Exposome, 2022, **1(1)**, osac001
https://doi.org/10.1093/exposome/osac001
Commentary

Chemical contact tracing for exposomics

Ken Liu , Phd*

Clinical Biomarkers Laboratory, Emory University School of Medicine, Atlanta, GA, USA

*To whom correspondence should be addressed: Email: hkliu@emory.edu

Abstract

Human health and disease reflects a complex interplay between the genome and the exposome. High-resolution mass spectrometry (HRMS)-based metabolomics routinely measures thousands of endogenous, dietary and xenobiotic chemicals. However, confident identification of exposure-related chemicals remains a challenge as a significant portion of chemical signals detected in metabolomics analyses remains uncharacterized. Illuminating the "dark matter" of the exposome cannot be accomplished efficiently if the prevailing approach depends on the use of purified authentic standards that are not readily accessible for most laboratories. An alternative approach involves chemical exposure "contact tracing" analogous to contact tracing used to track the spread of infectious disease. For transmissible diseases, contact tracing identifies sets of potentially infected individuals that are linked by close contact to a confirmed positive case. Similarly, chemical exposures can be identified by establishing sets of xenobiotic metabolites that are linked to the original exposure via enzymatic biotransformation. Here, we provide a commentary on how incorporating enzyme-based strategies for chemical contact tracing enables -omics scale characterization of chemical exposures to further illuminate the "dark matter" of the exposome.

Keywords: mass spectrometry; exposomics; metabolomics; metabolism; exposure science

Introduction

We live in a chemical world, and chemicals live within us. Over the course of our lives, humans are intentionally or unintentionally exposed to chemical compounds in the spaces we inhabit, air we breathe, liquids we drink, foods we eat, drugs and supplements we take, and the products we use. Currently, over 350 000 compounds are documented within the chemical space of potential human exposures,[1] with approximately 40 000 compounds having been previously reported in human blood.[2] Following exposure, these chemicals can distribute throughout our bodies and be converted into different chemicals via enzymatic biotransformation, light or metal-catalyzed reactions or other spontaneous mechanisms, with both precursors and products potentially impacting cellular functions with implications for human health. Because the accumulation of complex, low-level environmental exposures over the lifespan contributes to human disease risk,[3] there is a need for analytical methods capable of delivering "-omics" scale information on environmental exposures.[4-6]

High-resolution mass spectrometry (HRMS)-based metabolomics, perhaps the most widely used platform for exposomics, offers the potential to measure tens of thousands of endogenous, microbial, and xenobiotic chemicals from single samples.[7-10] Routine HRMS analyses from human plasma capture over 20 000 detected chemical features, providing information on a scale comparable to the 20 000+ genes analyzed by genomics.[11] Individual chemical features may represent multiple ions corresponding to a single chemical.[12,13] However, workflows that automatically group features (isotopes and adducts) together may miss physiologically important information since sample-specific differences in alkali metal ion concentrations can affect metabolite adduct formation[14] and differences in carbon and/or nitrogen assimilation pathways can affect expected natural isotope distributions.[15] Despite the continual growth of computational tools and chemical databases to aid chemical annotation, the majority of detected chemical features are not confidently identified.[8] Current standard practice for chemical identification require multiple orthogonal properties (e.g. accurate mass *m/z*, retention time [RT], ion fragmentation spectra, and collisional cross section [CCS]) of an experimental peak of interest to match properties obtained from analysis of an authentic reference material, usually in the form of a purified chemical standard.[16,17] However, the limited availability of authentic standards and the inability to collect orthogonal information on low-level chemical exposures in real samples poses a real challenge for exposome epidemiology. For metabolomics to provide useful information on the "dark matter" of the exposome, alternative strategies for identifying low abundance chemical exposures are needed.

