Environmental Science & Technology

Article

pubs.acs.org/est

# Chemical Isotope Labeling Exposome (CIL-EXPOSOME): One High-Throughput Platform for Human Urinary Global Exposome Characterization

Shenglan Jia,[†,‡,§] Tengfei Xu,[†,§] Tao Huan,[∥] Maria Chong,[†] Min Liu,[†,‡] Wenjuan Fang,[‡] and Mingliang Fang*[,†,§,⊥]

[†]School of Civil and Environmental Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798

[‡]Residues and Resource Reclamation Centre, Nanyang Environment and Water Research Institute, Nanyang Technological University, 1 Cleantech Loop, CleanTech One, Singapore 637141

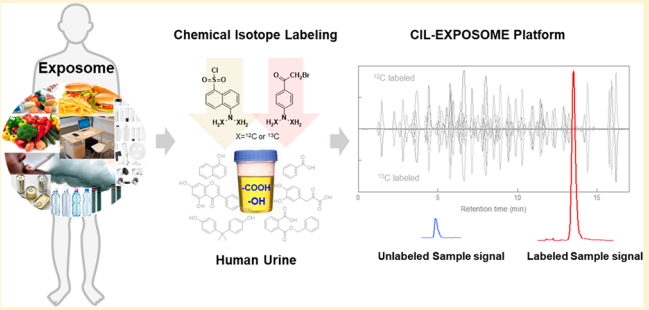
[§]Advanced Environmental Biotechnology Centre, Nanyang Environment & Water Research Institute, Nanyang Technological University, Singapore 637141

[∥]Department of Chemistry, University of British Columbia, Vancouver Campus, 2036 Main Mall, Vancouver, BC, Canada V6T 1Z1

[⊥]Analytics Cluster, Nanyang Environment and Water Research Institute, Nanyang Technological University, 1 Cleantech Loop, CleanTech One, Singapore 637141

Supporting Information

**ABSTRACT:** Human exposure to hundreds of chemicals, a primary component of the exposome, has been associated with many diseases. Urinary biomarkers of these chemicals are commonly monitored to quantify their exposure. However, because of low concentrations and the great variability in physicochemical properties of exposure biomarkers, exposome research has been limited by low-throughput and costly methods. Here, we developed a sensitive and high-throughput exposome analytical platform (CIL-EXPOSOME) by isotopically labeling urinary biomarkers with common functional groups (hydroxyl/carboxyl/primary amine). After a simple cleanup, we used mass spectrometry to perform a screening for both targeted and untargeted biomarkers, which was further processed by an automatic computational pipeline method for qualification and quantification. This platform has effectively introduced an isotope tag for the absolute quantification of biomarkers and has improved sensitivity of 2−1184 fold compared to existing methods. For putative identification, we built a database of 818 urinary biomarkers with MS/MS fragmentation information from either standards or in silico predictions. Using this platform, we have found 671 urinary exposure biomarker candidates from a 2 mL pooled urine sample. The exposome data acquisition and analysis time has also been greatly shortened. The results showed that CIL-EXPOSOME is a useful tool for global exposome analysis of complex samples.




## INTRODUCTION

The concept of the "exposome", which is a measure of the effects of life-long environmental exposure on health, has been a hot topic since its introduction in 2005.[1] Recent initiatives in environmental exposome research have attempted to map the complex relationships between biomarkers and disease risks.[2−4] Therefore, measurable and characteristic changes of all biomarkers taken together form the key foundation of exposome research.[5] Exposure biomarkers possess enormous structural diversity and have a broad concentration range (from $10^{-7}$ $\mu$M to 1 mM for environmental pollutants).[6,7] Meanwhile, we are constantly exposed to an array of ever-increasing chemicals, both endogenous and exogenous, from our diet, the environment, and our behavior.[8] As such, profiling a large number of exposure biomarkers requires powerful analytical techniques with broad coverage and high sensitivity.

Current exposure analytical methods include both traditional targeted measurement and untargeted discovery of chemicals with biological importance (mainly endogenous metabolites).[9−11] For targeted appraoches, high sensitivity, wide dynamic range, and good reproducibility are achieved with complex cleanup steps and expensive chemical isotope-labeled (CIL) analogues.[12] Moreover, the targeted method can measure only several biomarkers at one time and requires large

Received: January 14, 2019
Revised: March 27, 2019
Accepted: April 3, 2019
Published: April 3, 2019





DOI: 10.1021/acs.est.9b00285
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

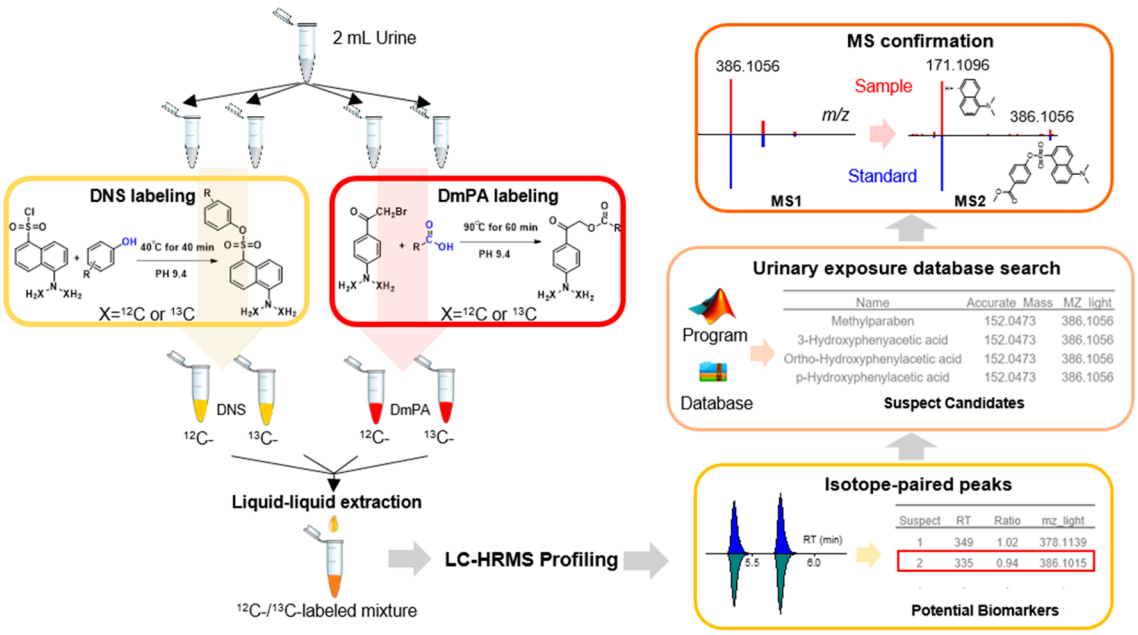



**Figure 1.** Workflow for the determination and quantitation of hydroxyl/primary amine and carboxyl exposure biomarkers in human urine by the CIL-EXPOSOME platform. Following derivatization and sample processing, the LC-qTOF raw data were imported into the IsoMS and suspect exposure biomarkers were searched against the compounds in the CIL-EXPOSOME database under different mass accuracy tolerances. The resulting information (name, mass tolerance, accurate mass, and $m/z$ light) was exported. The exposure biomarker candidates were further confirmed with MS/MS (i.e., $MS^2$) fragmentation.

volumes of sample. Untargeted approaches have primarily evolved from global metabolomics methods, by combining advanced analytical and computational tools.[13] Untargeted metabolomics retain a broad coverage of analytes by sacrificing sensitivity, making the approach inefficient at measuring low-concentration environmental chemicals.[5,14] Therefore, there is a great demand for a sensitive and high-throughput platform, capable of capturing both endogenous and low-level exogenous exposure biomarkers.

