



Open, High-Resolution EI+ Spectral Library of Anthropogenic Compounds

Elliott J. Price^{1,2*}, Jirí Palát², Katerina Coufaliková², Petr Kukučka², Garry Codling², Chiara Maria Vitale², Štěpán Koudelka² and Jana Klánová²

¹ Faculty of Sports Studies, Masaryk University, Brno, Czechia, ² RECETOX Centre, Masaryk University, Brno, Czechia

OPEN ACCESS

Edited by:

Benedikt Warth,
University of Vienna, Austria

Reviewed by:

Biswapriya Biswas Misra,
Independent Researcher,
Visakhapatnam, India
Karl Jobst,
Memorial University of
Newfoundland, Canada

*Correspondence:

Elliott J. Price
elliott.james.price@gmail.com

Specialty section:

This article was submitted to
Environmental Health and Exposome,
a section of the journal
Frontiers in Public Health

Received: 28 October 2020

Accepted: 08 February 2021

Published: 09 March 2021

Citation:

Price EJ, Palát J, Coufaliková K,
Kukučka P, Codling G, Vitale CM,
Koudelka Š and Klánová J (2021)
Open, High-Resolution EI+ Spectral
Library of Anthropogenic Compounds.
Front. Public Health 9:622558.
doi: 10.3389/fpubh.2021.622558

To address the lack of high-resolution electron ionisation mass spectral libraries (HR-[EI+]-MS) for environmental chemicals, a retention-indexed HR-[EI+]-MS library has been constructed following analysis of authentic compounds via GC-Orbitrap MS. The library is freely provided alongside a compound database of predicted physicochemical properties. Currently, the library contains over 350 compounds from 56 compound classes and includes a range of legacy and emerging contaminants. The RECETOX Exposome HR-[EI+]-MS library expands the number of freely available resources for use in full-scan chemical exposure studies and is available at: <https://doi.org/10.5281/zenodo.4471217>.

Keywords: electron ionization [EI+], spectral library, gas chromatography mass spectrometry, chemical exposure, high-resolution

INTRODUCTION

Since commercial release in 2015, the high-resolution gas chromatography Orbitrap mass spectrometer (GC Orbitrap MS) has been evidenced as a valuable tool in metabolomics (1–6), environmental (7–10), clinical (11), and forensic analysis (12). The enhanced mass accuracy and greater achievable linear dynamic range (13), when operating in full-scan mode, makes GC Orbitrap MS particularly suited to chemical characterization of complex samples with unknown composition.

The most commonly applied ionization for screening of environmental contaminants is electron ionization (EI+), typically operated at 70 electron volts (eV), favored for robust fragmentation. When coupled with retention index information, the matching of EI+ spectra can enable structural annotation of relatively high confidence (14).

However, high-resolution electron ionization mass spectral (HR-[EI+]-MS) libraries are currently limited (1), particularly for environmental chemicals. This hinders the application of GC Orbitrap MS without prior generation of in-house spectral libraries, which requires substantial resources; or the purchase of commercial libraries that are often tied to proprietary data formats and software.

Whilst matching to low resolution (LR) spectra is possible and additional accurate mass information utilized [e.g., via High Resolution Filtering (15) (HRF)], freely available LR-[EI+]-MS libraries are equally limited in coverage of environmental chemicals. Furthermore, scan time has a significant impact of fidelity of isotopic abundance (16) and specific chemical gas-phase reactions in the trap (17) can lead to Orbitrap system-specific spectra. In addition, it is known that spectra are source-dependent, even at standardized 70 eV (18). These additional information are needed to

be overlooked when matching to LR-[EI+]-MS spectra and entail discrepancy to current spectral predictions and substructure characterizations.

Accurate spectral prediction is particularly crucial for the generation of “suspect” libraries for GC-[EI+]-MS screening of environmental contaminants, to improve identification of unknowns (19). The lack of [EI+]-MS spectra for environmental chemicals is a constraint for spectral prediction (20) by limiting inputs for machine-learning methods and preventing validation of computed spectra.

Herein, the RECETOX Exposome HR-[EI+]-MS library has been generated for free distribution to enable accelerated application of GC-Orbitrap MS for identification of environmental contaminants.

