A scalable workflow for the human exposome

Xin Hu¹, Douglas I. Walker², Yongliang Liang¹, M. Ryan Smith¹, Michael L. Orr¹,
Brian D. Juran³, Chunyu Ma⁴, Karan Uppal¹, Michael Koval¹, Greg S. Martin¹, David C.
Neujahr¹, Carmen J. Marsit⁵, Young-Mi Go¹, Kurt Pennell⁶, Gary W. Miller⁷, Konstantinos N.
Lazaridis³, Dean P. Jones¹

¹Division of Pulmonary, Allergy, Critical Care, and Sleep Medicine, Department of Medicine,
School of Medicine at Emory University, Atlanta, GA;
²Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount
Sinai, New York, NY;
³Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN;
⁴Huck Institute of the Life Sciences, Penn State University
⁵Department of Environmental Health, Rollins School of Public Health at Emory University,
Atlanta, GA;
⁶School of Engineering, Brown University, Providence, RI;
⁷Department of Environmental Health Sciences, Mailman School of Public Health, Columbia
University, New York, NY.

*Corresponding author; dpjones@emory.edu

615 Michael St. 205P Whitehead Building, Emory University, Atlanta GA, 30322
Abstract

Complementing the genome with an understanding of the human exposome is an important challenge for contemporary science and technology. Tens of thousands of chemicals are used in commerce, yet cost for targeted environmental chemical analysis limits surveillance to a few hundred known hazards. To overcome limitations which prevent scaling to thousands of chemicals, we developed a single-step express liquid extraction (XLE), gas chromatography high-resolution mass spectrometry (GC-HRMS) analysis and computational pipeline to operationalize the human exposome. We show that the workflow supports quantification of environmental chemicals in small human plasma (200 μL) and tissue (≤ 100 mg) samples. The method also provides high resolution, sensitivity and selectivity for exposome epidemiology of mass spectral features without a priori knowledge of chemical identity. The simplicity of the method can facilitate harmonization of environmental biomonitoring between laboratories and enable population level human exposome research with limited sample volume.
Introduction

Humans have cumulative lifelong exposure to a million or more commercial, occupational and environmental chemicals (Figure 1A). Forty-seven percent of the 86,405 chemicals registered with the United States Toxic Substances Control Act (TSCA) inventory as of June 2020 are actively manufactured, processed or imported (1), and each of these has manufacturing impurities and conversion products. Mass spectrometry (MS) provides a powerful chemical analysis platform, and targeted assays are available or possible for almost any chemical. Major, unmet analytical challenges exist for exposome research, however, as a consequence of the number of environmental chemicals and metabolites, chemical diversity, low abundance (2) and lack of readily available authentic standards (3-5). As a result, few are routinely biomonitored in humans, and many of the commercial chemicals, along with legacy pollutants from prior commercial uses, biotransformation products and impurities, exist as "dark matter" of the human exposome (2).

We focus here on an analytical workflow to operationalize untargeted environmental biomonitoring to gain information on known as well as unknown exposures for human exposome research. In contrast to targeted MS analysis, which is developed to measure specific chemicals, untargeted exposome analysis includes measures of known chemicals which are "identified" by MS criteria, and also other MS signals which are unidentified because they have not been associated with known chemicals by MS criteria (6). These unidentified signals also include chemical contaminants that are known and uncharacterized, as well as reaction products that are effectively unknown to science; capability to measure these unidentified chemicals in biologic samples is essential to enable population health studies of the dark matter of the exposome.
To address the challenge to measure large numbers of identified as well as un-identified environmental chemicals in human samples, we sought to develop a workflow with gas chromatography-high-resolution mass spectrometry (GC-HRMS) which minimizes operator and instrument variation and can be applied consistently for untargeted analysis of tens of thousands of samples. GC-coupled analysis is important because many environmental chemicals are hydrophobic, semi-volatile and do not ionize well with popular liquid chromatography (LC)-MS methods. GC-MS is robust, with universally applicable retention time indices and highly reproducible spectra for database development (7, 8). The high mass accuracy and mass resolution of GC-HRMS (9) in full scan mode further enhances resolving power to obtain extensive chemical coverage in complex biological matrices. As opposed to more targeted acquisitions with single-ion-monitoring (SIM) or data-independent-acquisitions (DIA), collection of all spectral features enables measurement of known targets based upon libraries of authentic standards, while reproducibly measuring and preserving information for unidentified MS features.

With recognition that GC-HRMS in full-scan mode provides reproducibility, extensive coverage, and quantifiable data, we focused on key obstacles to implementation of GC-HRMS in exposome epidemiology, specifically variability in sample extraction and automated extraction and assembly of the complex data. In LC-HRMS, a single-step sample extraction procedure (10) improved delivery of data following FAIR principles (11), especially interoperability of data, by eliminating differences due to multistep sample processing. As a result, LC-HRMS is transforming environmental health research through delivery of omics scale exposure and biologic response data with improved sensitivity, throughput and affordability (12-14). In contrast, workflows for targeted GC-MS of environmental chemicals, such as polychlorinated
biphenyls (PCBs), polybrominated diphenyl ethers (PBDEs), polybrominated biphenyls (PBBs) and chlorinated pesticides, use multistep processing to remove biologic matrix effects and enrich the targeted chemicals (Figure 1B) (15, 16). Losses of semi-volatile chemicals can occur from dry-down steps following solid-phase extraction (SPE) and liquid-liquid extraction (LLE). Variable loss and contamination can also occur at each processing step. For targeted analyses, inclusion of stable isotopic internal standards overcomes these limitations. For unidentified MS features, however, variability in recovery or loss cannot be evaluated directly (Figure 1B), limiting use for discovery of unidentified environmental chemicals associated with health outcomes.