Among all biological samples, urine samples have a great potential to be used in long-term exposome characterization because of their noninvasive nature and easy collection.[15] Most environmental chemical exposure biomarkers are excreted and can be found in the urine following oxidation, reduction, or hydrolysis and further phase II conjugations.[16] However, high concentrations of salts and the diversity of chemical compounds present make the analysis of urinary exposure biomarkers across a wide range of concentrations challenging.

To address this analytical challenge and establish a sensitive, high-throughput exposome platform, we applied a chemical-derivatization method for exposome profiling in urinary samples. Chemical isotope labeling combined with liquid chromatography−mass spectrometry (CIL-LC-MS) analysis was used to enhance the detection of both endogenous and exogenous compounds, simplifying both data acquisition and processing during exposome analysis. CIL-LC-MS could significantly improve chemical detection and quantification of endogenous metabolites, with sensitivity increases of 10−1000-fold observed in previous studies.[17−19] Furthermore, isotope forms of the derivatization reagents can be used as an economical internal standard for absolute quantification.[20,21] After the metabolites were categorized into several subgroups based on their functional moieties (i.e., carboxyl, carbonyl, amine, etc.), different labeling chemistries can be used to profile different submetabolomes in order to provide better coverage of the entire compound range of interest.[22,23] For most exogenous biomarkers in the urinary exposome, either the parent compound or metabolites commonly contain hydroxyl or carboxylic acid functional groups to enhance their aqueous solubility, making the CIL method promising for this application.[24]

Here, we present a urinary exposome profiling workflow that employs CIL and automatic data processing to examine hydroxyl/primary amine and carboxyl biomarkers originating from both endogenous and exogenous chemicals. This approach provides a solution to several challenges currently limiting exposure biomarker profiling. First, CIL methods were used to increase sensitivity and greatly improve the coverage of exposure biomarkers. Second, an automatic data processing approach was built up to confirm potential exposure biomarkers, enabling fast screening and compound identification. Finally, the method's performance was validated in a pooled human urine sample, demonstrating that the CIL-EXPOSOME workflow is a promising tool for future global exposome analysis.

## MATERIALS AND METHODS

**Chemicals and Stock Preparation.** All chemicals and reagents were purchased from Sigma-Aldrich Singapore (Singapore) except those otherwise noted. The labeling reagent, *p*-dimethylaminophenacyl bromide (DmPA), was purchased from Apollo Scientific Ltd. (Stockport, U.K.). The isotope reagents, $^{13}C_2$-danysl chloride (DNS) and $^{13}C_2$-*p*-dimethylaminophenacyl bromide, were purchased from TMIC (Edmonton, Canada). $^{13}C_{12}$-BPA, $^{13}C_6$-monobenzyl phthalate, and $^{13}C_4$-methyl paraben were purchased from Cambridge Isotope Laboratories (Andover, MA, United States). HPLC grade water, ethyl acetate, and acetonitrile (ACN) were purchased from Thermo Fisher Scientific (Singapore).


DOI: 10.1021/acs.est.9b00285
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

*Standard Solution Preparation.* Twenty-two hydroxyl- or carboxyl- containing standards were chosen to optimize the analytical method. Additional information is provided in Table S1. Each compound dissolved in ACN at a concentration of 10 mg/mL. Seven-level mixed standard working solutions (10, 1, 0.1, 0.01, 0.001, 0.0001, 0.00001 μg/mL), also serving as calibration standards, were prepared with ACN through serial dilution of stock solutions for method validation. In addition, three-level quality control (QC) working solutions were prepared in a similar manner to provide low (0.0001 μg/mL), medium (0.01 μg/mL), and high (1 μg/mL) level controls. All stock and working solutions were stocked at −40 °C and diluted with ACN to the desired level before use.

**Urine Sample Collection and Preparation.** A pooled human urine sample was prepared by mixing first morning voids of 35 volunteers (healthy, ages 22−37) which were immediately frozen at −20 °C after sampling. Informed written consent was obtained from all participants. Institutional Review Board approval for human specimen analysis were obtained from Singapore (IRB-2017-02-023). The pooled urine sample was buffered to pH 5.5 and treated with an enzyme mixture (β-glucuronidase/sulfatase) for 12 h at 37 °C. After centrifugation (10 min, 16000*g*, 4 °C), the urine was processed with labeling reagents for method development. Spiking experiments were performed by adding a mixed working solution of the selected compounds or surrogate standards to the pooled samples before sample preparation.

**CIL-EXPOSOME Procedure.** *Work Flow.* In this study, two CIL methods were used to profile hydroxyl/primary amine/carboxyl endogenous and exogenous biomarkers. Figure 1 illustrates the overall workflow of the method and data processing design with the following steps: (1) dansylation or *p*-dimethylaminophenacyl bromide (DmPA) labeling of the deconjugated urine sample; (2) equal amounts of $^{12}$C-/$^{13}$C-dansylation/DmPA were mixed with labeled samples; (3) liquid−liquid extraction was performed to enrich the labeled biomarkers from above complex mixtures; (4) raw LC-qTOF analysis data were imported into IsoMS (Test S1) to get $^{12}$C-/$^{13}$C- peak pairs; (5) the exported paired suspect biomarkers were mass matched to the database under mass accuracy tolerance (e.g., < 20 ppm); (6) the resulting information (biomarker candidate name, mass tolerance, and $m/z$) were exported; and (7) the exported biomarker candidates were processed and putatively identified with standard or in silico predicted MS/MS fragmentation pattern.

*Reaction Method Optimization.* To optimize the derivatization conditions, three-level quality control (QC) mixed standard solutions reflecting human relevant levels were utilized. The mixtures were allowed to react under different concentrations of derivatization reagent, preparation times of derivatization reagent, ACN/$H_2O$ percentages, reaction temperatures, and incubation times individually. A detailed description of the labeling reaction can be found in Text S2.

*LC-qTOF Profiling and MS/MS Fragmentation Acquisition.* Standard and urine analyses were performed using a high-performance liquid chromatography (HPLC) system (1200 series, Agilent Technologies) coupled to an Agilent 6550 quadrupole time-of-flight (qTOF) mass spectrometer (Agilent, Singapore). Samples were injected for reversed-phase analysis in electrospray ionization (ESI) positive and negative modes, as detailed in Text S3. The instrument was set to acquire over a mass-to-charge ratio ($m/z$) range from 50 to 1700 with the MS acquisition rate of 2 Hz. For the acquisition of MS/MS spectra of selected precursors, the default isolation width was set to 1.3 Da with MS and MS/MS acquisition rates of 4 Hz. The collision energy was set to 20−40 eV. Data were processed by using the CIL-EXPOSOME data processing approach as described in Results and Disucssion.

