Large-scale, enzyme-based xenobiotic identification for exposomics

Ken Liu
Emory University https://orcid.org/0000-0002-9736-5828
Choon Lee
Emory University School of Medicine
Grant Singer
Emory University School of Medicine
Michael Woodworth
Emory University
Thomas Ziegler
Emory University School of Medicine
Colleen Kraft
Emory
Gary Miller
Columbia University
Shuzhao Li
The Jackson Laboratory for Genomic Medicine
Young-Mi Go
Emory University
Edward Morgan
Emory University School of Medicine
Dean Jones (dpjones@emory.edu)
Emory University

Article

Keywords: exposomics, xenobiotic metabolites, high-resolution mass spectrometry, enzyme-based identification

Posted Date: November 5th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-77801/v1

Version of Record: A version of this preprint was published at Nature Communications on September 14th, 2021. See the published version at https://doi.org/10.1038/s41467-021-25698-x.

Abstract

Advances in genomics have revealed many of the genetic underpinnings of human disease, but exposomics methods are currently inadequate to obtain a similar level of understanding of environmental contributions to human disease. Exposomics methods are limited by low abundance of xenobiotic metabolites and lack of authentic standards, which precludes identification using solely mass spectrometry-based criteria. Here, we develop and validate a method for enzymatic generation of xenobiotic metabolites for use with high-resolution mass spectrometry (HRMS) for chemical identification. Generated xenobiotic metabolites were used to confirm identities of respective metabolites in mice and human samples based upon accurate mass, retention time and co-occurrence with related xenobiotic metabolites. The results establish a generally applicable enzyme-based identification (EBI) for mass spectrometry identification of xenobiotic metabolites.

Introduction

Humans are exposed to tens of thousands of xenobiotic chemicals from diet, drugs, environmental and occupational exposures, commercial products, and the microbiome[1]. Together with endogenously generated metabolites of the human metabolome, these are predicted to include over one million distinct chemical entities, representing a substantial analytical challenge for human exposome research[2,3]. To overcome the limitation in number of chemicals measured with targeted methods, high-resolution metabolomics (HRM) uses liquid chromatography with ultra-high resolution mass spectrometry and computational methods to measure chemicals as mass spectral signals without *a priori* knowledge of chemical identity[3]. This fills an important need to deliver -omics scale chemical data for environmental epidemiology to detect unrecognized health hazards[4] but increases the need for new approaches for chemical identification.

Xenobiotics are typically three to five orders of magnitude lower in abundance than endogenous metabolites and may be present in only a small fraction of human samples[5]. To address these challenges, HRM uses a dual liquid chromatography protocol with positive electrospray ionization (ESI) and hydrophilic interaction chromatography (HILIC) paired with negative ESI and C18 chromatography. Each sample is analyzed with three technical replicates, and routine analyses provide relative quantification of over 35,000 accurate mass spectral features [defined by accurate mass mass-to-charge ratio (*m/z*) (± 3 ppm), retention time, and intensity]. Chemical identification is severely limited because mass spectrometry-based criteria require authentic standards and ion dissociation spectra (MS/MS or MS^n).[6,7] Standards are not available for most xenobiotic metabolites and only the top 1000–2000 highest intensity features have useful MS^2 or MS^n spectra. Therefore, additional procedures are needed to complement accurate mass *m/z* for chemical identification.

Human enzymes convert xenobiotics to phase I and II metabolites. Usually, these related xenobiotic metabolites are expected to be present in samples with real exposures. With HRM, computational approaches provide a strategy to test for co-detection of related xenobiotic metabolites by use of

biological precursor-product relationships and metabolite-metabolite correlations. For example, multiple metabolites of naphthalene and other polycyclic aromatic hydrocarbon (PAH) were correlated and detected in military personnel exposed to burn pits[8], providing credibility for PAH exposure despite signal intensities too low for MS/MS analysis. Similarly, a metabolome-wide association study of the dichlorodiphenyltrichloroethane (DDT) metabolite chlorophenyl acetic acid identified two other correlated DDT metabolites[9]. While *in silico* tools for prediction of xenobiotic metabolites[10] and associated fragmentation spectra[11–13] are available, the lack of readily available authentic standards for most xenobiotic biotransformation products limits our ability to identify xenobiotic metabolites in real samples.

In the present study, we developed a system to generate xenobiotic metabolites in a high-throughput manner to enhance mass spectrometry capabilities to identify xenobiotic metabolites in human samples. We used biological preparations with xenobiotic biotransforming enzyme activities (pooled human liver S9 fractions) to generate metabolites from a panel of xenobiotics with known biotransformation products. We characterized metabolic products by accurate mass *m/z*, retention time, MS^n and stable isotope methods and matched these to xenobiotic metabolites detected in mouse and human circulation. Our data show that related xenobiotic metabolites co-occur in human samples with documented exposures, and we apply this principle to identify undocumented environmental exposures. The results establish a method to provide authentic xenobiotic metabolites and associated stable isotopic forms for use with HRM to improve confidence in identification of low abundance xenobiotic metabolites in humans.

Results

Development of enzyme-based xenobiotic metabolism for production of xenobiotic metabolites

We aimed to develop a platform that enabled high-throughput enzymatic production of phase I and phase II xenobiotic metabolites. We used pooled human liver S9 fractions because they are more representative of hepatocyte metabolism compared to microsomes and are more amenable for high-throughput screening applications compared to hepatocytes[6]. S9 fractions contain both microsomal and cytosolic subcellular fractions and include most phase I (cytochrome P450s, flavin monooxygenases) and phase II (sulfotransferases, aldehyde oxidases, methyl transferases, glutathione S-transferase, N-acetyl transferase, UDP-glucuronosyltransferases) biotransformation enzymes[14-20]. We prepared human liver S9 fractions with required cofactors in a 96-well plate format (Fig. 1a) which enabled the production of phase I and phase II metabolites (Fig. 1b) from multiple xenobiotics in a single plate. Reaction extracts were analyzed with HRM to build retention time and spectral libraries for use with HRM analyses of human samples (Figure 1c).

In vitro production of phase I and phase II metabolites by human liver S9 fractions

To test whether standard conditions for S9 incubations could be used to produce biotransformation products of diverse xenobiotics, we tested 9 xenobiotics with well-characterized phase I and phase II

metabolism. Expected phase I and phase II metabolites were not detected at time 0 and increased with incubation time. For example, MS^1 peak areas for hydroxy beta-naphthoflavone, beta-naphthoflavone dihydrodiol, hydroxy beta-naphthoflavone glucuronide, and hydroxy beta-naphthoflavone sulfate (Fig. 2a-d) increased with time and were not detected in the 0 hour sample. Overall, more than 90% of expected metabolites were detected in incubation extracts after 24 hours (Fig. 2e). It is important to note that not all metabolites were not detected at sufficient levels to collect ion dissociation spectra. Extracted ion chromatograms and summarized data (*m/z*, adduct form, retention time, analytic method) for other xenobiotic precursors (acetaminophen, caffeine, nicotine, chlorzoxazone, tolbutamide, bupropion, benzo[a]pyrene) are available (Supplemental Fig. 1, Supplemental Table 1). Thus, the present results show that use of standardized human liver S9 fraction incubations is suitable as a platform for production of expected phase I and phase II metabolites of diverse xenobiotics.