In the present study, we developed a single-step sample preparation method, which we term express liquid extraction (XLE), for use with GC-HRMS to minimize recovery variability and provide extensive coverage of both identified and unidentified environmental chemicals. We evaluated chemical recovery and used the National Institute of Standards and Technology (NIST) Standard Reference Material-1958 (SRM-1958) to test quantification of environmental chemicals with stable isotopic standards. We established validity of single-point quantification by reference standardization (17) and show that XLE with GC-HRMS supports quantification of environmental chemicals in diverse human samples, including plasma, lung, thyroid and stool. We further show a computational workflow which enables untargeted analysis of identified and unidentified environmental chemicals in a form suitable for exposome epidemiology.
Results

Sample preparation for GC-HRMS analysis. Starting initially with a QuEChERS procedure (19, 20), we systematically varied solvent composition, volume and extraction time to obtain a simplified procedure with a minimal number of steps and possibilities for contamination and variability in recovery (Figure 2A). In testing these procedures, results showed that at the low levels of abundance in many samples, analyses were subject to contamination by environmental chemicals in solvents and reagents used for QuEChERS, emphasizing the need for blank analyses in quality control (Figure 2A). The final procedure that allowed high reproducibility, minimal contamination, high sample throughput and maximal coverage of chemicals for XLE used formic acid and hexane:ethyl acetate (2:1) with internal standards, vigorous shaking, centrifugation, and transfer of the organic phase to a new tube, which contained pure MgSO₄ to remove water. Results of total ion intensity calculated by the sum of all MS peak intensity showed that, the signals in saline extracted by XLE (i.e. method blank) matched the baseline signals found in directly injected isooctane solvent (Figure 2B).

Validation of XLE quantification using standard reference material. High recovery of [13C] labelled chemicals was obtained for important classes of environmental chemicals (PCB, PBDE, PAH, chlorinated pesticides) in NIST SRM-1958 (Figure 2C). Recoveries ranged from 90±5% for [13C₆]anthracene to 78 to 87% for congeners of universally [13C] labeled PCBs, PBDEs and chlorinated pesticides, with only [13C₁₂]p,p’-dichlorodiphenyldichloroethylene (p,p’-DDE) having recovery <70% (Figure 2C). The simplified extraction procedure therefore provides a generally efficient recovery of environment chemicals in the organic phase. Chemicals with less recovery such as p,p’-DDE can be quantified as long as sample processing is consistent between operators and at different processing times. We evaluated interoperability
and found that comparable results (no differences shown by all raw $P > 0.05$, one-way ANOVA) were obtained with the procedure at two different times, 7 months apart, and by two different operators (Figure 2D).

We evaluated quantification using XLE by testing 54 different chemicals (PCB, PBDEs, chlorinated pesticides) in SRM-1958 using external calibration curves (0.05 to 2 ng/mL) and comparing to the known concentrations (21). We identified all 43 PCBs that are reported in the range of 46.6 to 490 ng/kg in SRM-1958 certificate of analysis (issue date: 11 October 2018; Figure 2F); of the 38 PCBs with certified values, quantification without adjustment for recovery was reproducible (19 PCBs presented in Figure 2E, all 38 in Supplemental Table S1). Eleven out of 13 PBDE/PBBs and all 17 organochlorine pesticides were identifiable (Figure 2F); 16 were reproducibly quantified in this experiment (Figure 2E and Supplemental Table S1).

Therefore, XLE provides sufficient recovery to support accurate absolute quantification of a broad range of environmental chemicals. Overall, XLE supported measurement of 71 out of the 73 chemicals that are in the ng/kg range in SRM-1958 (Figure 2F and Supplemental Figure S1).

Reference standardization for XLE-based exposome analysis. Absolute quantification of chemicals in human samples is often complicated by ion suppression effects of the biologic matrix on chemical detection by mass spectrometry. Stable isotope dilution addresses both recovery issues described above and matrix effects on ionization efficiency and is therefore ideal for quantification in targeted chemical analysis. Use of stable isotopic standards is not practical for untargeted analysis of large numbers of environmental chemicals or for un-identified environmental chemicals. For LC-HRMS analyses, single-point quantification by reference standardization has been established as a useful alternative. To provide this utility for GC-
HRMS (Figure 3A), we tested the ability of reference standardization (18) as a simple and practical approach in untargeted exposomics to estimate chemical concentrations using a single point calibration. We analyzed 20 human plasma samples and performed reference standardization of 17 chemicals based on SRM-1957 and SRM-1958 that were processed in parallel to the plasma samples. The selected chemicals included 7 PCBs, 7 chlorinated pesticides and 3 PBDEs that are detected in human biomonitoring studies (22, 23). To provide additional confirmation, the chemicals were quantified by external standard curves (ranging from 0.05 to 2 ng/mL) with recovery efficiency determined relative to spiked internal standards (Figure 3B). Among the 20 plasma samples, the measured concentrations from two quantification methods were similar (Figure 3B), with |ρ|>80% for all 17 chemicals, and |ρ|>90% for 14 chemicals (Figure 3C, Supplemental Table S2). Thus, the results validate use of reference standardization as a simple approach for quantification in a high throughput XLE workflow with GC-HRMS.

Application of the XLE workflow for analysis of environmental chemicals in human samples. We analyzed 80 archival samples from individuals (57 females, 23 males; aged 41 to 68 y) without known disease or occupational or environmental exposures of concern as a pilot to test the utility of XLE in large-scale human biomonitoring studies. For targeted environmental chemical analysis, we selected 378 chemicals from an in-house database for which dilution conditions (0.05 to 2 ng/mL) were relevant for general-population analyses (Supplemental Table S3). Using a requirement for at least 3 co-eluting accurate mass m/z features (± 5 ppm) within 30 s of database retention time, we identified 49 chemicals belonging to various environmental chemical classes. An unsupervised 2-way hierarchical cluster analysis (HCA) of log transformed intensity showed clustering according to chemical class (Figure 4A). In
particular, persistent chemicals were highly correlated with each other (all raw \( P < 0.001 \)), including \( p,p' \)-DDE, PCBs 153, 180, 138, 118 and 74, PBDE-47, hexachlorobenzene (HCB) and \textit{trans}-nonachlor (Figure 4B). Results showed a general increase of chemical levels with increasing age quartiles (Q3 and Q4: 53 to 68 versus Q1 and Q2: 41 to 52) using unsupervised clustering, a trend particularly evident for the cluster of \( p,p' \)-DDE, PCBs 153, 180, 138, 118 and 74, PBDE-47, HCB and \textit{trans}-nonachlor. Examination of data according to body mass index (BMI) showed that individuals with BMI ≥ 40 had lower levels of environmental chemicals, which may be attributed to high lipophilicity and propensity to distribute in adipose tissue versus plasma (Figure 4A). Quantification with reference standardization (Supplemental Table S4) showed that use of two SRM samples with differing environmental chemical concentrations can overcome variable batch effects in quantification for large-scale studies. Examples of the most frequently detected chemicals shows that overall distributions were positively skewed by a small subset of individuals with high concentrations (Figure 4C).