MATERIALS AND METHODS

Chemicals

All reagents were of GC grade (for pesticide residue analysis) or higher. Standards were of $\geq 98\%$ purity and stored as per manufacturer recommendations. Compounds were selected on the basis of being included in targeted environmental and biomonitoring analysis undertaken or under method development by the RECETOX Trace Analytical Laboratories (under EN ISO/IEC 17025:2005 accreditation), thus with known amenability for GC-[EI+]-MS. Standards were purchased in solution form and dilutions conducted following accredited trace analytical laboratory practice. Where necessary, solvent was switched to pyridine or hexane, under high purity N_2 . Individual aliquots of 50 μL were transferred into 2 mL amber vials with built-in 350 μL insert and stored at -20°C prior to injection.

Data Acquisition

Compounds or compound mixes were analyzed via GC Orbitrap MS comprising a Trace 1310 Series GC, Q Exactive GC-Orbitrap MS and TriPlus RSH Autosampler. Injections (1–2 μL , providing >100 pg on column per analyte) were made in splitless mode using split/splitless injector. Separation was performed on a 5-type MS column (30 m \times 0.25 mm, 0.25 μm i.d.; cross-linked 5% phenyl-95% methylpolysiloxane, Restek Rxi-5Sil MS) with guard (1 m \times 0.53 mm i.d.; non-polar deactivated fused silica, Restek Rxi guard) with helium as carrier gas (1.3 mL/min). The Orbitrap MS was operated in Full MS-SIM using 70 eV EI+ and data recorded in profile mode, scan range 70–700 m/z. Filament emission was 50 μA , MS transfer line at 250°C , and ion source at 280°C . Resolving power was 60,000 full-width at half maximum height at m/z 200, automatic gain control at $1\text{E}6$ and automatic max injection time. A C_7 - C_{40} alkane series was used for external non-isothermal Kováts retention-indexing (from temperature programming, using the definition of Van den Dool and Kratz) (21).

Library Construction

Vendor raw files were converted to mzML format using ProteoWizard MSConvert (22, 23) (ver 3) with vendor centroiding. Component peak identification and spectral deconvolution was performed using MS-DIAL (24, 25) (ver 4.20). Parameters were set as follows: minimum peak height:

50,000; mass slice width: 0.05; mass centroiding accuracy: 0.05; average peak width: 10; smoothing level: 3; sigma window: 0.3 and 1% spectra cut-off. Quality of deconvoluted spectra was manually checked (26) and acceptable spectra exported to MS-FINDER (25, 27) (ver 3.42). Precursor m/z was assigned as nearest ion in the spectra equal to or less than the compounds monoisotopic mass and fragments were annotated with a (5 ppm) tolerance. Spectra were saved in the MS Transfer File (MSP) format with a 1% relative abundance cut-off (28). Retention indices (RI) were retrieved from MS-DIAL and input to the MSP. Where possible, spectra were verified via similarity matching (forward search) (29) against LR-[EI+]-MS spectra of a composite library comprising NIST/EPA/NIH MS Library (NIST 14) (30), MS-DIAL MSP spectra kit of public EI-MS spectra (25, 31) (ver 2), SWGDRUG MS library (32, 33) (ver 3.6), Cayman Spectral Library (34) (v09112019), and Golm Metabolome Database (35) (v20112021). RIs were compared to consensus semi non-polar RIs (36). Spectral and RI matches were conducted via NIST MS Search (ver 2.3) (37), constrained to the 70–700 m/z scan range.

Database Management

Compound identifiers (InChI, InChIKey & SMILES) were retrieved via the Chemical Translation Service (38) (chemical name as input), United States Environmental Protection Agency (EPA) CompTox Chemicals Dashboard (39) (ver 3.5, chemical name and/or InChI as input) or generated in ACD/Chemsketch (40) (manually drawn structure). Predicted physico- and toxico-chemical properties were retrieved from the EPA CompTox Chemicals Dashboard (39) (ver 3.5, InChIKey as input); or generated via DataWarrior (41) (ver 5.2.1, SMILES as input). Natural product likeness scores were calculated via NP-Scout (42) on the NERDD portal (43) (SMILES as input). Structural classification was calculated via ClassyFire (44) (SMILES as input). Distribution plots were generated using plotly online (45) (available at <https://chart-studio.plotly.com/>). The database was compiled and exported in structure data formation (SDF) through DataWarrior (41).