In principle, a "contact tracing" strategy for chemical exposures analogous to contact tracing strategies implemented for infectious disease could improve identification of low abundance unidentified chemical exposures for metabolomics. In the context of infectious disease, contact tracing establishes a network of potentially infected individuals that are linked by close contact to a confirmed positive case.[18] By establishing a point of origin (confirmed positive case) and potential points of transmission (close contacts), higher risk "sets" of individuals can be identified for mitigation purposes. Similarly, exposure contact tracing for metabolomics would rely on metabolism to define the network of

Received: September 28, 2021. **Revised:** December 14, 2021. **Accepted:** January 03, 2022
© The Author(s) 2022. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

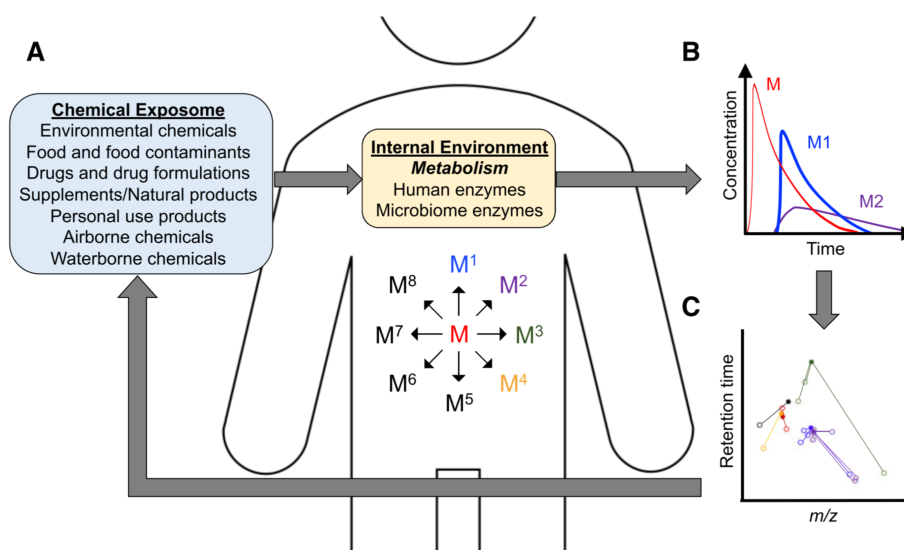

Figure 1. A theoretical basis for contact tracing chemical exposures. (**A**) Metabolism defines the internal exposure following exposure to environmental chemicals. A chemical that enters the body can be subjected to biotransformation by Phase I or II enzymes, as well as non-enzymatic reactions that may be catalyzed by light or metals. Thus, chemical contact tracing relies on identifying biotransformation products (or other non-enzymatic metabolites) that define the internal exposure to identify the original external environmental exposure. (**B**) Metabolism defines the sets of metabolites that can be detected at different time-points following exposure. Biotransformation enzymes catalyze Phase I and II metabolism to determine the internal exposure. (**C**) Reference xenobiotic metabolite sets can be used to identify relevant exposures in humans even if they are low abundance and cannot be identified by existing criteria. A xenobiotic metabolite set consists of precursor compounds (closed dots) that are linked to metabolites (open dots). For most chemical exposures, more than two related metabolites should co-occur in the same sample. Additional colors represent additional exposures and associated biotransformation products.

internal exposures (i.e. the set of downstream metabolites) that are linked to chemicals present in the original exposure (**Figure 1A**). Enzymes drive metabolic reactions and link external exposures to the internal environment. Over 130 human enzymes facilitate metabolism and biotransformation of xenobiotics.[19-22] These enzymes generally catalyze Phase I oxidation, reduction, hydrolysis, or Phase II conjugation reactions for bioactivation or detoxification. A literature review of the biotransformation of 1171 xenobiotic compounds concluded that xenobiotic metabolizing enzymes generate, on average, six metabolites per compound.[23] Recent studies of xenobiotic metabolism show that additional uncharacterized metabolites can be identified with HRMS, with some chemicals generating more than 25 metabolites.[24-27] Enzymatic biotransformation generates secondary, tertiary, and additional uncharacterized derivative internal exposures with different pharmacodynamic/toxicodynamic and pharmacokinetic/toxicokinetic properties compared with the original exposure (**Figure 1B**). Therefore, a sampling of the internal environment following a chemical exposure should reveal a subset of multiple co-occurring metabolites (arising from biotransformation, adducts, or other chemical transformations) that can be directly linked to the original xenobiotic exposure. Some chemicals and/or their metabolites may even persist in the body for weeks, months, or even years.[28] This is an important consideration for cross-sectional studies, where the timing of sample collection following acute or chronic exposures will determine the subset of co-occurring metabolites that can identify a chemical exposure. Including experimentally generated low abundance metabolites (which may only be characterized by its *m/z* and RT) in xenobiotic metabolite sets would allow corresponding low abundance and otherwise unidentified metabolites that co-occur with other known xenobiotic metabolites to be "contact traced" back to a particular chemical exposure (**Figure 1C**). It is important to note here that the goal of chemical contact tracing is not to identify every single xenobiotic metabolite in the sample, but rather to identify exposures based on co-detection of related metabolites characteristic of a particular exposure.