Use of stable isotopes and biochemical knowledge for metabolite identification

To further characterize S9-enzyme generated metabolites, we used isotopically labeled xenobiotic precursors to aid structure elucidation. Bupropion is mainly metabolized by CYP2B6, which performs a stereospecific hydroxylation to form 2,3-hydroxybupropion[21]. However, 4-hydroxybupropion, which shares the same exact mass, is also an expected minor product. Using d_9-bupropion with S9 enzymes, we can differentiate between the formation of 2,3-hydroxybupropion (Fig. 3a) and 4-hydroxybupropion (Fig. 3b). 2,3-hydroxybupropion is a heterocycle which loses one deuterium; whereas, 4-hydroxybupropion is formed without loss of the deuterium label. Our data show the major product is d_8-2,3-hydroxybupropion (264.1600 *m/z*) (Fig. 3c), with peak intensities 10x higher than d_9-4-hydroxybupropion (265.1666 *m/z*) after 24 hours (Fig. 3d). Using unlabeled bupropion, 256.1099 *m/z* alone does not discriminate between 2,3-hydroxybupropion and 4-hydroxybupropion (Fig. 3e). MS/MS spectra collected from 256.1099 *m/z* and 264.1600 *m/z* is consistent with 2,3-hydroxybupropion and d_8-2,3-hydroxybupropion (Fig. 3f).

Use of stable isotopes for identification of unreported metabolites

Full scan data collection for HRMS enables untargeted analysis to identify potentially unreported metabolites. Here, we tested whether other spectral features that increased with time of S9 enzyme incubation could be unreported xenobiotic metabolites. To distinguish non-specific reaction products from reaction products of test compound(s), we implemented an additional filter to account for mass differences due to labeling. For example, in the unlabeled caffeine reaction, 485 features were positively associated ($R > 0.9$) with time. Of these 485 features, we were able to identify two metabolites that were similarly increased with the addition of isotopically labeled ($^{13}C_3$ or D_3) caffeine. Paraxanthine is the expected major product of caffeine metabolism and we observed increases in both unlabeled (181.0720 *m/z* 31 second RT) and labeled ($^{13}C_2$ – 183.0786 *m/z* 31 second RT, D_3 – 184.0909 *m/z* 31 second RT) forms. Then, we identified an unexpected metabolite of caffeine with 213.0981 *m/z* (Fig. 4a, unlabeled, 31 second RT) with formation of the isotopically labeled forms ($^{13}C_3$ – 216.1080 *m/z* (Fig. 4b), D_3 –

216.1170 *m/z* (Fig. 4c)) at the same retention time. The metabolite 213.0981 *m/z* corresponds to an elemental composition $C_8H_{12}N_4O_3$, which corresponds to the addition of two hydrogens and one oxygen atom. MS/MS analysis of 213.0981 *m/z* (Fig. 4d), 216.1080 *m/z* (Fig. 4e), and 216.1170 *m/z* (Fig. 4f) were consistent with an oxidation (+O) and reduction (+2H) or hydration (+H_2O) of the caffeine imidazole. A proposed structure for this unreported caffeine metabolite is provided (Fig. 4g-i). This metabolite was detected in mice treated with caffeine and also correlated with caffeine in humans.

Identification of xenobiotics in vivo from documented exposures

Analysis of S9 reaction extracts provides authentic xenobiotic metabolites for generation of MS^1, MS/MS, and retention times, which can support identification of xenobiotic metabolites in human and mouse samples. Two of the criteria required for metabolite identification in human and mouse samples can be satisfied by use of *m/z* (± 3 ppm) and retention time windows (± 5 seconds) for matching compounds detected in study samples. If MS/MS cannot be collected for the study samples due to low abundance, a third criteria for identification can be satisfied if expected ions at their characteristic retention times are detected only in animal or human samples with documented exposure. For example, in mice treated with bupropion, bupropion (Fig. 5a), hydroxybupropion (Fig. 5b), and hydrobupropion (Fig. 5c) are detected at the same accurate mass *m/z* and retention time as S9-generated metabolites.

To test this concept in human samples, we used plasma samples collected from patients with well-documented pharmaceutical use (electronic medical records) receiving care at Emory University Hospital. HRM analysis shows that in samples collected from the one individual who was taking bupropion, bupropion (Fig. 5d), hydroxybupropion (Fig. 5e), and hydrobupropion (Fig. 5f) were detected at the same accurate mass *m/z* and retention time as S9-generated metabolites. Furthermore, we observed that hydroxybupropion, and hydrobupropion were only detected in the samples where bupropion was also detected (Fig. 5g). Acetaminophen, carvedilol, warfarin, and metoprolol were other medications that were used by individuals in this cohort. To evaluate whether expected metabolites of these medications were correlated with the initial parent compound, we used xMWAS software, a tool for multi-parameter metabolome-wide association analyses. Then, we applied a mass difference filter ($m/z_{feature} - m/z_{parent}$) on correlated metabolites to generate reaction-based networks. Partial-least squares regression analysis ($r > 0.4$) revealed 5 distinct clusters, each centered around a single parent compound. Each parent compound (bupropion Fig. 5a, acetaminophen Fig. 5i, metoprolol Fig. 5j, carvedilol Fig. 5k, warfarin Fig. 5l) was correlated with at least one expected biotransformation product and their associated source fragments, adducts, and isotopes (Supplemental Fig. 2). These metabolites were all detected at the same accurate mass *m/z* and retention times as those produced from S9 reactions. Thus, use of enzyme-generated xenobiotic metabolites as authentic metabolites together with use of co-detection of related metabolites provides a strategy for pathway-level identification of xenobiotics and xenobiotic metabolites.

Identification of xenobiotics in vivo from unknown exposures

To test whether use of accurate mass *m/z*, RT and co-detection of related xenobiotics could be used for identification of xenobiotics in human samples with undocumented environmental exposures, we examined a healthy cohort with no documented xenobiotic exposures. Nicotine is a common environmental xenobiotic that humans can be exposed to through tobacco products and secondhand smoke. Nicotine, cotinine and hydroxycotinine were detected in human samples at the same accurate mass *m/z* and retention times as detected from analysis of S9 reaction extracts (Fig. 6a). Cotinine and hydroxycotinine were co-detected (Fig. 6b) and correlated (Fig. 6c) with nicotine, thereby providing confidence in the identification of the xenobiotic metabolites. In hospitalized patients receiving care at Emory Hospital without access to nicotine-containing products, a mass spectrometry signal was detected at 163.1230 *m/z* but cotinine and hydroxycotinine were not detected in these samples. Thus, co-detection of related xenobiotic metabolites at their established accurate mass *m/z* and retention time provides a strategy for confident identification. Coupled to use of high-throughput S9 enzymatic systems, this approach can be used to complement traditional approaches to MS-based metabolite identification.