**CIL-EXPOSOME Method Validation.** *Linearity and Sensitivity.* Calibration was conducted based on the analysis of derivatized 7-level labeled mix standard working solutions. Calibration curves were established by using linear regression. The LOD and LOQ were estimated as 3 and 10 times of the standard deviation calculated from 6 replicates at the lowest spiking level, respectively.

*Precision and Accuracy.* Three-level derivatization standard quality control (DQC) working solutions were prepared to provide low, medium, and high scenarios of human biomarker concentration. Intra- and interday precision and accuracy were assessed by using relative standard deviation (RSD) within a day ($n$ = 6) and on three consecutive days ($n$ = 9), respectively.

*Stability and Cleanup Efficiency.* Sample preparation stability of all derivatives was evaluated by analysis of the DQC solutions at 48 h (28 °C) and −40 °C at 30-day intervals. Pooled urine samples spiked with DQC solutions were prepared to evaluate the cleanup efficiency, which was defined as the response ratio of extracted concentrations from a spiked sample to the same concentrations of DQC. Each measurement was performed in triplicate.

*Matrix Effect.* In the absence of a biomarker-free urine matrix, a pooled urine sample was processed as described in Urine Sample Collection and Preparation and divided into two parts (A and B). Part A was dried to prepare the urine matrix, and an equal volume of DQC solutions were added to estimate the influence of the matrix effect. Part B was analyzed directly to determine the background concentration of each analyte. The matrix factor of each compound was calculated by difference between part A and DQC relative to the response of DQC in pure solvent. To further distinguish the matrix effect from derivatization efficiency, $^{13}C_6$-monobenzyl phthalate and $^{13}C_4$-methyl paraben were added before derivatization to examine the derivatization efficiencies on the two types of labeling. Labeled $^{13}C_{12}$-BPA was dansylated and added before injection as a recovery standard to assess the matrix effect.

*Statistical Analyses and Quality Assurance/Quality Control (QA/QC).* One-way ANOVA and Tukey's posthoc test were used to compare the response difference during optimization. Procedure MiliQ-water blanks were prepared to evaluate contamination arising from laboratory materials and solvents. An ACN blank was injected after 6 samples, and a calibration check standard was injected after every 12 samples.

## RESULTS AND DISCUSSION

**CIL-EXPOSOME Chemical Derivatization Method Development.** *Derivatization Reagent Selection.* One of the important considerations in developing labeling chemistry is how to tag the urinary biomarkers with differential isotopic groups in a robust way. In this study, we evaluated the functional groups of urinary biomarkers from a recent published comprehensive database (Exposome Explorer).[25] Together, biomarkers with hydroxyl, primary amine, and carboxylic functional groups consist of >70% of the total database. Hydroxyl compounds and primary amines are preferentially labeled with dansylation methods.[18,26,27] The derivatization of carboxylic acids can be conducted with a


DOI: 10.1021/acs.est.9b00285
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

variety of chemical reactions for analytical applications, and phenacyl bromide has been used to label the acids for improved HPLC and UV detection.[28] However, a new reagent, DmPA, was designed that allows for the introduction of a mass tag and concurrent improvement during LC-MS analysis.[19] Here, we used dansyl chloride (DnsCl) and DmPA to label hydroxyl/primary amine and carboxyl urinary biomarkers, respectively.

*Reaction Condition Optimization.* Although the stable isotope labeling method has been previously optimized for the endogenous metabolites,[29−31] the derivatization efficiency of high-concentration endogenous ($\mu$M to mM level) and low-concentration xenobiotic biomarkers (nM or pM level) may vary from each other. Furthermore, compared to endogenous metabolites, the chemical structure of xenobiotics and their biotransformation products are more diverse. To optimize the labeling conditions for urinary biomarkers, MiliQ water and urine spiked with three-level QC working solutions were labeled under different reaction conditions. Figure 2a,b shows

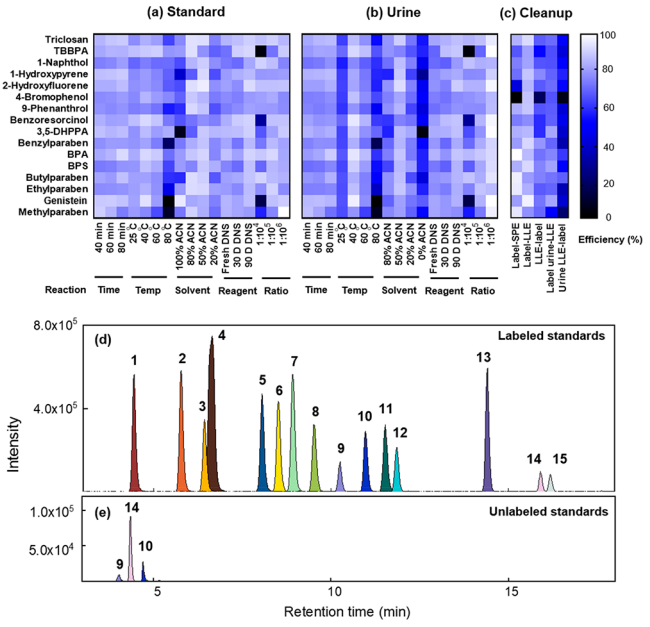



**Figure 2.** Comparison of labeling efficiencies for 15 hydroxyl standards (a) spiked in MiliQ water and (b) spiked in urine under different labeling reaction conditions. From left to right: reaction time, reaction temperature, incubation solvent, DnsCl preparation time, and compound:DnsCl mole ratio. (c) Comparison of cleanup method optimization for standards under different cleanup procedures. All data in panels a−c are presented as the mean from three separate experiments ($n$ = 3) of three levels of QC working solution (0.1, 10, and 1000 ppb). Extracted ion chromatograms of labeled standard mixture (d) after derivatization and (e) before derivatization (10 ppb each). The compounds are shown as follows: 1, methylparaben; 2, ethylparaben; 3, 4-bromophenol; 4, 1-naphthol; 5, benzylparaben; 6, 2-hydroxyfluorene; 7, butylparaben; 8, 9-phenanthrol; 9, bisphenol S; 10, benzoresorcinol; 11, 1-hydroxypyrene; 12, triclosan; 13, bisphenol A; 14, genistein; 15, 3,5-DHPPA.

the effects of reaction solvent, time, temperature, DnsCl/compound mole ratio, and DnsCl preparation time (from left to right) on the efficiency of standard dansylation. For the pure chemical optimization, results indicate that the effect of water in the incubation solvent on the labeling efficiency is compound-dependent. After the optimization of all the tested compounds, ACN/$H_2O$ (1:1, v/v) was selected as the reaction solvent for dansylation. The reaction temperature and time also had an effect on the labeling efficiency. We found that the optimal conditions were 40 min at 40 °C, which differs from the previous method for endogenous metabolites.[18] The optimal mole ratio of compound to DnsCl was 1:$10^5$. The optimization results for urine were the same as those for the MiliQ water, except that the mole ratio of compound to DnsCl increased to 1:$10^6$.