We tested the general utility of XLE in a variety of human biological samples by analyzing human lung and thyroid tissues and stool samples. We quantified 32 environmental chemicals in 11 human lungs, with HCB, PCB-28 and PCB-18 being most frequently detected (10 out of 11) (Supplemental Table S5). The commonly detected chemicals in human plasma were detected less frequently in the lung. For the 11 lungs, \( p,p' \)-DDE was detected in eight, PCB-153 in five, PBDE-47 and PCB-138 in four and PCB-180 in three. Although the plasma samples were from non-diseased individuals and the lungs were both diseased and non-diseased individuals, HCA results suggest that environmental chemical profiles in human lung may be very different from plasma. Indeed, quantification showed levels of PCB-18, PCB-28 and HCB, which are relatively volatile, were 3 to 10-fold higher in the lung than the plasma (PCB-18:
0.033 vs 0.004 ng/g, PCB-28: 0.050 vs 0.005 ng/g, HCB: 0.102 vs 0.032 ng/g, Figure 4C), indicating a potential contribution of inhalation exposure to the more volatile environmental contaminants.

In the small number of thyroids that was analyzed with XLE, 14 environmental chemicals were quantified (Supplemental Table S6). The most prevalent was p,p’-DDE, detected in 4 out of 5 thyroid samples, with median concentration (2.20 ng/g). The amounts of individual chemicals were highly variable among the individuals, and the small number of samples precludes any generalization. Nevertheless, HCA of correlation matrix showed high correlation of chemicals measured in the thyroid samples was similar to that in the lung and plasms (Figure 4B).

Human stool samples, as a noninvasive matrix, have unique value in exposome research (24, 25) but have not been extensively studied for environmental chemical exposures. For lipophilic and unabsorbed dietary environmental chemicals, stool is a primary route of elimination (25) and can therefore provide useful information on body burden and clearance of chemicals (24). In a pilot analysis of six human stool samples, we detected 52 and quantified 21 environmental chemicals, with HCB found in all samples (Supplemental Table S7). Quantification of HCB showed a median concentration of 0.057 ng/g. HCA of correlation matrix showed co-exposures of chemicals are likely as shown in the plasma, lung and thyroid (Figure 4B). Along with the pilot studies of plasma, lung and thyroid, these results establish feasibility to use XLE with GC-HRMS for high-throughput quantification of environmental chemicals in biomonitoring.

*Analysis of unidentified environmental chemicals.* An important goal of untargeted biomonitoring is to develop procedures to study the unknown exposures of the human exposome.
Suspect screening with chemical databases enables collection of information on known chemicals but not on other chemicals, contaminants and transformation products. In studies of LC-HRMS of human plasma, half of the m/z features associated with health outcomes are unidentified (14). In principle, statistical and bioinformatics analyses can be performed on accurate MS features obtained from GC-HRMS without chemical identification, and these features can then be used with index chemicals to define retention to obtain characterization for identification by database searching or deposition into data libraries (Figure 5A).

We tested the feasibility of XLE with GC-HRMS to capture information on unidentified m/z in a usable form for entry into a data library. We selected 4,747 m/z features detected within a 2-min interval (18.00 to 20.00 min) from the human plasma analyses and performed unsupervised spectral clustering with RAMclustR (26). Results provided 2,413 features aligned into 341 spectral clusters, potentially representing 341 chemicals (Figure 5B). The majority of the 341 clusters were unidentified based on library search against the NIST/EPA/NIH Main Library containing more than 100,000 compound MS spectra (Xcalibur v 4.2.1). Selected examples show potential to use accurate MS signal to aid in chemical identification via computational tools and fragmentation database (Figure 5B). For instance, in Figure 5C, using MS-Finder (ver 3.42) (27, 28), we predicted that Cluster 74 (C74) may contain a series of hydrocarbons with predicted elemental compositions containing one oxygen. Cluster 106 (C106) may contain a series of m/z features with accurate mass matches to halogenated hydrocarbons containing Br and Cl. Cluster 208 (C208) may contain a series of hydrocarbons matching elemental compositions containing 2 oxygens. The results verify that XLE can be used with GC-HRMS to perform untargeted analyses of human biospecimens to characterize unidentified m/z with data output for cumulative libraries to support exposome research.
Discussion

XLE with GC-HRMS provides a high-resolution exposomics workflow to address a critical need for public health research, namely, an affordable, interoperable method for population research to study health effects of an extensive number of low abundance chemicals to which humans are exposed in a single biological sample. As much as 85% of chronic disease is determined by the exposome (29, 30), yet detailed understanding is only available for a relatively small number of highly hazardous chemicals. Detection of less hazardous exposures requires large populations, and cost for targeted assays of large numbers of chemicals generally precludes study of large populations.

The presented XLE method can overcome these limitations by providing a standardized method to obtain quantitative information on known environmental chemicals while also providing information on thousands of unidentified chemicals. Based upon a previously developed multidimensional framework using LC-HRMS to characterize the metabolome (14), XLE with GC-HRMS can be used to advance omics scale analysis of the chemical exposome (Figure 6). The simple XLE processing procedure improves interoperability, and the multidimensional framework supports collection and communication of MS information for unidentified, as well as identified, chemicals for epidemiologic research. The information structure preserves key information about unidentified signals for subsequent chemical identification, i.e., accurate MS allows prediction of possible elemental compositions, isotopic internal standards are useful as index chemicals to facilitate retention time alignment, and pooled reference materials provide quantitative reference and allow simultaneous identification and quantification of known environmental chemicals (Figure 6). Storage of this information, along
with other available chemical and physical data, into standardized cumulative libraries will provide a resource for exposome research.