Discussion

The purpose of this study was to develop strategies to aid identification of low abundance xenobiotic metabolites detected by HRM. These features, which comprise a majority of detected features, are not easily identified using only mass spectrometry-based criteria which require matching MS^1, MS^2 and/or MS^n spectra and retention times of unknown metabolites with those obtained from analysis of authentic standards[6,7]. Confidence in metabolite identification increases when multiple lines of evidence are provided. For example, while use of a single accurate mass MS^1 is not sufficient to assign a single elemental composition[22], co-detection of naturally occurring isotopic ions and adducts increases confidence in assigning elemental compositions to detected ions. Along the same vein, incorporating biological criteria based on co-detection of related xenobiotic metabolites increases confidence in identification of xenobiotic exposures in humans. Since MS^2 and MS^n cannot always be collected from low abundance xenobiotics, generation of xenobiotic pathway metabolites using enzymatic systems enables co-detection of related xenobiotics and xenobiotic metabolites (at characteristic *m/z* and RT) to be used as criteria for metabolite identification. These strategies provide a complementary approach to increase confidence of metabolite identifications, especially for low abundance xenobiotic metabolites.

Here, we demonstrate that pooled human liver S9 fraction preparations in 96-well formats provide a generalized reaction system suitable for generating Phase I and Phase II metabolites of several xenobiotics in a single experiment. We demonstrate that addition of stable-isotope precursors to metabolite identification workflows aids structural determinations and identification of unreported metabolites. Furthermore, we show that xenobiotic metabolites produced from S9 reactions can be considered authentic metabolites for identification. Using five-minute dual chromatography HRMS methods aimed at increasing sample throughput, each 96-well experiment requires one day of duplicate sample analysis and can test eleven com-pounds with four time points. Using this strategy with the

assumption that each parent xenobiotic reaction generates five metabolites, fifty-five biotransformation products plus eleven parent xenobiotics could be characterized in a single day of sample analysis. Adoption of pooled xenobiotic mixtures into experimental workflows could, in principle, increase reaction throughput to more efficiently cover exposome space. As compound libraries are screened with this approach, it will be necessary to automate data acquisition and analysis pipelines to enhance data sharing across laboratories.

While pooled human liver S9 fractions are useful models to predict primary human metabolites *in vivo*[23], other biology-based systems using cell lines, primary human hepatocytes, or other co-cultured cell systems[24] could also be used for this application. Future studies could perform a systematic comparison of the performance of our S9 incubations with cell-based systems and the ability to recapitulate metabolites detected in human and mouse studies.

Some limitations to this approach require discussion. Not all xenobiotics undergo appreciable biotransformation from human enzymes. For these cases, use of co-detection of related xenobiotics as a criterion for identification cannot be used. Furthermore, heterogeneity in xenobiotic metabolizing enzyme expression in populations could cause differences in expression of Phase I and II metabolites that are co-detected in humans. Additionally, without use of appropriately labeled stable isotope precursors or targeted analytic strategies (improved chromatographic separation or NMR), complete characterization of stereochemistry or regiochemistry is not yet possible using mass spectrometry alone. For these cases, metabolites should be identified with broader nomenclature (e.g. hydroxybupropion versus 2,3-hydroxybupropion or 4-hydroxybupropion) without use of positional modifications or stereochemistry, unless additional evidence is provided. Use of MS/MS spectra or targeted chromatographic methods can be used to increase selectivity of identification in these cases.

In conclusion, use of biologically-based enzymatic systems such as S9 fractions to generate mixtures of authentic xenobiotic metabolites for identification of xenobiotic pathways provides a useful strategy for identifying low abundance xenobiotics detected in human samples. This approach could be adapted to work with cell-based systems and could be scaled to produce thousands of xenobiotic metabolites in a single experiment. Used alongside biological criteria based on co-detection of multiple verified xenobiotics within a pathway, adoption of a community-based xenobiotic identification strategy could be developed to address the current limitations in identifying metabolites of the tens of thousands of chemicals used in commerce[25]. Ultimately, combining the power of multiplexed enzymatic reactions of xenobiotics with computational strategies for xenobiotic pathway enrichment analysis under development for mass spectrometry data would enable higher confidence identification of unknown xenobiotic exposures.

Methods

Materials: Acetaminophen (APAP), β-naphthoflavone (β-NF), caffeine, 13C3-caffeine, nicotine, chlorzoxazone (CLZ), tolbutamide, bupropion, warfarin, carvedilol, metoprolol, adenosine 3'-phosphate 5'-

phosphosulfate lithium salt hydrate (PAPS), glutathione (GSH), acetyl-coenzyme A (acetyl-CoA), and HPLC grade acetonitrile were purchased from Sigma-Aldrich (St. Louis, MO). D3-Caffeine and D9-bupropion were purchased from CDN isotopes (Point-Clare, Quebec, Canada). Pazopanib, benzo[a]pyrene, and uridine diphosphate glucuronic acid (UDPGA) were purchased from US Biological (Salem, MA). Pooled human liver S9 fraction: mixed gender (from 50 livers, H0640.S9) and NADPH regenerating system (K5000-10) were from Sekisui Xenotech (Kansas City, KS).

Xenobiotic incubations: Incubations of xenobiotics with human liver S9 fraction were carried out with minor modifications to the method of Richardson[23]. Pooled human liver S9 fractions (20 mg/ml protein) were aliquoted into 0.5 ml microcentrifuge tubes and, stored at − 80 °C, and then thawed at room temperature prior to use. Stock solutions of 5 mM β-NF, tolbutamide, bupropion, pazopanib, and benzo[a]pyrene were prepared in DMSO, while 5 mM stock solutions of APAP, caffeine, nicotine, and chlorzoxazone were prepared in water. Each test compound stock solution was diluted to 0.3 mM with water prior to incubation with S9 enzymes. The NADPH regenerating system was reconstituted with addition of 3.5 ml of water to make a final volume of 5 ml. Cofactors were combined to form a 4X cofactor stock as follows, prior to addition into the reaction mixture: 10 mM UDPGA, 2 mM GSH, 2 mg/ml PAPS, 0.1 mM acetyl-CoA, and NADPH regenerating system (1 mM NADP, 5 mM glucose-6-phosphate, 1 unit glucose-6-phosphate dehydrogenase). Reactions were carried out at 30 °C on 96-well plates (Fig. 1). S9 fraction was diluted 10-fold in water immediately before mixing with 0.2 M Tris-Cl, pH 7.5/2 mM $MgCl_2$ and 0.3 mM of the xenobiotic solution in a 1:1:1 ratio (15 µL each). and incubated at 30 °C for 5 min. To start the reaction, 15 µL of 4X cofactor stock was added, and incubation was carried out at 30 °C for the indicated times. To terminate the reaction, we added a three-fold volume of acetonitrile, covered the plate with parafilm, vortexed, and froze at − 20 °C to precipitate insoluble materials such as protein. After thawing and centrifugation of the incubation plate, the supernatants were transferred into polypropylene autosampler vials, which were stored at − 20 °C until instrumental analysis.