Under both MiliQ water and urine background, 80% ACN/$H_2O$ was selected as the reaction solvent for DmPA reaction, and the optimized reaction temperature and time were found to be 90 °C for 60 min (Figure S1a,b). The mole ratio of compound to DmPA was optimal at 1:$10^4$ for MiliQ water and 1:$10^5$ for urine.

*Cleanup Method Optimization.* A proper cleanup method is crucial to minimizing matrix effects. In the present study, we compared the recoveries of a few pretreatment methods: (1) extraction of the free biomarkers from MiliQ water followed by derivatization (LLE-Label); (2) biomarker derivatization in MiliQ water followed by extraction (Label-LLE, Label-SPE); (3) extraction of free biomarkers from urine followed by derivatization (urine LLE-Label); and (4) biomarker derivatization in urine followed by extraction (urine Label-LLE). As shown in Figure 2c, derivatized standards were more efficiently extracted by the liquid−liquid extraction (LLE) strategy than free standards in both MiliQ water and urine. This is probably because the derivatization step increases the hydrophobicity of hydroxyl, amines, and carboxylic acids in the urine sample. As a result, the chemicals are more likely to partition to the organic phase.[17] Sample processing was also streamlined by combining DnsCl- and DmPA-derivatized samples prior to extraction. Finally, the optimized cleanup in this study was simplified with a single liquid−liquid solvent extraction, which was enabled by the use of the nonpolar derivatization reagent. Details are as follows: (1) Equal amounts $^{12}C$-/$^{13}C_2$-DnsCl/DmPA labeled urine samples were combined to profile hydroxyl/primary amine and carboxyl biomarkers. (2) Saturated NaCl solution (v/v, 1:5) was added to the mixture then extracted with ethyl acetate twice (v/v, 1:3). (3) The extracted solutions were combined and dried, followed by reconstitution in ACN prior to LC-qTOF injection.

**Method Performance Validation.** *Sensitivity and Separation Improvement.* After derivatization, both the sensitivity and retention of all the tested compounds were greatly improved compared with their nonderivatized counterparts (Figure 2d,e). A similar finding was observed for the urine samples (Figure S2). The results showed that the detection sensitivities of these standard compounds generally increased by 2−1184 fold upon chemical labeling (Table S2). Before derivatization, only three of the free form hydroxyl compounds (bisphenol S, benzoresorcinol, and genistein) were detected at the concentration of 10 ppb. Chemical labeling not only will efficiently improve the detection sensitivities but also will lead to the discovery of more low-abundance biomarkers. The chromatographic separation also showed significant improvement after derivatization. Because of the diverse chemical structures and physicochemical properties among biomarkers, separation of these compounds in a complex urine sample is not readily accomplished with a single LC method. Derivatization can overcome this difficulty by introducing a hydrophobic functional group to the analytes. Following derivatization, sufficient separation can be accomplished by


DOI: 10.1021/acs.est.9b00285
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