Uncontrolled loss of chemicals is a major challenge in traditional sample preparation for analysis of environmental chemicals. Volatility, partitioning between solvents, and binding affinity to absorbents, vary substantially among different chemical classes. Multiple steps including evaporation of solvent, adsorption, reconstitution in traditional SPE and LLE methods introduce variable loss of different chemicals. While this can be addressed by using stable isotope labelled standards, this is impractical for omics scale analyses of environmental chemicals. XLE minimizes uncontrolled variations by minimizing the number of sample preparation steps. Results with three quantitative approaches, \(^{13}\text{C}\)-labeled internal standards, external calibration curves and reference standardization relative to NIST-1958, showed efficient recovery and quantification of PCBs, PBDEs and chlorinated pesticides with XLE. Lower efficiencies are expected for dissimilar structures with greater polarity, a limitation which can be overcome by parallel use of LC-HRMS methods for exposome analysis (13).

As previously shown for LC-HRMS data, reference standardization provides a practical approach for quantification relative to pooled reference material that is processed and analyzed concurrently with unknown samples. Reference standardization assumes a linear relationship between analyte concentration and instrument response, which was validated with authentic standards for the 378 chemicals reported here. We showed that this single-point calibration method performs comparably to quantification by traditional six-point calibration for GC-HRMS (Supplemental Table S2) as well as relative to internal stable isotopes. Using reference standardization, levels for environmental chemicals in the 80 CHDWB plasma samples were within the range of total population serum values reported by National Health and Nutrition
Examination Survey (NHANES) in 2003-2004, e.g., p,p’-DDE (0.971 [80 subject median] vs 1.29 [NHANES geometric mean] ng/g whole weight), HCB (0.032 vs 0.092 ng/g), trans-nonachlor (0.028 vs 0.089 ng/g), PCB-118 (0.021 vs 0.037 ng/g) and PCB-180 (0.066 vs 0.092 ng/g). The values for the CHDWB plasma are expected to be lower than 2003-2004 NHANES levels as the CHDWB participants were surveyed after 2006.

Human specimens other than blood samples are increasingly available from biorepositories and provide important opportunities for exposome research, as they may allow for assessments of exposure within target tissues of toxicity (31, 32). Historically, white adipose tissues were sampled as a storage and effector site for persistent environmental pollutants (33). As environmental factors are now recognized to contribute to the origins and expressions of many human diseases, more extensive analysis of clinically annotated tissue samples is needed. The present results showing quantification of multi-class chemicals demonstrate that XLE and reference standardization provide a generalizable approach for human specimens. Compared to human plasma, the lung had higher levels of the more volatile PCBs and HCB and lower levels of the less volatile chemicals indicating the potential effects of respiratory exposure to organ chemical profiles (Figure 4C). On the other hand, subjects with severe obesity showed lower plasma levels of common persistent pollutants than subjects with normal BMI ($P<0.01$ for HCB, PCB-138, 153, 170 and 180), indicating a negative effect of body fat on circulating levels of lipophilic chemicals (34). Stool had relatively low environmental chemical content, but is important as a route for elimination of lipophilic chemicals from fat reservoirs (24) and source for information on recent exposures from diet (35). The results obtained reflect inter-organ interactions in chemical uptake, distribution and clearance, and stress the value of XLE with GC-HRMS for analysis of diverse human tissue types.
XLE with GC-HRMS provides an important step forward for human exposome research by providing a method to maximize capture of information on unidentified chemicals in human samples. Use of such information is facilitated by available annotation and library search software and algorithms, such as RAMclustR and MS-Finder. These tools cluster spectral ions and predict potential structure information on unidentified spectral features. Even without identification, statistical tests can be applied to detect mass spectral features associated with health outcomes. A rigorous interoperable data reporting structure will enable research community efforts to advance chemical identification.

In conclusion, XLE with GC-HRMS addresses a critical need for methods to deliver omics-scale biomonitoring data for exposome epidemiology in an automated, high-throughput and affordable manner. The method provides measures for both known and unidentified MS features to test for associations with disease. For known chemicals, an automated workflow integrates computational methods for data extraction, pre-processing and spectral annotation; with reference standardization, the method supports quantification of environmental chemicals. Tests with plasma, lung, thyroid and stool samples showed that the method is suitable for multiple sample types. The simplicity of the method facilitates harmonization of exposome analyses, enabling development of cumulative human exposome databases to include information on tens of thousands of chemical exposures in tens of thousands of individuals.
Materials and Methods

Standards and reference materials. As an initial screening, we have purchased and examined spectra and chromatographic information for over 900 authentic chemical standards (1 to 20 ng/mL) in the form of single chemical or mixture of ≤40 chemicals, from Cambridge Isotopes (Tewksbury, MA) and AccuStandard (New Haven, CT). A total of 556 chemicals showed high detection sensitivity and linearity (|ρ|>0.98 over 0 to 20 ng/mL); 378 of these analyzed under dilution conditions in isooctane (0.05 to 2 ng/mL) relevant for human analyses were entered into a data library (Supplemental Table S3). In addition, two $^{13}$C labeled chemicals $[^{13}C_{12}]$PCB-28 and $[^{13}C_{12}]$PBB-153 were used as volumetric internal standards added to the final extract, and nine $^{13}$C labeled chemicals (99% isotope enrichment for each) were spiked as recovery standards to estimate chemical recovery efficiency by XLE: $[^{13}C_{12}]$PCB-101, $[^{13}C_{12}]$PCB-153, $[^{13}C_{12}]$PCB-180, $[^{13}C_{12}]$PBDE-47, $[^{13}C_{12}]$PBDE-99, $[^{13}C_{6}]$anthracene, $[^{13}C_{10}]$mirex, $[^{13}C_{6}]$cis-permethrin, and $[^{13}C_{12}]$p,p’-DDE.