High-resolution mass spectrometry

10 µL aliquots of sample extracts were analyzed using liquid chromatography (Dionex Ultimate 3000) and Fourier-transform high-resolution mass spectrometry (Thermo Scientific Fusion). The chromatography system was operated in a dual pump configuration to enable parallel analyte separation and column flushing. Sample extracts were injected and analyzed using hydrophilic interaction liquid chromatography (HILIC) with electrospray ionization (ESI) operated in positive mode and reverse phase (C18) chromatography with ESI operated in negative mode. For S9 reactions extracts, each analysis was performed using two biological replicates and a single technical replicate injection per column. For mouse and human samples, each analysis was performed using three technical replicate injections per column. Analyte separation for HILIC was accomplished using a Waters XBridge BEH Amide XP HILIC column (2.1 mm x 50 mm, 2.5 µm particle size) and eluent gradient (A = water, B = acetonitrile, C = 2% formic acid) consisting of an initial 1.5 min period of 22.5% A, 75% B, 2.5% C, followed by a linear increase to 75% A, 22.5% B, 2.5% C at 4 min and a final hold of 1 min. C18 chromatography was performed using Higgins Targa C18 2.1 mm x 50 mm, 3 µm particle size) column with an eluent gradient

(A = water, B = acetonitrile, C = 10 mM ammonium acetate) consisting of an initial 1 min period of 60% A, 35% B, 5% C, followed by a linear increase to 0% A, 95% B, 5% C at 3 min and held for the remaining 2 min. For both methods, mobile phase flow rate was held at 0.35 ml/min for the first 1 min, then increased to 0.4 ml/min for the final 4 min. The high-resolution mass spectrometer collected data at 120,000 resolution for MS^1 data collection. Data were collected from 85 – 1,275 *m/z* for MS^1 analysis with additional fragmentation spectra collected for predicted xenobiotic metabolites of a test substrate using a targeted inclusion list (Supplemental Table 2) with data-dependent MS/MS.

Data processing: Expected metabolites for each xenobiotic substrate were identified using literature searches and the web-based metabolite prediction tool Biotransformer (http://biotransformer.ca)[7]. Raw MS^1 and MS^2 data were processed in xCalibur QualBrowser. Extracted ion chromatograms were generated from accurate mass MS^1 spectra for expected M + H for HILIC + or M-H for C18- ions ± 3 ppm and plotted using GraphPad Prism 6.0.

Metabolite Identification Criteria

At least three orthogonal criteria were used for metabolite identification. Authentication of xenobiotics was based on 1) presence of appropriate MS signals only in samples where xenobiotic was added, 2) accurate mass MS^1 ± 3 ppm of expected masses, and MS^2 spectra matching known or predicted patterns. Identification of xenobiotic metabolites produced from S9 was based on 1) presence of predicted xenobiotic metabolite only in samples where xenobiotic was added, 2) accurate mass MS^1 ± 3 ppm of expected mass for predicted enzymatic product and 3) time-dependent increase in MS^1 intensity only in samples where co-substrate for enzymatic reaction was added or absence at time zero. Identification in clinical and experimental samples was based upon 1) detection of the expected MS^1 mass only in the samples with documented xenobiotic exposure matching the accurate mass MS^1 ± 3 ppm of observed mass from authentic xenobiotic 2) coelution (± 5 seconds) with the authentic xenobiotic and 3) Co-occurrence of more than one related xenobiotic at characteristic *m/z* and retention time as established from analysis of S9 extracts. Ion dissociation spectra (MS^2) were collected when available and provided a fourth criterion for identification.

Identification of unreported metabolites with stable isotopes

Caffeine, D_3-caffeine, $^{13}C_3$-caffeine, bupropion, and D_9-bupropion were incubated separately using the S9 fraction conditions described with samples collected at 0 hours and 24 hours. Following instrumental analysis, data from HILIC/ESI + analysis were extracted and aligned using mzMine2[26] with extraction and alignment parameters provided in Supplemental Table 3. Metabolites increased at 24 hours were identified using a Pearson's correlation (R > 0.9 for *m/z* peak intensity vs. time) for labeled and unlabeled experiments. A mass difference filter (based on mass differences between expected isotope-labeled forms and native forms) was applied to identify metabolites with increases in a corresponding labeled

form. Features that did not have a corresponding increase in a labeled form which co-eluted with the unlabeled form were considered non-specific metabolites.

Human samples: EDTA plasma samples (n = 51) from non-smoking individuals with documented pharmaceutical use were from 9 patients participating in a clinical trial at Emory University Hospital (ClinicaTrials.gov Identifier: NCT02922816, IRB00090101). Of these 9 patients, 6 (67/%) were female, 7 (78%) were white, 3 (33%) were Hispanic, and their median age was 62 years old (IQR 53–70). Repeat samples from each patient were collected at day 1, day 15 and day 36 of each study visit cycle. During the trial, each individual continued medications prescribed as part of routine clinical care. Each individual had a distinct set of medications taken as prescribed. Other demographic information on these samples were not available. EDTA plasma samples from individuals with no documented pharmaceutical use (n = 214) were from a subset of samples from adults enrolled in the Emory-Georgia Tech Centers for Health Discovery and Well-Being (CHDWB) study[27–29] (IRB00007243). These individuals were primarily employed by Emory University or Emory healthcare systems. Spectral feature tables from these studies were extracted and aligned using apLCMS[30,31] and xMSanalyzer[32]. Correlation networks of xenobiotics were generated using xMWAS[33] using the PLS regression function with a correlation threshold of 0.4.

Declarations

ACKNOWLEDGMENT

This work was supported by U2C-ES030163, P30 ES019776, RC2-DK118619, S10-OD018006, UL1-TR002378 Georgia Clinical and Translational Science Alliance (Georgia CTSA), UM1AI104681, K23AI144036, SSCI Research Scholar Award.

