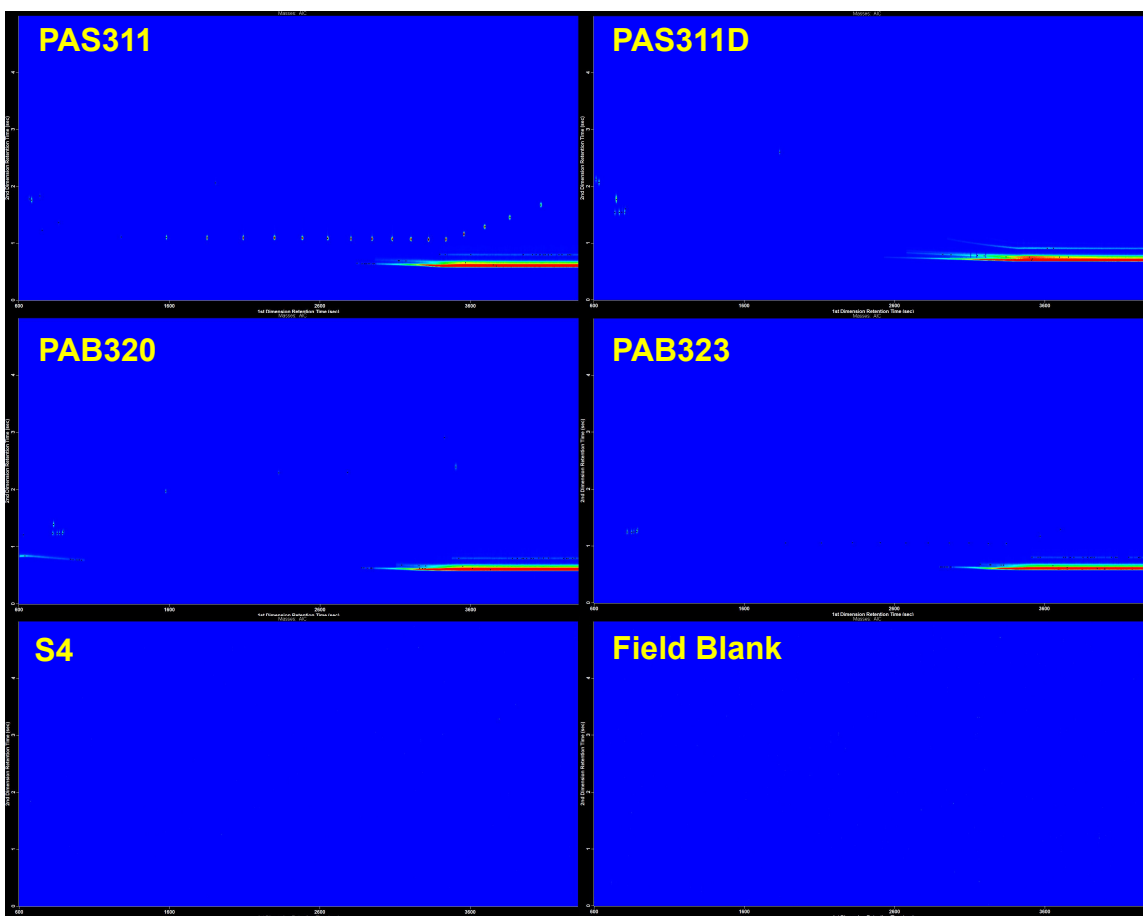


Figure S12. GCxGC-TOF-MS analysis of five (out of 23 total) low-but-detectable-DRO containing groundwater samples and a field blank indicated no detectable levels of bis(2-ethylhexyl) phthalate. These results demonstrate that phthalates, while ubiquitous in industrial products (e.g., plastics and piping material), do not often appear in groundwater in this region. The complimentary analysis of two high-DRO containing groundwater samples appears in the main text and shows the presence bis(2-ethylhexyl) phthalate.

References

1. Getzinger et al. (2015) Natural Gas Residual Fluids: Sources, Endpoints, and Organic Chemical Composition after Centralized Waste Treatment in Pennsylvania. *Environ Sci Technol* 49(14):8347-8355.
2. US Environmental Protection Agency (2003) *Method 8015D Nonhalogenated Organics Using GC/FID*. (US EPA, Washington DC).
3. Warner NR, et al. (2012) Geochemical evidence for possible natural migration of Marcellus Formation brine to shallow aquifers in Pennsylvania. *Proc Natl Acad Sci USA* 109(30):11961-11966.
4. Jackson RB, et al. (2013) Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proc Natl Acad Sci USA* 110(28):11250-11255.
5. Darrah TH, Vengosh A, Jackson RB, Warner NR, & Poreda RJ (2014) Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales. *Proc Natl Acad Sci USA* 111(39):14076-14081.
6. Faill RT (1997) A geologic history of the north-central Appalachians. 1. Orogenesis from the mesoproterozoic through the tectonic orogeny. *Am J Sci* 297(6):551-619.
7. Faill RT (1997) A geologic history of the north-central Appalachians. 2. The Appalachian basin from the Silurian through the Carboniferous. *Am J Sci* 297(7):729-761.
8. Rogers JD, Burke TL, Osborn SG, & Ryan JN (2015) A Framework for Identifying Organic Compounds of Concern in Hydraulic Fracturing Fluids Based on Their Mobility and Persistence in Groundwater. *Environ Sci Technol Letters* 2(6): 158-164.
9. Gelhar LW, Welty C, & Rehfeldt KR (1992) A Critical Review of Data on Field-Scale Dispersion in Aquifers. *Water Resour Res* 28(7):1955-1974.
10. Schwarzenbach RP, Gschwend PM, & Imboden DM. *Environmental Organic Chemistry*. 2nd ed. 2003, Hoboken, NJ: John Wiley & Sons, Inc.
11. Todd DK & Mays LW. *Groundwater Hydrology*. 3rd ed. 2005, Hoboken, NJ: John Wiley & Sons, Inc.
12. Agency for Toxic Substances and Disease Registry (ATSDR). *Toxicological Profile for Di(2-ethylhexyl)phthalate*. Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA. 1993.

- 317 13. US Environmental Protection Agency (2015). Estimation Programs Interface Suite
318 for Microsoft® Windows, v 4.11. (US EPA, Washington DC).
319
- 320 14. Crank J. *The Mathematics of Diffusion*. 2nd ed. 1975, Bristol, England: Clarendon
321 Press.
322
- 323 15. Rusina TP, Smedes F, Klanova J, Booij K, & Holoubek I. (2007) Polymer selection
324 for passive sampling: A comparison of critical properties. *Chemosphere* 68(7):1344-
325 1351.
326
- 327 16. Fernandez LA, MacFarlane JK, Tcaciuc AP, & Gschwend PM (2009) Measurement
328 of Freely Dissolved PAH Concentrations in Sediment Beds Using Passive Sampling
329 with Low-Density Polyethylene Strips. *Environ Sci Technol* 43(5):1430-1436.