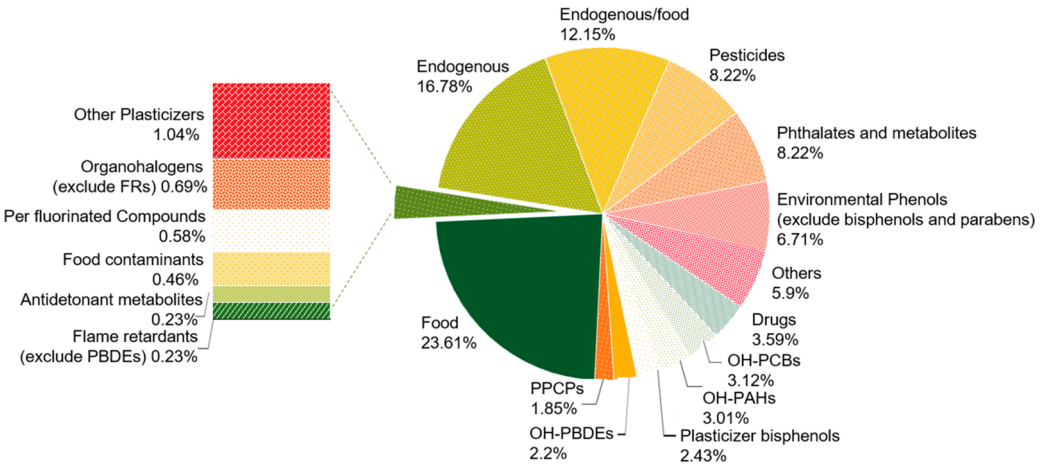



**Figure 3.** Composition of CIL-EXPOSOME biomarker database: number of compounds ($n$ = 818) by chemical classes.

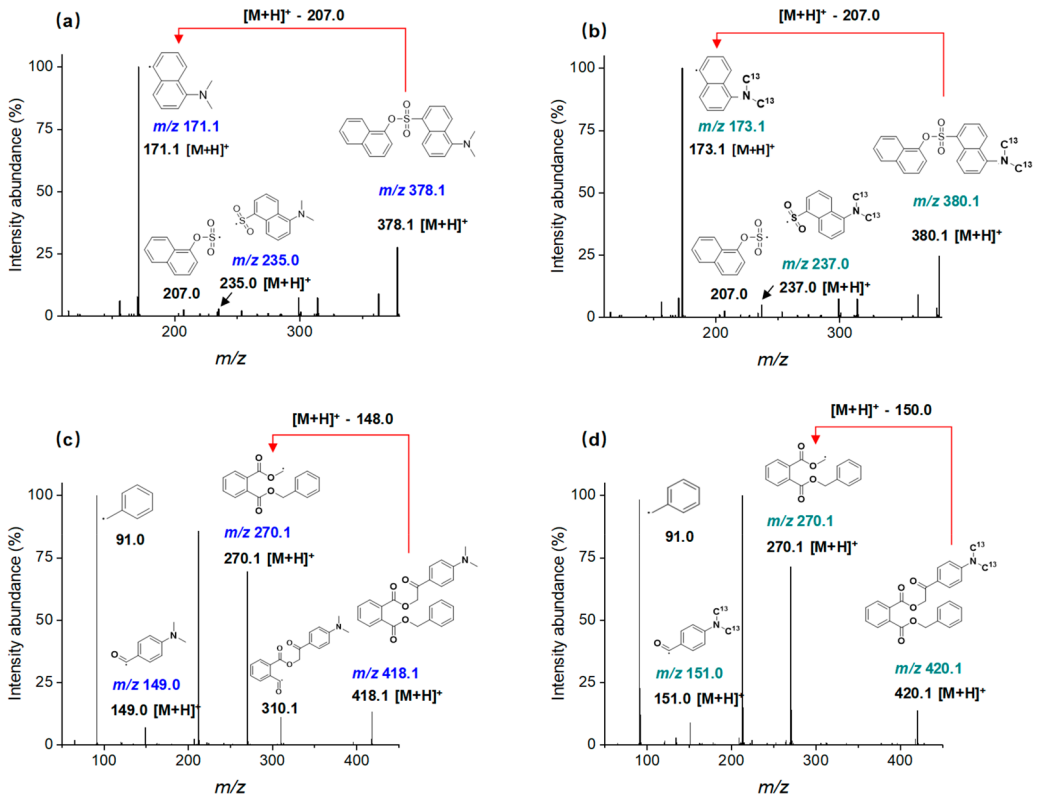



**Figure 4.** Fragmentation pattern of chemically labeled products. (a) DNS-labeled hydroxyl compound; (b) $^{13}C_2$-DNS-labeled hydroxyl compound; (c) DmPA-labeled carboxylic acid; (d) $^{13}C_2$-DmPA-labeled carboxylic acid. Highlighted in blue and green are the theoretical and labeled $m/z$, respectively.

using simple reverse phase chromatography. By chromatographic optimization, the derivatization methods greatly reduced the analysis time in the traditional targeted analysis, especially for polar biomarkers that are not retained by regular reverse-phase chromatography.

*Analytical Figures of Merit.* We estimated a series of attributes of the proposed quantification method, such as calibration range, linearity, sensitivity, precision, accuracy, stability, extraction efficiency, and matrix effect. The linearities of calibration curves for all analytes were excellent among the validated concentration range of 100 ppt to 1 ppm, with determination coefficients ($R^2$) greater than 0.99. Because of the diversity of biomarker concentrations, we considered two injection volumes (1 and 10 $\mu$L) to avoid response saturation. As shown in Table S2, all analytes had an LOQ lower than 1 ppb. The analytical method was shown to be highly reproducible, with all of the relative standard deviations (RSDs) less than 20% for both intraday and interday variations (Table S3). The stability of all derivatized standards had coefficients of variation (CVs) that were ≤18% and ≤16% for storage at 28 °C for 48 h or at −20 °C for 30 days, respectively (Table S4). Extraction efficiencies of three-level DQC using the above optimized cleanup method were approximately 73−92% (Table S5). The interference of the matrix with the analytes was within 40% (Table S5).


DOI: 10.1021/acs.est.9b00285
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

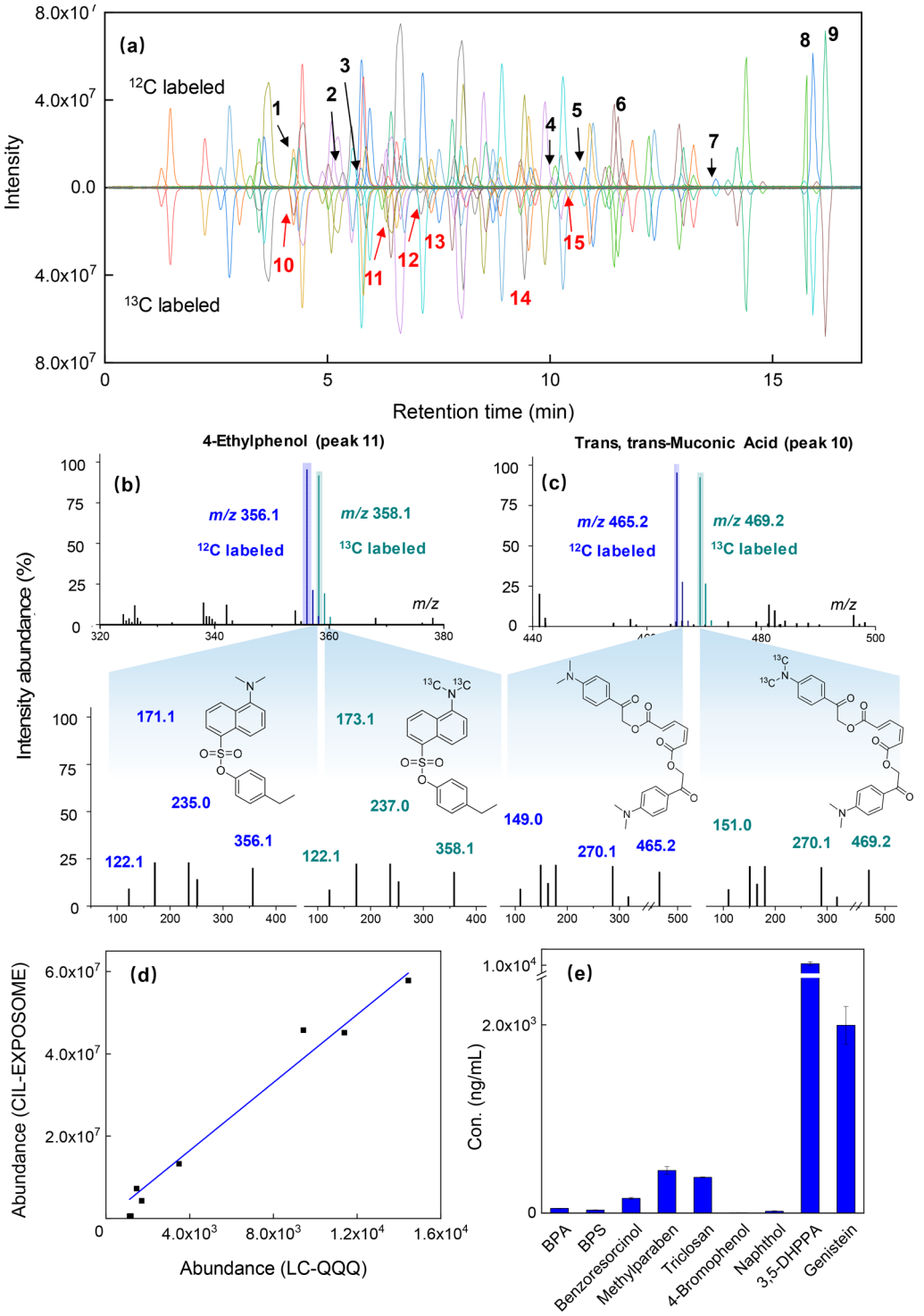



**Figure 5.** Chemical isotope-labeled urine sample analyzed by the CIL-EXPOSOME platform. Extracted ion chromatograms of (a) exposure biomarker examples from ~671 potential hydroxyl/primary amine and carboxylic acid metabolites from combined DNS/$^{13}C_2$-DNS and DmPA/$^{13}C_2$-DmPA-labeled urine sample. Fifteen exposure biomarkers have been confirmed with authentic standards, namely, 1, methylparaben; 2, 4-bromophenol; 3, naphthol; 4, BPS; 5, benzoresorcinol; 6, triclosan; 7, BPA; 8, genistein; 9, 3,5-DHPPA; 10, *trans,trans*-muconic acid; 11, 4-ethylphenol; 12, 2,4-di-*tert*-butylphenol; 13, 2-methylphenol; 14, 2-isopropoxyphenol; and 15, 4-propoxyphenol. An example of MS/MS from the newly identified (b) hydroxyl compound 4-ethylphenol and (c) carboxylic acid *trans,trans*-muconic acid. (d) Correlation of BPA abundance between LC-QQQ and the CIL-EXPOSOME platform ($R^2$ = 0.98, $p$ < 0.001) from 9 urine samples. (e) Absolute quantification of 9 targeted compounds detected in the pooled mixture of urine samples ($n$ = 35).