References

1 Idle, J. R. & Gonzalez, F. J. Metabolomics.

2 Uppal, K. *et al.* Computational Metabolomics: A Framework for the Million Metabolome. *Chem Res Toxicol* 29, 1956-1975, doi:10.1021/acs.chemrestox.6b00179 (2016).

3 Jones, D. P. Sequencing the exposome: A call to action. *Toxicol Rep* 3, 29-45, doi:10.1016/j.toxrep.2015.11.009 (2016).

4 Jones, D. P. & Cohn, B. A. A vision for exposome epidemiology: The pregnancy exposome in relation to breast cancer in the Child Health and Development Studies.

5 Rappaport, S. M. Implications of the exposome for exposure science. *Journal of Exposure Science & Environmental Epidemiology* 21, 5-9, doi:10.1038/jes.2010.50 (2011).

6 Schymanski, E. L. *et al.* Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol* 48, 2097-2098, doi:10.1021/es5002105 (2014).

7 Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 3, 211-221, doi:10.1007/s11306-007-0082-2 (2007).

8 Walker, D. I. *et al.* Deployment-Associated Exposure Surveillance With High-Resolution Metabolomics. *J Occup Environ Med* 58, S12-21, doi:10.1097/JOM.0000000000000768 (2016).

9 Walker, D., Go, Y.-M., Liu, K., Pennell, K. & Jones, D. 167-211 (2016).

10 Djoumbou-Feunang, Y. *et al.* BioTransformer: a comprehensive computational tool for small molecule metabolism prediction and metabolite identification. *J Cheminform* 11, 2, doi:10.1186/s13321-018-0324-5 (2019).

11 Ruttkies, C., Schymanski, E. L., Wolf, S., Hollender, J. & Neumann, S. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* 8, 3-3, doi:10.1186/s13321-016-0115-9 (2016).

12 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 Research* 42, W94-W99, doi:10.1093/nar/gku436 (2014).

13 Djoumbou-Feunang, Y. *et al.* CFM-ID 3.0: Significantly Improved ESI-MS/MS Prediction and Compound Identification. *Metabolites* 9, doi:10.3390/metabo9040072 (2019).

14 Kolrep, F., Rein, K., Lampen, A. & Hessel-Pras, S. Metabolism of okadaic acid by NADPH-dependent enzymes present in human or rat liver S9 fractions results in different toxic effects.

15 Li, Y. *et al.* Current trends in drug metabolism and pharmacokinetics. *Acta Pharmaceutica Sinica B* 9, 1113-1144, doi:https://doi.org/10.1016/j.apsb.2019.10.001 (2019).

16 Rudolph, W., Remane, D., Wissenbach, D. K. & Peters, F. T. Comparative study on the metabolism of the ergot alkaloids ergocristine, ergocryptine, ergotamine, and ergovaline in equine and human S9 fractions and equine liver preparations.

17 Varkhede, N. R., Jhajra, S., Ahire, D. S. & Singh, S. Metabolite identification studies on amiodarone in in vitro (rat liver microsomes, rat and human liver S9 fractions) and in vivo (rat feces, urine, plasma) matrices by using liquid chromatography with high-resolution mass spectrometry and multiple-stage mass spectrometry: Characterization of the diquinone metabolite supposedly responsible for the drug's hepatotoxicity. *Rapid Communications in Mass Spectrometry* 28, 311-331, doi:10.1002/rcm.6787 (2014).

18 Caspar, A. T., Meyer, M. R., Westphal, F., Weber, A. A. & Maurer, H. H. Nano liquid chromatography-high-resolution mass spectrometry for the identification of metabolites of the two new psychoactive substances N-(ortho-methoxybenzyl)-3,4-dimethoxyamphetamine and N-(ortho-methoxybenzyl)-4-methylmethamphetamine.

19 Singh, R., Chang, S. Y. & Taylor, L. C. E. In vitro Metabolism of a Potent HIV-protease Inhibitor (141W94) Using Rat, Monkey and Human Liver S9. *Rapid Communications in Mass Spectrometry* 10, 1019-1026, doi:10.1002/(sici)1097-0231(19960715)10:9<1019::Aid-rcm618>3.0.Co;2-j (1996).

20 Chalbot, S. & Morfin, R. Human liver S9 fractions: Metabolism of dehydroepiandrosterone, epiandrosterone, and related 7-hydroxylated derivatives. *Drug Metabolism and Disposition* 33, 563-569, doi:10.1124/dmd.104.003004 (2005).

21 Coles, R. & Kharasch, E. D. Stereoselective metabolism of bupropion by cytochrome P4502B6 (CYP2B6) and human liver microsomes. *Pharm Res* 25, 1405-1411, doi:10.1007/s11095-008-9535-1 (2008).

22 Kind, T. & Fiehn, O. Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinformatics* 7, 234, doi:10.1186/1471-2105-7-234 (2006).

23 Richardson Sj Fau - Bai, A., Bai A Fau - Kulkarni, A. A., Kulkarni Aa Fau - Moghaddam, M. F. & Moghaddam, M. F. Efficiency in Drug Discovery: Liver S9 Fraction Assay As a Screen for Metabolic Stability.

24 Peisl, B. Y. L., Schymanski, E. L. & Wilmes, P. Dark matter in host-microbiome metabolomics: Tackling the unknowns-A review. *Anal Chim Acta* 1037, 13-27, doi:10.1016/j.aca.2017.12.034 (2018).

25 Liu, K. H. *et al.* Reference Standardization for Quantification and Harmonization of Large-Scale Metabolomics. *Analytical Chemistry*, doi:10.1021/acs.analchem.0c00338 (2020).

26 Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11, 395, doi:10.1186/1471-2105-11-395 (2010).

27 Rask, K. J., Brigham Kl Fau - Johns, M. M. E. & Johns, M. M. Integrating comparative effectiveness research programs into predictive health: a unique role for academic health centers.

28 Brigham, K. L. Predictive health: the imminent revolution in health care. *J Am Geriatr Soc* 58 Suppl 2, S298-S302, doi:10.1111/j.1532-5415.2010.03107.x (2010).

29 Tabassum, R. *et al.* A Longitudinal Study of Health Improvement in the Atlanta CHDWB Wellness Cohort.

30 Yu, T. & Jones, D. P. Improving peak detection in high-resolution LC/MS metabolomics data using preexisting knowledge and machine learning approach.

31 Yu, T., Park, Y., Johnson, J. M. & Jones, D. P. apLCMS--adaptive processing of high-resolution LC/MS data. *Bioinformatics (Oxford, England)* 25, 1930-1936, doi:10.1093/bioinformatics/btp291 (2009).

32 Uppal, K. S., Q. A.; Strobel, F. H.; Pittard, W. S.; Gernert, K. M.; Yu, T.; Jones, D. P. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC Bioinformatics* 14, doi:10.1186/1471-2105-14-15 (2013).