National Institute of Standards & Technology [NIST] Standard Reference Materials (SRM) 1958 was analyzed with ≥3 aliquoted replicates per experiment to validate and test reproducibility of detection and quantification. Two sets of SRM, NIST SRM-1958 (human serum fortified with environmental chemicals) and NIST SRM-1957 (non-fortified human serum) were run in every batch of 20 samples to support quality control and quantification using reference standardization, a protocol that was previously validated in high-resolution mass spectrometry data for LC methods (18). In this protocol, individual chemical concentrations in unknown samples are estimated by comparison to a concurrently analyzed, pooled reference sample with known chemical concentrations.
**Human materials.** 200 µL plasma samples from 80 individuals without known disease were randomly selected from archival samples obtained from the Center for Health Discovery and Well Being (CHDWB) cohort of approximately 750 individuals. The original study was conducted under Emory Investigational Review Board (IRB approval No. 00007243) and included both genders and individuals self-identifying as white, black, Hispanic and Asian. Whole human lungs were from eleven individuals; 4 were end-stage diseased lungs acquired from the Emory Transplant Center (IRB approval No. 00006248), one was from Cystic Fibrosis Biospecimen Registry at Emory University (IRB approval No. 00095116) and 6 non-diseased post-mortem lungs were obtained through the International Institute for the Advancement of Medicine (IIAM, Edison, NJ) or Novabiosis (Morrisville, NC). Whole human thyroids from five individuals were post-mortem tissues that were acquired from National Disease Research Interchange (NDRI, Philadelphia, PA). Human stool samples were obtained from PROGRESS, a biobank for cholestatic liver disease (IRB 670-02; Mayo Clinic, Rochester, MN), to test suitability of XLE for environmental chemical measurements in stool of patients with primary sclerosing cholangitis (PSC).

**XLE sample extraction.** For plasma samples, 50 µL formic acid (Emprove® Essential DAC, Sigma-Aldrich) was added to 200 µL plasma and immediately followed by addition of 200 µL hexane – ethyl acetate (2:1 v/v, ≥99% pure, Sigma-Aldrich) containing the internal standards (final concentration: 1 ng/mL). The mixture was shaken vigorously on ice using multi-tube vortexer (VWR VX-2500) for 1 h and centrifuged at 1000 g, 4 °C for 10 min. The organic supernatant was transferred to a new tube with 25 mg MgSO₄ (≥99.99% pure, Sigma-Aldrich) and vortexed vigorously to remove water. After 10 min centrifugation at 1000 g, 80 µL of the final supernatant was spiked with instrumental internal standards (final concentration: 1 ng/mL).
for analysis. Other materials were processed similarly, i.e., 100 mg lung was homogenized in 300 µL water and extracted with 150 µL formic acid and 400 µL hexane-ethyl acetate mixture, while 40 mg thyroid was homogenized in 250 µL water and extracted with 50 µL formic acid and 200 µL solvent. Stool samples (100 mg) were homogenized and extracted directly in 50 µL formic acid and 200 µL solvent and then processed as plasma samples. Samples were prepared in batches containing 20 samples along with two SRM samples, two iso-octane, two solvent blank containing internal standards and two saline samples that went through solvent extraction, as part of quality control measures.

*Instrumental analysis.* Samples were analyzed with three injections using GC-HRMS with a Thermo Scientific Q Exactive GC hybrid quadrupole Orbitrap mass spectrometer with 2 µL per injection. A capillary DB-5MS column (15 m × 0.25 mm × 0.25 µm film thickness) was used with the following temperature program: hold 75 °C for 1 min, 25 °C/min to 180 °C, 6 °C/min to 250 °C, 20 °C/min to 350 °C and hold for 5 min. The flow rate of the helium carrier gas was 1 mL/min. Ion source and transfer line temperatures were 250°C and 280°C, respectively. Data were collected from 3 to 24.37 min with positive electron ionization (EI) mode (+70 eV), scanning from \( m/z \) 85.0000 to 850.0000 with a resolution of 60,000.

*Data extraction and pre-processing.* Raw data were examined by checking signal-to-noise ratio, peak shape and spectral information for surrogate and internal standards using a 5 ppm \( m/z \) tolerance and 30 s retention time window in xCalibur Qualbrowser software. TraceFinder software version 4.1 (Thermo Fisher Scientific) was tested with mixtures of standards and found challenging to simultaneously detect >250 chemicals. Thus, data extraction was performed by XCMS (36) to generate about 40,000 chemical features identified by spectral \( m/z \) and retention time; extraction with apLCMS (37) generated more than 200,000 \( m/z \) features.
which were considered too many for current needs. Data were pre-filtered to retain around 25,000 features that had average peak intensities for non-blank samples that were 10-fold greater than saline method blanks.

For targeted quantitation, we used the library of 378 chemical standards consisting of the spectral information (the five most abundant \textit{m}/\textit{z}) and retention time (\textbf{Supplemental Table S3}). Features were selected with tolerance of ±5 ppm \textit{m}/\textit{z} and 30 s retention time, and further clustered by RAMclustR (26) based on feature similarity in retention time and correlation across samples. Features were matched to chemical spectra for identification, and intensities of the most abundant \textit{m}/\textit{z} fragments were used for quantification. Alternatively, as an untargeted approach, features were clustered with RAMclustR without matching to target chemicals to support biostatistics and bioinformatics analysis before chemical annotation and identification.

\textit{Metabolite quantification in NIST reference serum using external standard curves.}

Recoveries of each recovery standard was determined after normalizing to \[^{13}\text{C}_{12}\]PCB-28 and \[^{13}\text{C}_{12}\]PBB-153 added as volumetric internal standard in SRM-1958 samples (n≥3). For each authentic environmental chemical, a serial dilution of 0, 0.05, 0.1, 0.25, 0.5, 1 and 2 ng/mL in isoocetane (n=2 each) were run and analyzed through data extraction and pre-processing together with SRM-1958 samples (n=3). Instrument detection limits and linearity of response were calculated from the diluted standards (38) using the most abundant spectral \textit{m}/\textit{z} (\textbf{Supplemental Table S3}). Absolute quantification of chemicals in SRM-1958 was determined from external calibration curves using the most abundant \textit{m}/\textit{z} fragment in full scan mode.

\textit{Reference standardization.} Reference standardization is a simple quantification protocol in which individual chemical concentrations in unknown samples are estimated by single-point calibration to a concurrently analyzed, pooled reference sample with known chemical
concentrations. This method has been previously validated for LC-HRMS data. To validate the use in GC-HRMS, a randomly selected set of human plasma (n=20) were quantified by two methods: 1) Reference standardization using NIST SRM-1958 and NIST SRM-1957 that were run in parallel to unknown samples; for chemicals that were detected in both SRM-1957 and SRM-1958, the certified mass fractions from both references were used; for chemicals that were only detected in SRM-1958, the certified mass values from SRM-1958 was used. 2) External calibration curves constructed by known concentrations of standards (0, 0.05, 0.1, 0.25, 0.5, 1 and 2 ng/mL) and adjusted for recoveries measured by $^{13}$C labelled internal standards.