*CIL-EXPOSOME Biomarker Database.* To further facilitate the identification of urinary biomarkers, we established a urinary CIL-EXPOSOME biomarker database. Each entry in the exposure database included the chemical name, accurate molecular weight, exposure source, chemical class, predicted exact $m/z$ after being labeled, MS/MS fragmentation, and a detailed information link (HMDB, KEGG, or PUBCHEM). To create a broad exposome database, the CIL-EXPOSOME


DOI: 10.1021/acs.est.9b00285
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

biomarker database was compiled from a few existing data sources: Exposome-Explorer database,[25] Human Metabolome Database (HMDB),[32] the U.S. National Health and Nutrition Examination Survey (NHANES),[33] and MyCompoundID database.[34] The MS/MS fragmentation of urinary biomarkers was either generated using authentic standards or predicted using CFM-ID (see below). As a result, the CIL-EXPOSOME biomarker database that was established (Table S6) included 818 reported hydroxyl/primary amine and carboxyl biomarkers from food, drugs, endogenous metabolites, plasticizers, personal care products (PPCPs), polybrominated diphenyl ethers (PBDEs), polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), phthalates, pesticides, and flame retardants (excluding PBDEs) (Figure 3).

**CIL-EXPOSOME MS/MS Fragmentation Pattern and Database Development.** MS/MS fragmentation provides an important criteria for compound identification. In the present study, we created a MS/MS fragmentation library of ∼100 xenobiotics and their biotransformation products by conducting dissociation at different collision energies (0, 10, 20, and 30 eV). Together with ∼250 MS/MS spectra of endogenous metabolites from mycompound ID,[35] we have generated a MS/MS database composed of ∼350 authentic standards. For the rest of the compounds in the database, in silico MS/MS information was predicted using a machine-learning method (CFM-ID).[36] The database is available to the public via Google Drive (https://drive.google.com/drive/folders/1i1UNhfwMh_ry97TH6-FKGEm_A-i-oIeU?usp=sharing).

We further investigated the MS/MS fragmentation behavior of compounds labeled by two reagents (DnsCl and DmPA). As shown in Figure 4, the expected precursor ions were $m/z$ 378.1/380.1 for DNS/$^{13}C_2$-DNS-labeled naphthol and $m/z$ 418.1/420.1 for DmPA/$^{13}C_2$-DmPA-labeled monobenzyl phthalate. The DNS/$^{13}C_2$-DNS-labeled hydroxyl compounds generated two characteristic product ions, 171.1/173.1 and 235.0/237.0, by neutral loss from the precursor ion, $[M + H]^+$. For naphthol, another characteristic product ion was formed, $[M + H]^+$ 207.0, which is the parent ion of $[M + H]^+$ 171.1/173.1 (Figure 4a,b). The DmPA/$^{13}C_2$-DmPA-labeled carboxylic acids typically lose the common derivatization group ($m/z$ 149.0/151.0). The resulting product ion $[M + H]^+$ 270.1 for monobenzyl phthalate was formed (Figure 4c,d). Overall, we have observed $m/z$ 171.1 and 235.0 as two typical fragments for dansylation and $m/z$ 149.0 as a typical fragment for the DmPA label. The molecular MS fragment feature of the targeted compounds can also be observed in other MS/MS fragmentation patterns. The use of the selected paired peaks after chemical labeling as the targeted precursor ions facilitates subsequent MS/MS analysis. In addition, the characteristic fragmentation of metabolites after chemical labeling make identification feasible. This strategy can narrow down the candidate compounds from fragmentation ions.

**CIL-EXPOSOME Automatic Profiling Workflow.** In addition to working as an internal standard, $^{13}C$ reagent-labeled biomarkers can play a unique role in untargeted profiling and identification. Here, the obtained full scan (MS) raw data was processed by IsoMS to extract $^{12}C/^{13}C$ peak pairs from mass spectra. The paired peaks with a defined mass difference (2.0067 Da for one light/heavy labeled metabolites), tolerance window (e.g., <20 ppm), and intensity ratios (0.80−1.20) were extracted from the spectra. Information on detected compounds, including accurate $m/z$, retention time ($t$), and peak intensity, was generated for each exposure biomarker candidate (Table S7).

We further designed a MATLAB program to automatically search the exposure biomarker candidate peak pairs against the CIL-EXPOSOME biomarker database; the detailed manual can be found in Google Drive. On the basis of the IsoMS preselected mass list, the prospective candidates of each exposure biomarker were output with mass tolerance, name, database ID, accurate mass, and labeled accurate mass (Figure S3). Users can select different values to use for mass tolerance. For the identification, targeted MS/MS fragmentation of the potential biomarker was generated with a qTOF mass spectrometer. qTOF parameters included width of 1.3 $m/z$ units and a collision energy of 20 V. The obtained MS/MS fragmentation spectra were then further matched with prospective candidates in the established MS/MS library. Additionally, compounds were excluded that did not meet the following specifications: (1) targeted MS/MS of the potential biomarker included 171.1/173.1 or 149.0/151.0 in the native and isotope-labeled compounds and/or (2) the MS/MS fragment matched with the precursor compounds.

**CIL-EXPOSOME Platform Application.** With the established method, we determined the chemical and exposure biomarkers in the pooled urine sample. Using this integrated CIL-EXPOSOME automatic profiling method, 671 potential biomarkers (505 hydroxyl and 166 carboxylic acid biomarkers) were found, as partially illustrated in Figure 5a. For example, peaks 10 and 11 in Figure 5a are representative of extracted ion chromatograms at $m/z$ 356.1345/358.1397 and $m/z$ 465.2120/469.2265, indicating possible hydroxyl and carboxylic biomarkers, respectively. The MS spectra of the light and heavy labeled samples (Figure 5b,c) further confirmed the precursor ions of the paired peaks from the extracted ion chromatograms. The use of the selected paired peaks after chemical labeling as the targeted precursor ions facilitated the subsequent MS/MS analysis. The biomarker in Figure 5b ($m/z$ 356.1345/358.1397) was identified as 4-ethylphenol, an exposure that results from food ingestion. The carboxylic acid in Figure 5c ($m/z$ 465.2120/469.2265) identified as *trans,trans*-muconic acid, which is a urinary metabolite of benzene. In addition to these compounds, the identities of four other suspected biomarkers (listed in Table S8) that are commonly used in consumer products as plasticizers and preservatives were confirmed.