Our group recently introduced a framework for enzyme-based compound identification (EBI) for chemical exposure contact tracing.[29] EBI complements existing chemical identification workflows by incorporating precursor–product relationships inherent to metabolism as a criterion for metabolite identification. EBI is based on the curation of xenobiotic metabolite sets, which are defined by a xenobiotic precursor and their product metabolites, each with associated *m/z* and RT at a minimum, and ion fragmentation spectra and/or CCS values, if available. To some extent, xenobiotic metabolism by host and microbial enzymes can be predicted,[30] allowing hypothetical xenobiotic metabolite sets to be generated computationally, with each metabolite potentially having in silico predicted ion fragmentation spectra,[1,31,32] RTs,[33-35] and/or CCS values.[36,37] These computational methods facilitate the process of metabolite annotation, with annotations essentially serving as a "testable hypothesis" about the identity of a detected metabolite. Curating an experimentally validated xenobiotic metabolite set then would involve empirical evaluation of precursor xenobiotic compounds (defined as reference materials) and their enzymatically synthesized metabolites (biological derivatives of a reference material) to obtain accurate mass *m/z*, RTs, ion fragmentation spectra (if possible), or CCS values (if possible) for xenobiotics and their associated metabolites. Incorporating these criteria into MS-based chemical identification workflows have been discussed previously[38] and implementation will require large-scale curation of experimentally validated xenobiotic metabolism pathways[39,40] and assembly of this information into publicly accessible data formats.

To accomplish this on a large-scale, we developed a practical method to curate experimentally validated sets of xenobiotic

metabolites for EBI purposes.[29] Because liver metabolic activity is a major determinant of the set of xenobiotic biotransformation products generated in vivo following exposures, we incubated xenobiotic compounds with pooled human liver S9 enzymes to generate xenobiotic metabolites in vitro for HRMS characterization. Pooled human liver S9 fractions contain cytochrome P450 (P450), flavin monooxygenase (FMO), aldehyde oxidase, xanthine oxidase, uridine glucuronosyltransferase (UGT), sulfotransferase (SULT), *N*-acetyltransferase (NAT), methyltransferase, and glutathione transferase (GST) enzymes, providing Phase I and II biotransformation activities similar to human hepatocytes.[41] By transforming a single xenobiotic compound into a set of enzymatically generated xenobiotic metabolites, the analysis of a single chemical standard is elevated into a mixtures analysis of compounds representative of the internal exposure. HRMS analysis of S9 reaction extracts showed that S9 enzymes produced expected and also previously uncharacterized metabolites from diverse xenobiotics. Accurate mass *m/z*, RTs, and ion fragmentation spectra (for higher abundance metabolites) for each set of xenobiotic metabolites (includes *m/z* and RT for low abundance metabolites) were used to identify corresponding undocumented diet, drug, and environmental chemical exposures from human samples. For example, we detected unique patterns of chemical exposures in people, with multiple co-occurring Phase I or II xenobiotic metabolites corresponding to drug use, naphthalene exposure, diet-derived piperine (from black pepper), or other xenobiotics found in human plasma and urine samples. Some exposures, like piperine, were detected in more than 80% of samples, and others, like naphthalene, were detected in less than 10% of samples. These metabolites were detected in human samples at their characteristic accurate mass *m/z* and RT values as generated from analysis of S9-enzyme generated xenobiotic metabolites. Additionally, at least one xenobiotic or xenobiotic metabolite from each detected xenobiotic metabolite set was confirmed from matching ion fragmentation spectra and RTs from experimental samples to those obtained from analysis of S9 reaction mixtures. In principle, our work shows that an EBI workflow can be used to contact trace chemical exposures for mass-spectrometry-based exposomics. However, statistical models to establish identification probabilities (for exposures) will need to be evaluated using real world data. Because the complete set of downstream metabolites arising from an exposure may not be detectable in a given experimental sample, an empirical probability for the expected number of co-occurring metabolites within each xenobiotic metabolite set could be employed. This could be accomplished by establishing population means (±standard deviation) for the number of biotransformation products detected per confirmed xenobiotic exposure in large-scale human metabolomics studies.