33 Uppal, K., Ma, C., Go, Y. M., Jones, D. P. xMWAS: a data-driven integration and differential network analysis tool.

Supplementary Tables

Supplemental Tables 1-3 are not available with this version.

Figures

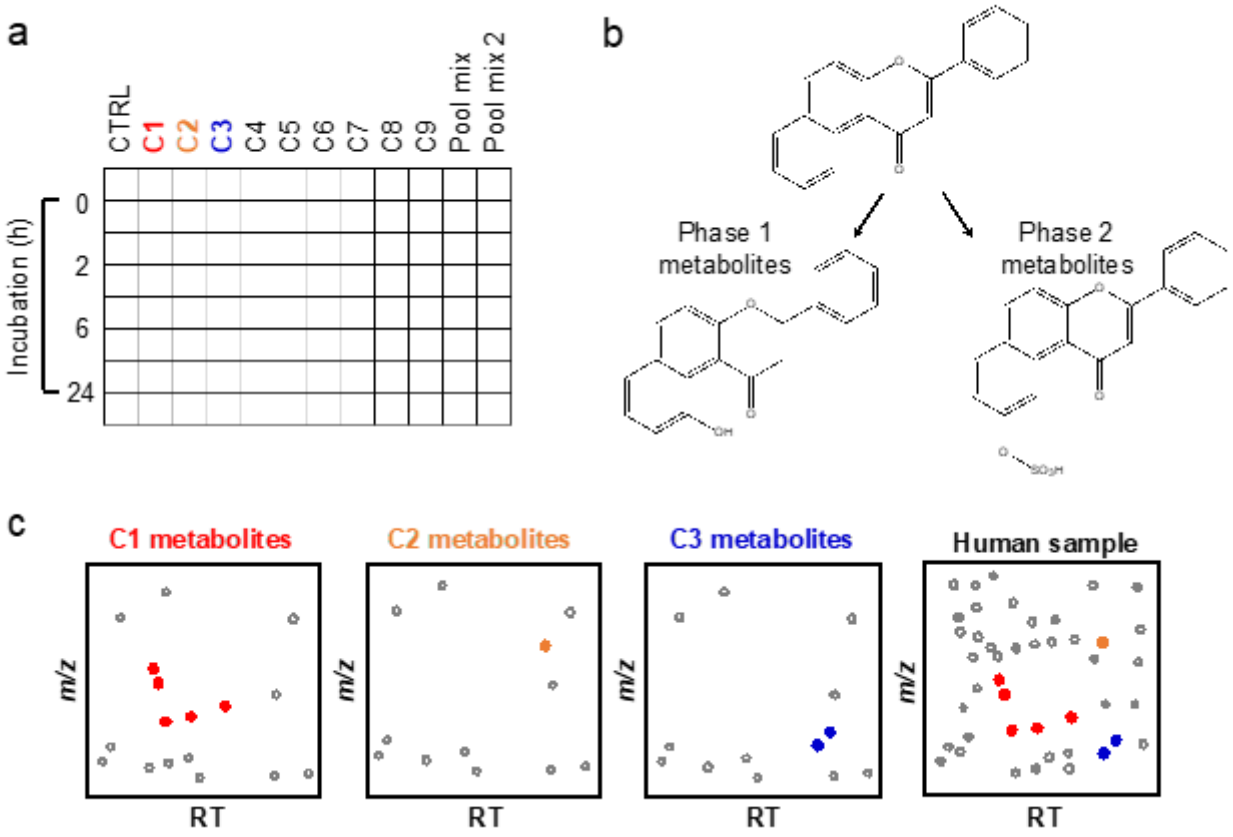

Figure 1

Enzymatic generation of xenobiotic metabolites for mass spectrometry-based chemical identification. a Human liver S9 enzymes prepared in a 96-well plate format capable of performing multiple xenobiotic reactions in a single plate. b Human liver S9 enzymes perform Phase I and II biotransformations to generate xenobiotic metabolites or other downstream adducts from reactive intermediates. c HRM analysis of enzyme-generated xenobiotic metabolites provide authentic metabolites for matching accurate mass m/z and retention time (RT) to detected metabolites in humans. Colored dots in each m/z RT dot plot represent spectral features that are identified from in vitro S9 reactions that could be matched to features detected in clinical or experimental studies.