We further used the current method to measure several targeted biomarkers in 9 individual urine samples. After derivatization, the signal intensities for targeted compounds showed significant increases compared to the signal intensities for their free forms. The detection limit in our method was comparable to that obtained with LC-QQQ, which is considered as the gold standard for quantitation (Figure 5d). For example, the LOD values of BPA are 0.39 and 0.15 ng/mL for CIL-EXPOSOME and LC-QQQ, respectively. Furthermore, isotope labeling creates a convenient internal standard for the absolute quantification of all biomarkers, especially for the urine samples which exhibit significant and sample-dependent matrix effects. Because the internal standard coelutes with the analyte, it compensates for either ionization enhancement or suppression caused by the matrix. In conclusion, our method not only showed robust reliability in quantification but also encompassed a broader coverage of biomarkers compared with previous methods.


DOI: 10.1021/acs.est.9b00285
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Using this method, a high concentration of benzoresorcinol, methylparaben, 3,5-DHPPA, and genistein were detected in a pooled mixture of urine. As shown in Figure 5e, two abundant compounds, namely, 3,5-DHPPA (10 186.69 ± 262.04 ng/mL) and genistein (1 995.32 ± 202.04 ng/mL), were food additives. Another 7 hydroxyl compounds were also well-detected, including plasticizers BPA (49.21 ± 1.21 ng/mL), BPS (32.20 ± 2.76 ng/mL), and benzoresorcinol (157.11 ± 10.74 ng/mL); personal care products methylparaben (453.04 ± 41.89 ng/mL) and triclosan (379.65 ± 4.65 ng/mL); pesticide 4-bromophenol (4.03 ± 0.48 ng/mL); and PAH exposure biomarker naphthol (19.92 ± 3.76 ng/mL). The ubiquity of these biomarkers indicates that people are exposed to xenobiotics on a daily basis. Notably, most of these xenobiotics are biologically active as either endocrine-disrupting chemicals (e.g., BPA, BPS, and triclosan) or carcinogens (e.g., 4-bromophenol and naphthol).[37,38]

In conclusion, a high-coverage urinary exposure biomarker profiling method based on a stable isotope derivatization approach was established in this study and provided enhanced sensitivity and selectivity compared to existing methods. CIL-EXPOSOME has overcome several hurdles encountered in traditional exposome profiling platforms such as low sensitivity, complex sample preparation, high cost of isotope standards, and low-throughput profiling. Specifically, $^{12}C/^{13}C_2$-danysl chloride (DNS) and $^{12}C/^{13}C_2$-*p*-dimethylaminophenacyl bromide (DmAP) were applied to label hydroxyl/amines/carboxyl biomarkers, respectively, which are common functional groups in most urinary biomarkers. This labeling method, together with a simplified cleanup step, is effective not only in introducing an economical isotope tag for accurate biomarker quantification but also in improving the signal by up to several orders of magnitude in electrospray ionization. In addition, this strategy facilitated untargeted biomaker identification by selecting the paired peaks with a defined mass tolerance. Furthermore, we established a urinary exposome database with MS and MS/MS information for more than 800 compounds compiled from previous studies in human biological fluids to facilitate exposome characterization. Operation using this database has been automated using computational tools. Using the CIL-EXPOSOME platform, 671 exposure biomarker candidates from one injection were profiled in a 2 mL urine sample. In sum, CIL-EXPOSOME provides a sensitive, cost-effective, high-coverage, and automated exposome platform for the analysis of both known and unknown biomarkers. We believe that this platform provides a novel way to circumvent the bottlenecks of current exposome approaches and contributes to the future of exposome characterization.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.9b00285.

Detailed experimental results in Tests S1−S3, Tables S1−S8, and Figures S1−S3 (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

*Phone: +65 6790 5331; e-mail: mlfang@ntu.edu.sg.

### ORCID

Tao Huan: 0000-0001-6295-2435

Mingliang Fang: 0000-0002-2204-9783


### Notes

The authors declare no competing financial interest.

The CIL-EXPOSOME database, software, and usage protocol may be accessed at https://drive.google.com/drive/folders/1i1UNhfwMh_ry97TH6-FKGEm_A-i-oIeU?usp=sharing.

## ■ ACKNOWLEDGMENTS

This work is funded by Singapore Ministry of Education Academic Research Fund Tier 1 (M4011732.030), Start Up Grant of Nanyang Technological University (M4081915, M060030008), and Singapore National Environment Agency (M4061617). We greatly appreciate Dr. Li Liang's group from University of Alberta, Canada for the synthesis and shipment of the $^{13}C_2$-danysl chloride and $^{13}C_2$-*p*-dimethylaminophenacyl bromide standards. Dr. Dai Chao, Wang Mengjing, and Tan Jeremy from Nanyang Technological University, Singapore and Dr. Allison Phillips from Duke University (Durham, NC, United States) are gratefully acknowledged for helping with the program, revising English grammar, and database setup.

## ■ REFERENCES


(1) Wild, C. P. *Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology*. In AACR: 2005.

(2) Rappaport, S. M. Biomarkers intersect with the exposome. *Biomarkers* **2012**, *17* (6), 483−489.

(3) Patti, G. J.; Yanes, O.; Siuzdak, G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* **2012**, *13* (4), 263.

(4) Johnson, C. H.; Ivanisevic, J.; Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **2016**, *17* (7), 451.

(5) Dennis, K. K.; Marder, E.; Balshaw, D. M.; Cui, Y.; Lynes, M. A.; Patti, G. J.; Rappaport, S. M.; Shaughnessy, D. T.; Vrijheid, M.; Barr, D. B. Biomonitoring in the era of the exposome. *Environ. Health Perspect.* **2017**, *125* (4), 502−510.

(6) Rappaport, S. M.; Barupal, D. K.; Wishart, D.; Vineis, P.; Scalbert, A. The Blood Exposome and Its Role in Discovering Causes of Disease. *Environ. Health Perspect.* **2014**, *122* (8), 769−774.

(7) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. Global metabolic profiling procedures for urine using UPLC−MS. *Nat. Protoc.* **2010**, *5* (6), 1005.

(8) Athersuch, T. J.; Keun, H. C. Metabolic profiling in human exposome studies. *Mutagenesis* **2015**, *30* (6), 755−762.

(9) Miller, G. W.; Jones, D. P. The nature of nurture: refining the definition of the exposome. *Toxicol. Sci.* **2014**, *137* (1), 1−2.

(10) Rappaport, S. M.; Smith, M. T. Environment and disease risks. *Science* **2010**, *330* (6003), 460−461.

(11) Wild, C. P. The exposome: from concept to utility. *International journal of epidemiology* **2012**, *41* (1), 24−32.

(12) Griffiths, W. J.; Koal, T.; Wang, Y.; Kohl, M.; Enot, D. P.; Deigner, H. P. Targeted metabolomics for biomarker discovery. *Angew. Chem., Int. Ed.* **2010**, *49* (32), 5426−5445.