Unannotated chemical features that are likely to be identifiable should be biochemically related (linked by a one or two step reaction) away from a known compound. Therefore, curating experimentally validated xenobiotic metabolite sets not only increases the number of metabolites that can be validated from a single starting compound (accelerating coverage of chemical space), but also allows for identification of low abundance exposures that cannot be otherwise identified using widely adopted chemical identification criteria. For example, our S9 incubations with caffeine showed production of a metabolite with accurate mass 213.0981 *m/z* ($C_8H_{12}N_4O_3$) which corresponds to an unexpected hydration product of caffeine. Although the structure could not be fully elucidated using mass spectrometry alone, we saw that 213.0981 *m/z* was detected in mice following oral gavage of caffeine, and that 213.0981 *m/z* was correlated with caffeine in human plasma, establishing multiple lines of evidence to identify 213.0981 *m/z* as a caffeine-related metabolite in humans. While earlier research on caffeine metabolism[42] shows that 213.0981 *m/z* could be 6-amino-5-(formyl-*N*-methylamino)-1,3-dimethyluracil, this was not a predicted metabolite of caffeine using Biotransformer,[30] and also not present in common metabolite databases as a caffeine-associated metabolite. There are likely many more unexpected (unpredicted or metabolites that are not present in chemical databases) or otherwise uncharacterized metabolites that can be identified by this contact-tracing approach, since biotransformation studies of thousands of xenobiotics will likely yield additionally thousands of unexpected and/or uncharacterized metabolites that may be detectable in the human exposome.

In principle, use of enzyme-based identification will enhance interoperability of metabolomics data. Because enzymes catalyze discrete modifications with characteristic mass shifts, these precursor–product relationships are preserved across different analytical platforms. For example, we are able to detect acetaminophen ($C_8H_9NO_2$) and its glucuronide ($C_{14}H_{17}NO_8$) using both HILIC chromatography with positive electrospray ionization and C18 chromatography with negative electrospray ionization with each pair of detected ions differing by the mass of a glucuronide ($C_6H_8O_6$), 176.0321 *m/z*.[29] While the RTs for these two metabolites were different from one another on each platform, the RTs on both columns for acetaminophen and its glucuronide in human samples matched those observed from analysis of S9 enzyme incubations with acetaminophen. Employing RT shifts with mass defect filtering[43] could also be applied to link precursor–product pairs in HRMS data.

To facilitate community use, open format mass spectrometry data are available at Metabolomics Workbench[44] and curated xenobiotic metabolite sets are published online at metabolomics.cloud.[29] Certainly, there remain substantial challenges to harmonizing exposomics data collected across multiple laboratories across multiple methods. Pre-analytical (sample storage, processing, and preparation) factors influence metabolite stability and recovery (the chemical space that is sampled) prior to instrumental analysis. How the data are collected (instrumentation [mass analyzer, e.g. Time-of-flight, Orbitrap, Fourier Transform and Quadrupole], separation mechanisms [chromatography and ion mobility], ion source parameters [mass ranges, ionization mechanism, polarity and fragmentation energy]) influences sensitivity and selectivity of metabolite detection, as well as metabolite quantification. Furthermore, several post-analytical data processing workflows for peak extraction, alignment, formula assignment, and annotation can be used. In principle, common analyses of shared samples across different laboratories (using different methodologies) should identify similar sets of exposures. However, because some exposures generate common metabolites (e.g. benzoic acid could arise from dietary benzoic acid or biotransformation of toluene) or isobaric xenobiotics being converted into common masses (e.g. alpha-naphthoflavone and beta-naphthoflavone), establishing the prevalence of a particular exposure and its distribution within a population could enable probability-based ranking of exposures. Alternatively, some xenobiotic exposures may require specific targeted approaches that capture additional information (MS^n, appropriate chromatographic conditions, ion mobility) to achieve adequate selectivity to confidently identify an exposure. Nonetheless, continued curation of xenobiotic sets and integration of these for broader use with pathway enrichment analysis

tools[45,46] or the GNPS infrastructure[47-49] will help to facilitate interoperability between laboratories within a framework for community-based exposure assessments.