After reference standardization based on SRM-1957 and SRM-1958 was validated, quantification results of human plasma, tissue and stool samples were obtained using this approach (18). In each batch, SRM-1957 and SRM-1958 were extracted and analyzed in parallel to study samples to support batch-wise quantification.

Statistical analysis. Data are presented as the mean and standard error of the mean. Welch’s $t$ test was used to determine significant differences between two groups, and one way ANOVA was used among multiple groups. Mann-Whitney U test was used as a nonparametric test when normality assumption failed by Shapiro-Wilk test ($P < 0.05$). Pearson’s correlation analysis was performed with SigmaPlot 14.0 (Systat Software, Inc). All other bioinformatics analyses were performed in R Studio version 1.1.447 (RStudio, Inc). The significance level was $p < 0.05$ for all tests.
Acknowledgments

**Funding:** This study was supported by the NIEHS, U2C ES030163 (DPJ), U2C ES030859 (DIW) and P30 ES019776 (CJM), NIDDK RC2 DK118619 (KNL), NHLBI R01 HL086773 (DPJ), US Department of Defense W81XWH2010103 (DPJ), and the Chris M. Carlos and Catharine Nicole Jockisch Carlos Endowment Fund in Primary Sclerosing Cholangitis (PSC) (KNL).

**Author contributions:** XH, DIW, CJM, YMG, KP, GWM, KNL, DPJ discussed and designed the study. XH, DIW, YLL tested the analytical method and performed experiment. MRS, MLO, BDJ, MK GSM, DCN, KNL collected and provided human material. XH and DPJ analyzed and interpreted data. XH, DIW, CJM, KP, GWM, KNL, DPJ prepared and edited the manuscript text. All authors reviewed the manuscript.

**Competing interests:** The authors do not have any competing interests.

**Data and materials availability:** Raw data of this study is in the process of being uploaded to Metabolomics Workbench (www.metabolomicsworkbench.org).
Figure 1. New analytic approach is needed to support health research on extensive and diverse chemical exposures of humans. A. Humans are exposed to chemicals from diet, personal use products, drugs and dietary supplements, microbial metabolites, air pollution, occupational exposures and other environmental exposures, which are distributed throughout the body and sampled in different ways for biomonitoring. B. Targeted analyses for a limited
number of chemicals are available using mass spectrometry. Left: Multistep processing is used for targeted detection of low abundance of environmental chemicals. Right: Recovery of chemicals differs by sample preparation procedures. Standard reference material NIST SRM-1958 was analyzed using solid-phase extraction (SPE) and liquid-liquid extraction (LLE). Middle: Addition of stable isotopic internal standards during the procedure allows correction for variability of recovery due to losses during processing. The multistep processing has potential for contamination at each step, decreases sample throughput, increases cost and are not suitable for untargeted analysis because there is no way to correct for variable recovery.
Figure 2. Express liquid extraction (XLE) provides single-step extraction to minimize variation in chemical recovery for untargeted environmental biomonitoring. **A.** XLE minimizes the number of sample processing steps for use with GC-HRMS for measurement. **B.** Inclusion of method blanks showing similar signal level as directly-injected isooctane solvent supports quality control (QC) and standard reference material (NIST SRM-1957, 1958) in each
sample batch supports quality assurance (QA). Total ion intensity was calculated by the sum of intensity of all detected peaks (n=6). C. XLE supports recovery of stable isotopic standards calculated as percent recovered in XLE-processed SRM-1958. D. XLE is reproducible by independent operators at different analysis time. Selected chemicals that are commonly detected in human biomonitoring studies (22) are presented in raw intensity (average ± SE, n≥4) and illustrate comparability of results as well as long-term robustness. Left to right chemicals: \( p,p' \)-DDE, PCB-138, PCB-118, HCB, PBDE-47, PBDE-153, PBB-153. E. Quantification of selected chemicals in SRM-1958 (average ± SE, n=3) using external standard curves (0.05 to 2 ng/mL). Certified values of chemical mass fraction are presented by open circle (average ± SE (21)). Because unidentified environmental chemicals will not have recovery data, these values are expressed without correction for recovery to illustrate usefulness for untargeted analysis (e.g., compare height of bar to open circle). F. Chemical chromatography in XLE analyzed SRM-1958 shows detection of extensive number of environmental chemicals for quantification (peaks are color coded by primary m/z, with chemical details in Supplemental Figure S1).
Figure 3. Adaptation of Reference Standardization to GC-HRMS provides automated workflow for quantification of identified environmental chemicals. A. Reference standardization is a validated single-point quantification method used in LC-HRMS for automated measurement in high-throughput analyses. The method relies upon pooled reference materials that are calibrated relative to external calibration curves and relative to other pooled reference materials. These pooled reference materials are analyzed before and after each batch of samples so that concentrations of environmental chemicals detected in the samples can be quantified relative to concentrations in the pooled reference materials. B. We used dilution series of environmental chemical standards to validate quantification by reference standardization; mean concentrations of 20 plasma samples are presented with positive standard errors by two methods. C. Pearson correlation coefficients of concentrations across 20 plasma samples quantified by two methods.
Figure 4. XLE with GC-HRMS supports detection and quantification of environmental chemicals in different biologic materials. A. Human plasma from 80 individuals without known occupational or environmental exposures of concern were analyzed for 49 chemicals and visualized by unsupervised two-way hierarchical clustering of log-transformed intensities. Subjects were color-coded into three BMI groups and age quantiles (Q1 to Q4). Results show that high-throughput analysis of environmental chemicals enables study of relationships of chemical distributions and associations with health characteristics. B. XLE supports quantification of environmental chemicals in human plasma (n=80), lung (n=11), thyroid (n=5) and stool (n=6) samples. Hierarchical cluster analysis of Spearman’s correlation among quantified environmental chemicals shows co-exposure of different chemical classes and illustrates that XLE supports quantification of multiple classes of chemicals in different tissue types for integrative analyses of diverse chemical exposures. C. Concentrations and distribution of chemicals prevalent in plasma (n=80) and lung (n=11) samples shows that XLE is suitable to
support quantification in human organs and compare results to plasma (white bar – median; black dot – individual sample concentration).
Figure 5. XLE with GC-HRMS measures unidentified environmental chemicals to support exposome epidemiology. 