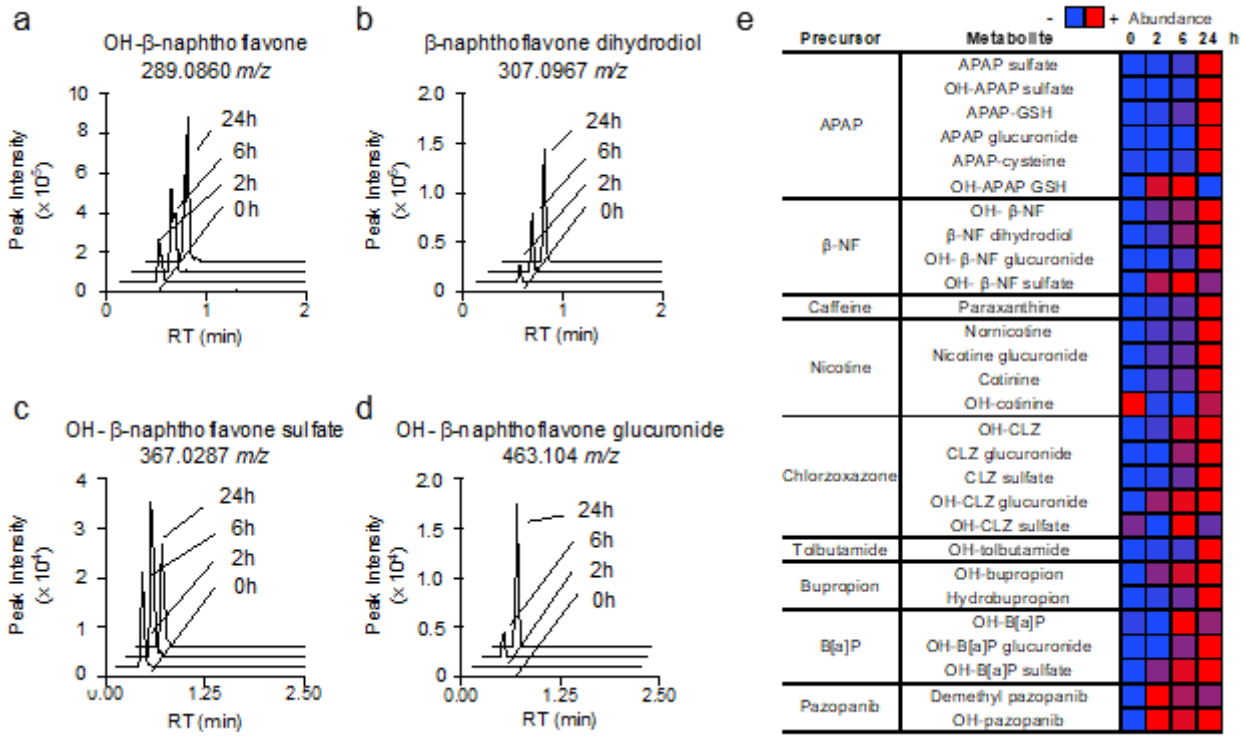

Figure 2

Human liver S9 enzymes generate Phase I and II xenobiotic metabolites in a time-dependent manner. Using standard conditions for S9 reactions, expected Phase I (a, b) and II (c, d) metabolites are formed in a time-dependent manner. Each plot depicts the extracted ion chromatograms of predicted metabolites of beta-naphthoflavone at 0, 2, 6, and 24 hours. Metabolites were identified if they produced a time-dependent increase in MS1 signal corresponding to the accurate mass of a predicted or expected metabolite. e Formation of expected metabolites in a time-dependent manner for a variety of xenobiotics using S9 enzymes.

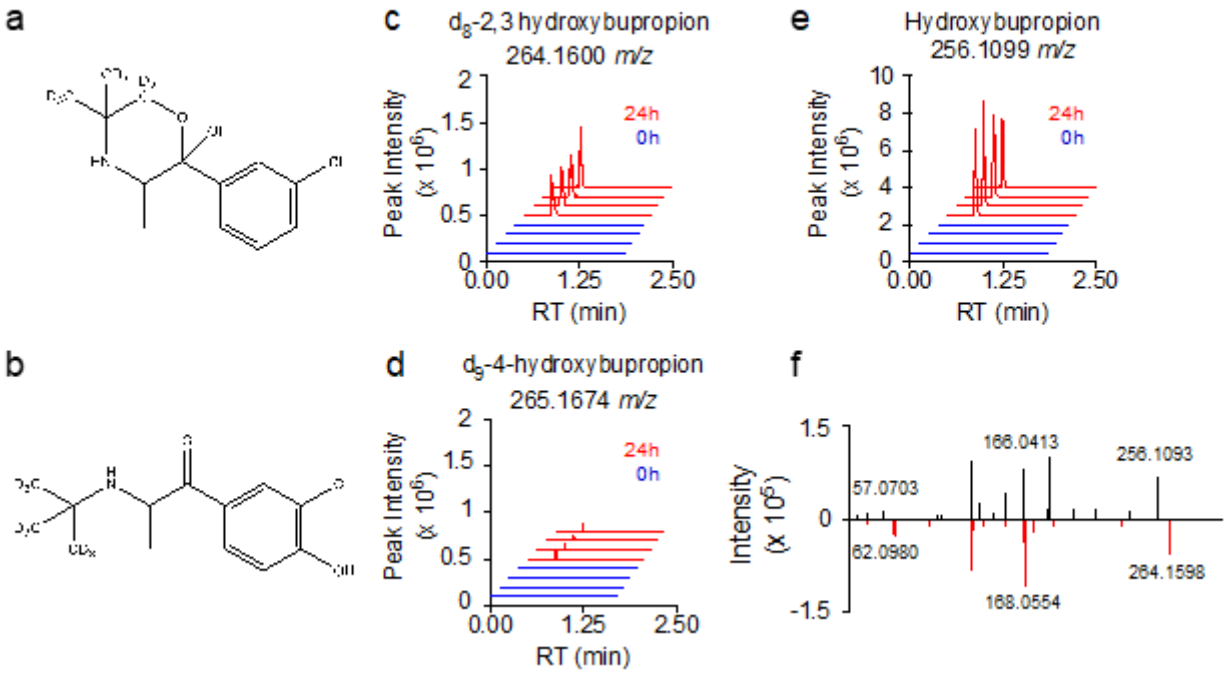

Figure 3

Stable-isotope assisted metabolite identification. Use of d9 bupropion with S9 enzyme system produces either a d8-2,3 hydroxybupropion or b d9-4 hydroxybupropion. c Extracted ion chromatogram of 264.1600 m/z (d8-2,3 hydroxybupropion) at 0 and 24 hours with addition of d9 bupropion. D) Extraction ion chromatogram of 265.1674 m/z (d9-4 hydroxybupropion) at 0 and 24 hours with addition of d9 bupropion. e Extracted ion chromatogram of 256.1099 m/z (hydroxybupropion) at 0 and 24 hours with addition of bupropion. f Ion dissociation spectra (MS/MS) of 256.1099 m/z (black) and 264.1600 m/z (red) is consistent with 2,3 hydroxybupropion.