Chemical contact tracing, with each metabolite directing focus back to a chemical exposure, provides an important step forward for exposome research. Consequently, adopting enzyme-based identification concepts as outlined here should enhance capabilities to link cumulative human chemical exposures to health outcomes. As additional in vitro systems for generating metabolites from lung, skin, kidney, intestinal, or even microbial enzymes are developed, tissue-specific metabolite sets can be curated to further characterize internal exposures. Furthermore, coupling enzymatic biotransformation to large-scale analysis of complex environmental mixtures would allow curation of sets of chemical co-exposures and associated metabolites to define metabolic signatures of complex human exposures. In principle, larger quantities of metabolites derived from any particular exposure can be generated for isolation and characterization with NMR and bioactivity screens. Ultimately, as the field moves to adopt chemical contact tracing and bioactivity-guided exposomics,[50] potential causal environmental determinants of the human condition can be identified, even if they are currently considered “dark matter.”

Acknowledgments

The author acknowledges helpful discussions from Professors Dean P. Jones (Emory University), Edward T. Morgan (Emory University), and Shuzhao Li (Jackson Laboratories) as well as grant support from National Institutes of Health U2C-ES030163.

Conflict of interest statement

The author declares no conflicts of interest.

References

1. Schymanski EL, Kondić T, Neumann S, *et al.* Empowering large chemical knowledge bases for exposomics: PubChemLite meets MetFrag. J Cheminform. 2021;13(1):19. doi:10.1186/s13321-021-00489-0
2. Barupal DK, Fiehn O. Generating the blood exposome database using a comprehensive text mining and database fusion approach. Environ Health Perspect. 2019;127(9):97008.doi:10.1289/EHP4713
3. Rappaport SM, Smith MT. Epidemiology. Environment and disease risks. Science. 2010;330(6003):460–461. doi:10.1126/science.1192603
4. Jones DP. Sequencing the exposome: A call to action. Toxicol Rep. 2016;3:29–45. doi:10.1016/j.toxrep.2015.11.009
5. Jones DP, Cohn BA. A vision for exposome epidemiology: The pregnancy exposome in relation to breast cancer in the Child Health and Development Studies. Reprod Toxicol. 2020;92:4–10. doi:10.1016/j.reprotox.2020.03.006
6. Vermeulen R, Schymanski Emma L, Barabási A-L, Miller Gary W. The exposome and health: Where chemistry meets biology. Science. 2020;367(6476):392–396. doi:10.1126/science.aay3164
7. Johnson JM, Yu T, Strobel FH, Jones DP. A practical approach to detect unique metabolic patterns for personalized medicine. Analyst. 2010;135(11):2864–2870. doi:10.1039/c0an00333f
8. Liu KH, Walker DI, Uppal K, *et al.* High-resolution metabolomics assessment of military personnel: Evaluating analytical strategies for chemical detection. J Occup Environ Med. 2016;58(8 Suppl 1):S53–61. doi:10.1097/JOM.0000000000000773