(13) Warth, B.; Spangler, S.; Fang, M.; Johnson, C. H.; Forsberg, E. M.; Granados, A.; Martin, R. L.; Domingo-Almenara, X.; Huan, T.; Rinehart, D.; Montenegro-Burke, J. R.; Hilmers, B.; Aisporna, A.; Hoang, L. T.; Uritboonthai, W.; Benton, H. P.; Richardson, S. D.; Williams, A. J.; Siuzdak, G. Exposome-Scale Investigations Guided by Global Metabolomics, Pathway Analysis, and Cognitive Computing. *Anal. Chem.* **2017**, *89* (21), 11505−11513.

(14) Zhou, J.; Yin, Y. Strategies for large-scale targeted metabolomics quantification by liquid chromatography-mass spectrometry. *Analyst* **2016**, *141* (23), 6362−6373.

(15) Thomas, C. E.; Sexton, W.; Benson, K.; Sutphen, R.; Koomen, J. Urine collection and processing for protein biomarker discovery and




DOI: 10.1021/acs.est.9b00285
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

quantification. *Cancer Epidemiol., Biomarkers Prev.* **2010**, *19* (4), 953−959.
(16) Xu, C.; Li, C. Y.-T.; Kong, A.-N. T. Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Arch. Pharmacal Res.* **2005**, *28* (3), 249.
(17) Zhao, S.; Luo, X.; Li, L. Chemical isotope labeling LC-MS for high coverage and quantitative profiling of the hydroxyl submetabolome in metabolomics. *Anal. Chem.* **2016**, *88* (21), 10617−10623.
(18) Guo, K.; Li, L. Differential 12C-/13C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal. Chem.* **2009**, *81* (10), 3919−3932.
(19) Guo, K.; Li, L. High-performance isotope labeling for profiling carboxylic acid-containing metabolites in biofluids by mass spectrometry. *Anal. Chem.* **2010**, *82* (21), 8789−8793.
(20) Arrivault, S.; Guenther, M.; Fry, S. C.; Fuenfgeld, M. M.; Veyel, D.; Mettler-Altmann, T.; Stitt, M.; Lunn, J. E. Synthesis and use of stable-isotope-labeled internal standards for quantification of phosphorylated metabolites by LC−MS/MS. *Anal. Chem.* **2015**, *87* (13), 6896−6904.
(21) Mung, D.; Li, L. Development of chemical isotope labeling LC-MS for milk metabolomics: comprehensive and quantitative profiling of the amine/phenol submetabolome. *Anal. Chem.* **2017**, *89* (8), 4435−4443.
(22) Yuan, B.-F.; Zhu, Q.-F.; Guo, N.; Zheng, S.-J.; Wang, Y.-L.; Wang, J.; Xu, J.; Liu, S.-J.; He, K.; Hu, T.; et al. Comprehensive Profiling of Fecal Metabolome of Mice by Integrated Chemical Isotope Labeling-Mass Spectrometry Analysis. *Anal. Chem.* **2018**, *90* (5), 3512−3520.
(23) Jiang, R.; Jiao, Y.; Zhang, P.; Liu, Y.; Wang, X.; Huang, Y.; Zhang, Z.; Xu, F. Twin Derivatization Strategy for High-Coverage Quantification of Free Fatty Acids by Liquid Chromatography−Tandem Mass Spectrometry. *Anal. Chem.* **2017**, *89* (22), 12223−12230.
(24) Parkinson, A. O. Biotransformation of Xenobiotics. In *Casarett & Doull's Essentials of Toxicology*; Klaassen, C. D., Watson, J. B., Eds.; McGraw-Hill: New York, 2010.
(25) Neveu, V.; Moussy, A.; Rouaix, H.; Wedekind, R.; Pon, A.; Knox, C.; Wishart, D. S.; Scalbert, A. Exposome-Explorer: a manually-curated database on biomarkers of exposure to dietary and environmental factors. *Nucleic Acids Res.* **2017**, *45* (D1), D979−D984.
(26) Seiler, N.; Deckardt, K. Determination of amines and amino acids in sugar-containing samples by dansylation. *Journal of chromatography* **1975**, *107* (1), 227−229.
(27) Minocha, R.; Long, S. Simultaneous separation and quantitation of amino acids and polyamines of forest tree tissues and cell cultures within a single high-performance liquid chromatography run using dansyl derivatization. *Journal of Chromatography A* **2004**, *1035* (1), 63−73.
(28) Toyo'oka, T. Fluorescent tagging of physiologically important carboxylic acids, including fatty acids, for their detection in liquid chromatography. *Anal. Chim. Acta* **2002**, *465* (1−2), 111−130.
(29) Chen, D.; Han, W.; Su, X.; Li, L.; Li, L. Overcoming Sample Matrix Effect in Quantitative Blood Metabolomics Using Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry. *Anal. Chem.* **2017**, *89* (17), 9424−9431.
(30) Huan, T.; Troyer, D. A.; Li, L. Metabolite analysis and histology on the exact same tissue: comprehensive metabolomic profiling and metabolic classification of prostate cancer. *Sci. Rep.* **2016**, *6*, 32272.
(31) Luo, X.; Zhao, S.; Huan, T.; Sun, D.; Friis, R. M. N.; Schultz, M. C.; Li, L. High-performance chemical isotope labeling liquid chromatography−mass spectrometry for profiling the metabolomic reprogramming elicited by ammonium limitation in yeast. *J. Proteome Res.* **2016**, *15* (5), 1602−1612.
(32) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; et al. HMDB: the human metabolome database. *Nucleic Acids Res.* **2007**, *35*, D521−D526.
(33) Woodruff, T. J.; Zota, A. R.; Schwartz, J. M. Environmental chemicals in pregnant women in the United States: NHANES 2003−2004. *Environ. Health Perspect.* **2011**, *119* (6), 878.
(34) Li, L.; Li, R.; Zhou, J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y.; Huan, T.; Zheng, J.; Shi, Y.; Wishart, D. S.; Lin, G. MyCompoundID: Using an Evidence-Based Metabolome Library for Metabolite Identification. *Anal. Chem.* **2013**, *85* (6), 3401−3408.
(35) Huan, T.; Tang, C.; Li, R.; Shi, Y.; Lin, G.; Li, L. MyCompoundID MS/MS Search: Metabolite Identification Using a Library of Predicted Fragment-Ion-Spectra of 383,830 Possible Human Metabolites. *Anal. Chem.* **2015**, *87* (20), 10619−10626.
(36) Allen, F.; Pon, A.; Wilson, M.; Greiner, R.; Wishart, D. CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Res.* **2014**, *42* (W1), W94−W99.
(37) vom Saal, F. S.; Nagel, S. C.; Coe, B. L.; Angle, B. M.; Taylor, J. A. The estrogenic endocrine disrupting chemical bisphenol A (BPA) and obesity. *Mol. Cell. Endocrinol.* **2012**, *354* (1), 74−84.
(38) Hecht, S. S. Human urinary carcinogen metabolites: biomarkers for investigating tobacco and cancer. *Carcinogenesis* **2002**, *23* (6), 907−922.


DOI: 10.1021/acs.est.9b00285
*Environ. Sci. Technol.* XXXX, XXX, XXX−XXX