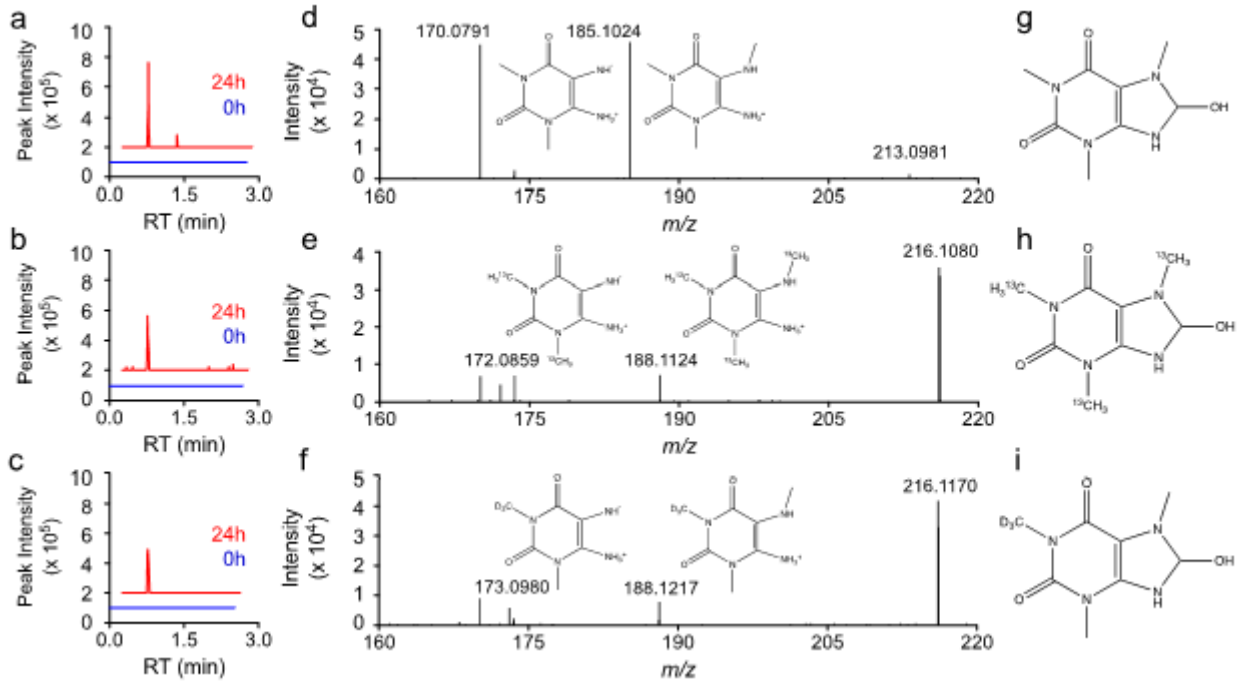

Figure 4

Stable isotope-assisted unexpected metabolite identification. a Time-dependent formation of 213.0981 m/z from caffeine. b Time-dependent formation of 216.1080 m/z from 13C3-caffeine. c Time-dependent formation of 216.1170 m/z from d3-caffeine. d MS/MS spectra of 213.0981 m/z. e MS/MS spectra of 216.1080 m/z. f MS/MS spectra of 216.1170 m/z. Proposed structures of unreported caffeine metabolite: g-i 213.0981 m/z, 216.1080 m/z, and 216.1170 m/z.

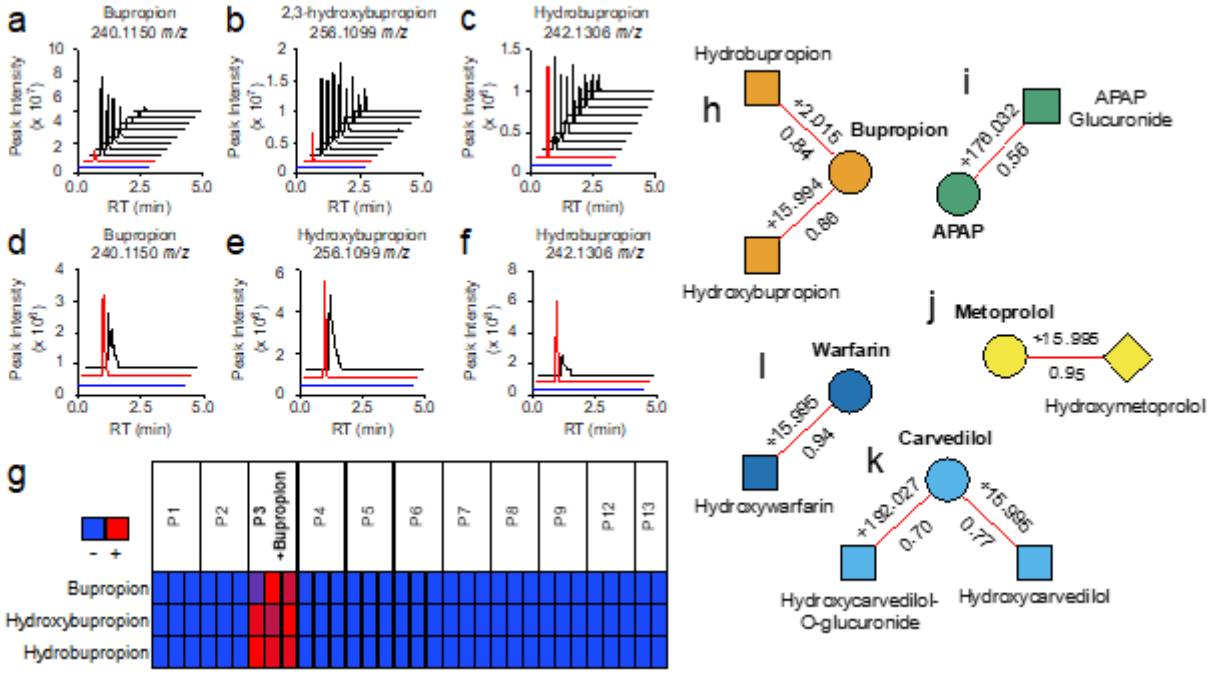

Figure 5

Identifying xenobiotics in samples with documented exposures. Extracted ion chromatograms of blood pharmacokinetic data from mice (black) that received bupropion shows a-c bupropion, hydroxybupropion, and hydrobupropion are detected at the same accurate mass m/z and retention time as S9-enzyme generated metabolites (red). These metabolites were not detected in a control sample (blue). Extracted ion chromatograms of plasma collected from an individual (black) taking bupropion shows d-f bupropion, hydroxybupropion, and hydrobupropion are detected at the same accurate mass m/z and retention time as enzyme-generated metabolites (red). g Co-occurrence of bupropion with its expected metabolites only in the samples with documented bupropion use. h-k Pathway-level biotransformation networking of parent xenobiotics shows expected metabolites with characteristic mass shifts (above red edge) are correlated (below red edge) with parent xenobiotic.

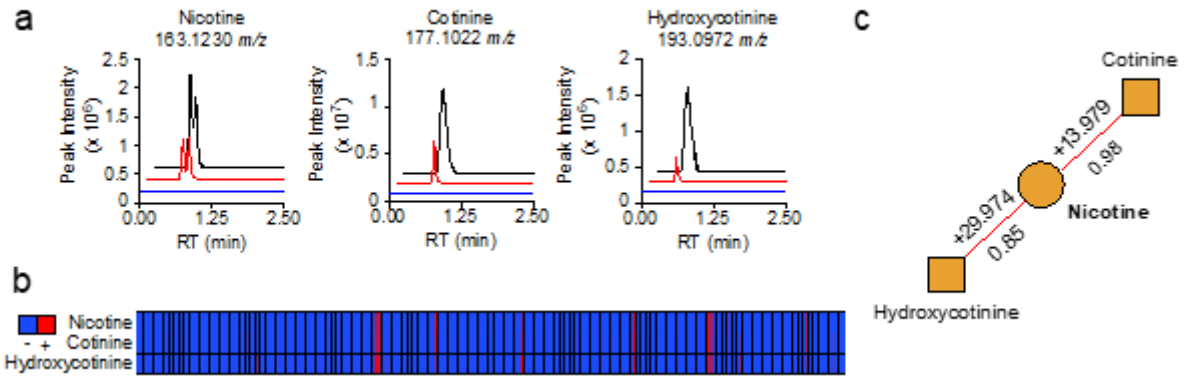

Figure 6

Identifying xenobiotics in samples from unknown exposures. a Nicotine, cotinine, and hydroxycotinine are detected at the same accurate mass m/z and retention time as S9-enzyme generated metabolites. b Nicotine is co-detected with cotinine and hydroxycotinine in 214 human samples. c Nicotine is correlated (below red edge) with cotinine and hydroxycotinine (R > 0.4) with characteristic mass shifts corresponding to known reactions (above red edge).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- U2Csupplemental.pdf
- U2CMetabolitetable.pdf