9. Go Y-M, Walker DI, Liang Y, *et al.* Reference standardization for mass spectrometry and high-resolution metabolomics applications to exposome research. Toxicol Sci. 2015;148(2):531–543. doi:10.1093/toxsci/kfv198
10. Liu KH, Nellis M, Uppal K, *et al.* Reference standardization for quantification and harmonization of large-scale metabolomics. Anal Chem. 2020;92(13):8836–8844. doi:10.1021/acs.analchem.0c00338
11. Uppal K, Walker DI, Liu K, *et al.* Computational metabolomics: A framework for the million metabolome. Chem Res Toxicol. 2016;29(12):1956–1975. doi:10.1021/acs.chemrestox.6b00179
12. Mahieu NG, Spalding JL, Gelman SJ, Patti GJ. Defining and detecting complex peak relationships in mass spectral data: The Mz.unity algorithm. Anal Chem. 2016;88(18):9037–9046. doi:10.1021/acs.analchem.6b01702
13. Schmid R, Petras D, Nothias L-F, *et al.* Ion identity molecular networking for mass spectrometry-based metabolomics in the GNPS environment. Nat Commun. 2021;12(1):3832. doi:10.1038/s41467-021-23953-9
14. Erngren I, Haglöf J, Engskog MKR, *et al.* Adduct formation in electrospray ionisation–mass spectrometry with hydrophilic interaction liquid chromatography is strongly affected by the inorganic ion concentration of the samples. J Chromatogr A. 2019;1600:174–182. doi:10.1016/j.chroma.2019.04.049
15. Huneau J-F, Mantha OL, Hermier D, *et al.* Natural isotope abundances of carbon and nitrogen in tissue proteins and amino acids as biomarkers of the decreased carbohydrate oxidation and increased amino acid oxidation induced by caloric restriction under a maintained protein intake in obese rats. Nutrients 2019;11(5):1087. doi:10.3390/nu1105
16. Sumner LW, Amberg A, Barrett D, *et al.* Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics. 2007;3(3):211–221. doi:10.1007/s11306-007-0082-2
17. Schymanski EL, Jeon J, Gulde R, *et al.* Identifying small molecules via high resolution mass spectrometry: Communicating confidence. Environ Sci Technol. 2014;48(4):2097–2098. doi:10.1021/es5002105
18. Eames KTD, Keeling MJ. Contact tracing and disease control. Proc Biol Sci. 2003;270(1533):2565–2571. doi:10.1098/rspb.2003.2554
19. Omiecinski CJ, Vanden Heuvel JP, Perdew GH, Peters JM. Xenobiotic metabolism, disposition, and regulation by receptors: From biochemical phenomenon to predictors of major toxicities. Toxicol Sci. 2011;120(Suppl. 1):S49–S75. doi:10.1093/toxsci/kfq338
20. Penner N, Woodward C, Prakash C. In: ed Surapaneni S, Zhang D, eds. ADME-Enabling Technologies in Drug Design and Development. Chap. Appendix. Wiley; 2012:545–65.
21. Guengerich FP. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. Chem Res Toxicol. 2001;14(6):611–650. doi:10.1021/tx0002583
22. McShan DC, Updadhayaya M, Shah I. Symbolic inference of xenobiotic metabolism. Pac Symp Biocomput. 2004:545–556. doi:10.1142/9789812704856_0051
23. Testa B, Pedretti A, Vistoli G. Reactions and enzymes in the metabolism of drugs and other xenobiotics. Drug Discov Today. 2012;17(11–12):549–560. doi:10.1016/j.drudis.2012.01.017
24. Kim JH, Choi WG, Lee S, Lee HS. Revisiting the metabolism and bioactivation of ketoconazole in human and mouse using liquid