A. Data from untargeted analyses can be used directly for biostatistical and bioinformatics analyses of relationships to health markers without chemical identification. By defining the mass spectral signals relative to known index chemicals, accurate mass m/z signals, along with retention time and ion intensity, results on unidentified signals can be incorporated into exposome reference databases and used for subsequent investigation, such as database searches.

B. Application of tools such as RAMclustR (26) to untargeted data allow co-eluting m/z features to be studied as possible products derived from an unidentified chemical. In this example of an analysis of human plasma (n=60), unidentified signals of a two-minute retention time interval are clustered into spectra and color-coded based on clusters. Size of circles reflect raw intensities.

C. Clustered m/z spectra are likely to include unidentified environmental chemicals and can be used for discovery of unidentified chemical structures. Examples are presented showing putative molecule formula assigned to spectra by MS-Finder.
ver 3.42 (28). Candidates with the highest formula scores (C74: 4.1; C106: 2.4; C208: 4.3) in MS-Finder were selected.
Figure 6. Application of XLE with GC-HRMS and reference standardization provides a framework for high-throughput exposome research. The analytic workflow with single-step extraction and analysis along with pooled reference materials provides a simple, automatable method for measurement of low abundance chemicals in biological materials. Left: The framework is anchored to accurate mass m/z signals which are clustered according to retention time and intensity correlations. Middle: These can be aligned relative to index chemicals and quantified relative to standard reference materials, thereby providing key criteria for interoperability and reproducibility. Right: Both targeted and untargeted chemical data obtained by this workflow are suitable for entry into cumulative data library to support exposome research.
Supplemental Figure S1a. Detected chromatographic traces of PCBs in NIST1958 using XLE.
Supplemental Figure 1b. Detected chromatographic traces of chlorinated pesticides in NIST1958 using XLE.
Supplemental Figure 1c. Detected chromatographic traces of PBDE Congeners and PBB 153 in NIST1958 using XLE.
More supplemental material (Excel Spreadsheet) not included in this file:

**Supplemental Table S1.** Quantification of chemicals with reference values in reconstituted standard reference material of human serum, NIST SRM-1958.

**Supplemental Table S2.** Comparison of quantification results in n=20 human plasma samples, by reference standardization using SRM-1957 and/or SRM-1958 and by external standard curves (0-2 ng/mL).

**Supplemental Table S3.** Information of 378 chemical authentic standards on retention time, spectral information (the most abundant m/z: mz1 to mz5), linearity and limit of detection (LOD).

**Supplemental Table S4.** Raw intensity and quantification of 80 human plasma samples collected via Center of Health Discovery and Well Being (CHDWB).

**Supplemental Table S5.** Concentration (ng/g tissue weight) of chemicals measured by reference standardization in 11 human lungs.

**Supplemental Table S6.** Concentration (ng/g tissue weight) of chemicals measured by reference standardization in 5 human thyroids.

**Supplemental Table S7.** Concentration (ng/kg wet weight) of chemicals measured by reference standardization in 6 human stool samples.
References

1. United States Environmental Protection Agency (EPA). TSCA chemical substance inventory. [accessed July 29th, 2020] http://www.epa.gov/tscs-inventory/.

2. S. M. Rappaport, D. K. Barupal, D. Wishart, P. Vineis, A. Scalbert, The blood exposome and its role in discovering causes of disease. Environmental health perspectives 122, 769-774 (2014).

3. S. M. Rappaport, M. T. Smith, Environment and disease risks. Science 330, 460-461 (2010).

4. C. P. Wild. Complementing the Genome with an “Exposome”: The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. Cancer Epidemiol Biomarkers Prev. 14(8), 1847-1850 (2005).

5. G. W. Miller, D. P. Jones, The nature of nurture: refining the definition of the exposome. Toxicological Sciences 137, 1-2 (2013).

6. R. Vermeulen, E. L. Schymanski, A.-L. Barabási, G. W. Miller, The exposome and health: Where chemistry meets biology. Science 367, 392-396 (2020).

7. W. J. Crinnion, The CDC Fourth National Report on Human Exposure to Environmental Chemicals: What it Tells Us About our Toxic Burden and How it Assists Environmental Medicine Physicians. Alternative medicine review 15, (2010).

8. L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W. M. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden, M. R. Viant, Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 3, 211-221 (2007).
9. T. Matsuo, H. Tsugawa, H. Miyagawa, E. Fukusaki, Integrated Strategy for Unknown EI-MS Identification Using Quality Control Calibration Curve, Multivariate Analysis, EI-MS Spectral Database, and Retention Index Prediction. *Anal Chem* **89**, 6766-6773 (2017).

10. D. Stettin, R. X. Poulin, G. Pohnert, Metabolomics Benefits from Orbitrap GC-MS-Comparison of Low- and High-Resolution GC-MS. *Metabolites* **10**, (2020).

11. A. C. Peterson, J.-P. Hauschild, S. T. Quarmby, D. Krumwiede, O. Lange, R. A. S. Lemke, F. Grosse-Coosmann, S. Horning, T. J. Donohue, M. S. Westphall, J. J. Coon, J. Griep-Raming, Development of a GC/Quadrupole-Orbitrap Mass Spectrometer, Part I: Design and Characterization. *Analytical Chemistry* **86**, 10036-10043 (2014).

12. D. P. Jones, Sequencing the exposome: a call to action. *Toxicology reports* **3**, 29-45 (2016).

13. M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* **3**, (2016).

14. D. I. Walker, Y.-M. Go, K. Liu, K. D. Pennell, D. P. Jones, "Population screening for biological and environmental properties of the human metabolic phenotype: implications for personalized medicine" in *Metabolic Phenotyping in Personalized and Public Healthcare* (Elsevier, 2016), pp. 167-211.

15. K. Uppal, D. I. Walker, K. Liu, S. Li, Y.-M. Go, D. P. Jones, Computational Metabolomics: A Framework for the Million Metabolome. *Chemical research in toxicology* **29**, 1956-1975 (2016).
16. D. P. Jones, B. A. Cohn, A vision for exposome epidemiology: The pregnancy exposome in relation to breast cancer in the Child Health and Development Studies. *Reproductive toxicology* **92**, 4-10 (2020).

17. A. Sjödin, R. S. Jones, C. R. Lapeza, J. F. Focant, E. E. McGahee, 3rd, D. G. Patterson, Jr., Semiautomated high-throughput extraction and cleanup method for the measurement of polybrominated diphenyl ethers, polybrominated biphenyls, and polychlorinated biphenyls in human serum. *Anal Chem* **76**, 1921-1927 (2004).

18. J. M. Keller, R. F. Swarthout, B. K. R. Carlson, J. Yordy, A. Guichard, M. M. Schantz, J. R. Kucklick, Comparison of five extraction methods for measuring PCBs, PBDEs, organochlorine pesticides, and lipid content in serum. *Analytical and bioanalytical chemistry* **393**, 747-760 (2009).

19. Y. M. Go, D. I. Walker, Y. Liang, K. Uppal, Q. A. Soltow, V. Tran, F. Strobel, A. A. Quyyumi, T. R. Ziegler, K. D. Pennell, G. W. Miller, D. P. Jones, Reference Standardization for Mass Spectrometry and High-resolution Metabolomics Applications to Exposome Research. *Toxicological Sciences* **148**, 531-543 (2015).

20. A. Srivastava, S. Rai, A. Kumar Sonker, K. Karsauliya, C. P. Pandey, S. P. Singh, Simultaneous determination of multiclass pesticide residues in human plasma using a mini QuEChERS method. *Analytical and bioanalytical chemistry* **409**, 3757-3765 (2017).

21. M. C. Bruzzoniti, L. Checchini, R. M. De Carlo, S. Orlandini, L. Rivoira, M. Del Bubba, QuEChERS sample preparation for the determination of pesticides and other organic residues in environmental matrices: a critical review. *Analytical and bioanalytical chemistry* **406**, 4089-4116 (2014).
22. National Institute of Standards & Technology, Certificate of Analysis Standard Reference Material® 1958 [accessed April 1 2020] https://www-s.nist.gov/srmors/certificates/1958.pdf.

23. Centers for Disease Control and Prevention. Fourth report on human exposure to environmental chemicals, updated tables, (January 2019). Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Accessed April 9th 2019: https://www.cdc.gov/biomonitoring/biomonitoring_summaries.html.

24. H. A. Anderson, M. S. Wolff, Environmental contaminants in human milk. Journal of Exposure Science & Environmental Epidemiology 10, 755-760 (2000).

25. L. M. O. Sahlström, U. Sellström, C. A. de Wit, S. Lignell, P. O. Darnerud, Feasibility Study of Feces for Noninvasive Biomonitoring of Brominated Flame Retardants in Toddlers. Environmental Science & Technology 49, 606-615 (2015).

26. D. B. Barr, R. Y. Wang, L. L. Needham, Biologic monitoring of exposure to environmental chemicals throughout the life stages: requirements and issues for consideration for the National Children's Study. Environmental health perspectives 113, 1083-1091 (2005).

27. C. D. Broeckling, F. A. Afsar, S. Neumann, A. Ben-Hur, J. E. Prenni, RAMClust: a novel feature clustering method enables spectral-matching-based annotation for metabolomics data. Anal Chem 86, 6812-6817 (2014).

28. H. Tsugawa, T. Kind, R. Nakabayashi, D. Yukihiro, W. Tanaka, T. Cajka, K. Saito, O. Fiehn, M. Arita, Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. Analytical chemistry 88, 7946-7958 (2016).

29. Z. Lai, H. Tsugawa, G. Wohlgemuth, S. Mehta, M. Mueller, Y. Zheng, A. Ogiwara, J. Meissen, M. Showalter, K. Takeuchi, T. Kind, P. Beal, M. Arita, O. Fiehn, Identifying
metabolites by integrating metabolome databases with mass spectrometry cheminformatics.

*Nature methods* **15**, 53-56 (2018).

30. S. M. Rappaport, Genetic factors are not the major causes of chronic diseases. *PloS one* **11**, e0154387 (2016).

31. X. Hu, S. Safley, M.R. Smith, Y. Liang, V. Tran, J. Sharma, D. P. Jones, C. J. Weber, Environmental chemicals and metabolic disruption in primary and secondary human parathyroid tumors. *Surgery*, in press (2020).

32. D. Valvi, D. I. Walker, T. Inge, S. M. Bartell, T. Jenkins, M. Helmrath, T. R. Ziegler, M. A. La Merrill, S. P. Eckel, D. Conti, Y. Liang, D. P. Jones, R. McConnell, L. Chatzi, Environmental chemical burden in metabolic tissues and systemic biological pathways in adolescent bariatric surgery patients: A pilot untargeted metabolomic approach. *Environ Int* **143**, 105957 (2020).

33. M. La Merrill, C. Emond, M. J. Kim, J.-P. Antignac, B. Le Bizec, K. Clément, L. S. Birnbaum, R. Barouki, Toxicological function of adipose tissue: focus on persistent organic pollutants. *Environmental health perspectives* **121**, 162-169 (2013).

34. M. Rönn, L. Lind, B. van Bavel, S. Salihovic, K. Michaëlsson, P. M. Lind, Circulating levels of persistent organic pollutants associate in divergent ways to fat mass measured by DXA in humans. *Chemosphere* **85**, 335-343 (2011).

35. C. Rose, A. Parker, B. Jefferson, E. Cartmell, The Characterization of Feces and Urine: A Review of the Literature to Inform Advanced Treatment Technology. *Crit Rev Environ Sci Technol* **45**, 1827-1879 (2015).
36. C. A. Smith, E. J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* **78**, 779-787 (2006).

37. T. Yu, Y. Park, J. M. Johnson, D. P. Jones, apLCMS—adaptive processing of high-resolution LC/MS data. *Bioinformatics* **25**, 1930-1936 (2009).

38. P. R. Loconto, *Trace environmental quantitative analysis: principles, techniques and applications* (2nd edition): Boca Raton, FL, Taylor & Francis (2006).