chromatography–mass spectrometry-based metabolomics. IJMS. 2017;18(3):621.doi:10.3390/ijms18030621
25. David A, Chaker J, Léger T, *et al.* Acetaminophen metabolism revisited using non-targeted analyses: Implications for human biomonitoring. Environ Int. 2021;149:106388. doi:10.1016/j.envint.2021.106388
26. Le Dare B, Ferron P-J, Allard P-M, Clement B, Morel I, Gicquel T. New insights into quetiapine metabolism using molecular networking. Sci Rep. 2020;10:19921. doi:10.1038/s41598-020-77106-x (2020)
27. Manier SK, Keller A, Schaper J, Meyer MR. Untargeted metabolomics by high resolution mass spectrometry coupled to normal and reversed phase liquid chromatography as a tool to study the in vitro biotransformation of new psychoactive substances. Sci Rep. 2019;9(1):2741.doi:10.1038/s41598-019-39235-w
28. Lee Y-M, Kim K-S, Jacobs DR Jr, Lee D-H. Persistent organic pollutants in adipose tissue should be considered in obesity research. Obes Rev. 2017;18(2):129–139. doi: https://doi.org/10.1111/obr.12481
29. Liu KH, Lee CM, Singer G, *et al.* Large scale enzyme based xenobiotic identification for exposomics. Nat Commun. 2021;12(1):5418. doi:10.1038/s41467-021-25698-x
30. Djoumbou-Feunang Y, Fiamoncini J, Gil-de-la-Fuente A, *et al.* BioTransformer: A comprehensive computational tool for small molecule metabolism prediction and metabolite identification. J Cheminform. 2019;11(1):2. doi:10.1186/s13321-018-0324-5
31. 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(Web Server issue):W94–W99. doi:10.1093/nar/gku436
32. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. MetFrag relaunched: Incorporating strategies beyond in silico fragmentation. J Cheminform. 2016;8:3–3. doi:10.1186/s13321-016-0115-9
33. Bonini P, Kind T, Tsugawa H, Barupal DK, Fiehn O. Retip: Retention time prediction for compound annotation in untargeted metabolomics. Anal Chem. 2020;92(11):7515–7522. doi:10.1021/acs.analchem.9b05765
34. Stanstrup J, Neumann S, Vrhovšek U. PredRet: Prediction of retention time by direct mapping between multiple chromatographic systems. Anal Chem. 2015;87(18):9421–9428. doi:10.1021/acs.analchem.5b02287
35. Low DY, Micheau P, Koistinen VM, *et al.* Data sharing in PredRet for accurate prediction of retention time: Application to plant food bioactive compounds. Food Chem. 2021;357:129757. doi:https://doi.org/10.1016/j.foodchem.2021.129757
36. Meier F, Köhler ND, Brunner A-D, *et al.* Deep learning the collisional cross sections of the peptide universe from a million experimental values. Nat Commun. 2021;12(1):1185. doi:10.1038/s41467-021-21352-8
37. Broeckling CD, Yao L, Isaac G, *et al.* Application of predicted collisional cross section to metabolome databases to probabilistically describe the current and future ion mobility mass spectrometry. J Am Soc Mass Spectrom. 2021;32(3):661–669. doi:10.1021/jasms.0c00375
38. Rochat B. Proposed confidence scale and ID score in the identification of known–unknown compounds using high resolution MS data. J Am Soc Mass Spectrom. 2017;28(4):709–723. doi:10.1007/s13361-016-1556-0
39. Baduel C, Lai FY, van Nuijs ALN, Covaci A. Suspect and nontargeted strategies to investigate in vitro human biotransformation products of emerging environmental contaminants: The benzotriazoles. Environ Sci Technol. 2019;53(17):10462–10469. doi:10.1021/acs.est.9b02429
40. Huber C, Müller E, Schulze T, Brack W, Krauss M. Improving the screening analysis of pesticide metabolites in human biomonitoring by combining high-throughput in vitro incubation and automated LC–HRMS data processing. Anal Chem. 2021;93(26):9149–9157. doi:10.1021/acs.analchem.1c00972
41. Richardson SJ, Bai A, Kulkarni AA, Moghaddam MF. Efficiency in drug discovery: Liver S9 fraction assay as a screen for metabolic stability. Drug Metab Lett. 2016;10(2):83–90. doi:10.2174/1872312810666160223121836
42. Blanchard J, Sawers SJ, Jonkman JH, Tang-Liu DD. Comparison of the urinary metabolite profile of caffeine in young and elderly males. Br J Clin Pharmacol. 1985;19(2):225–232. doi:10.1111/j.1365-2125.1985.tb02635.x
43. Zhang H, Zhang D, Ray K, Zhu M. Mass defect filter technique and its applications to drug metabolite identification by high-resolution mass spectrometry. J Mass Spectrom. 2009;44(7):999–1016. doi:10.1002/jms.1610
44. Sud M, Fahy E, Cotter D, *et al.* Metabolomics workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. Nucleic Acids Res. 2016;44(D1):D463–D470. doi:10.1093/nar/gkv1042
45. Xia J, Wishart DS. MSEA: A web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. Nucleic Acids Res. 2010;38(Web Server issue):W71–77. doi:10.1093/nar/gkq329
46. Li S, Park Y, Duraisingham S, *et al.* Predicting network activity from high throughput metabolomics. PLoS Comput Biol. 2013;9(7):e1003123. doi:10.1371/journal.pcbi.1003123
47. Wang M, Carver JJ, Phelan VV, *et al.* Sharing and community curation of mass spectrometry data with global natural products social molecular networking. Nat Biotechnol. 2016;34(8):828–837. doi:10.1038/nbt.3597
48. Quinn RA, Nothias L-F, Vining O, *et al.* Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy. Trends Pharmacol Sci. 2017;38(2):143–154. doi:https://doi.org/10.1016/j.tips.2016.10.011
49. Aron AT, Gentry EC, McPhail KL, *et al.* Reproducible molecular networking of untargeted mass spectrometry data using GNPS. Nat Protoc. 2020;15(6):1954–1991. doi:10.1038/s41596-020-0317-5
50. Chung MK, Rappaport SM, Wheelock CE, *et al.* Utilizing a biology-driven approach to map the exposome in health and disease: An essential investment to drive the next generation of environmental discovery. Environ Health Perspect. 2021;129(8):85001. doi:10.1289/EHP8327