RESULTS AND CONCLUSIONS

Authentic compounds have been analyzed in full-scan mode using GC-Orbitrap MS and the constructed RECETOX

TABLE 1 | Overview of current diversity of compounds included within HR-[EI+]-MS library.

Number of spectra	386
Unique compounds ^a (connectivity)	351
Number of chemical classes ^b	56
Monoisotopic mass range	108.06–715.45
boiling point ($^\circ\text{C}$) range ^c	170.51–549.62
Octanol water coefficient ($\log P$) range ^c	–0.84–9.20

^aCompounds with unique connectivity (46).

^bClassyFire taxonomy (44).

^cPredicted values retrieved from the EPA CompTox Chemicals Dashboard (39).

Exposome HR-[EI+]-MS library incorporates GC retention-index (alkane series, semi non-polar column) and theoretical fragment formula annotation.

The library contains compounds of broad physicochemical diversity (Table 1, Supplementary Figure 1) and toxicological importance. Of the 386 spectra collected, 336 are unique to the RECETOX Exposome HR-[EI+]-MS library with respect to the 31,491 contained in other freely available libraries (composite library excluding NIST 14; Supplementary Table 1). Notably, the majority of compounds (318 of 352) are listed on the Human Biomonitoring for Europe (HBM4EU) Screening List for Chemical of Emerging Concern (CECscreen) (47) (Supplementary Table 1).

The MSP format is widely used, readable and modifiable by commercial and freely available software tools (48), enabling easy incorporation into current annotation workflows. Spectral quality was ensured and comparison of the HR-[EI+]-MS entries to LR-[EI+]-MS libraries generated an average forward match score of 841 (Supplementary Table 1). In use, adequate spectral matches to the HR-[EI+]-MS library enhanced with RI match on similar 5-type semi-non polar columns (49) would warrant a level 2 “putative” annotation (50, 51) (example in Supplementary Figure 2). Furthermore, the high degree of compound diversity is beneficial for integration with EI+ spectral similarity networking via GNPS-MSHub (52) online workflows or offline via MetGEM (53) to assign compound class (level 3 annotation) (54).

The accompanying SDF database accompanies structures with structural classifiers and predicted physico- and toxico-chemical properties. Sharing facilitates ease of insight into chemical properties (example in Supplementary Figure 3) and future use of compound data for modeling, e.g., retention prediction (55).

The RECETOX Exposome HR-[EI+]-MS library is freely provided to enable broad usability and promotes open science in environmental research (56). We hope the RECETOX Exposome HR-[EI+]-MS library provides a valuable resource for those seeking to screen environmental exposures and chemical contaminants.

REFERENCES

- Stettin D, Poulin RX, Pohnert G. Metabolomics benefits from orbitrap GC-MS—comparison of low- and high-resolution GC-MS. *Metabolites*. (2020) 10:143. doi: 10.3390/metabo10040143
- Misra BB, Olivier M. High resolution GC-orbitrap-MS metabolomics using both electron ionization and chemical ionization for analysis of human plasma. *J Proteome Res*. (2020) 19:2717–31. doi: 10.1021/acs.jproteome.9b00774
- Qiu Y, Moir RD, Willis IM, Seethapathy S, Biniakewitz RC, Kurland JJ. Enhanced isotopic ratio outlier analysis (IROA) peak detection and identification with ultra-high resolution GC-orbitrap/MS: potential application for investigation of model organism metabolomes. *Metabolites*. (2018) 8:9. doi: 10.3390/metabo8010009
- Miller IJ, Peters SR, Overmyer KA, Paulson BR, Westphall MS, Coon JJ. Real-time health monitoring through urine metabolomics. *NPJ Digit Med*. (2019) 2:109. doi: 10.1038/s41746-019-0185-y
- Weidt S, Haggarty J, Kean R, Cojocariu CI, Silcock PJ, Rajendran R, et al. A novel targeted/untargeted GC-Orbitrap metabolomics methodology applied to *Candida albicans* and *Staphylococcus aureus* biofilms. *Metabolomics*. (2016) 12:1–10. doi: 10.1007/s11306-016-1134-2
- Shen S, Li L, Song S, Bai Y, Liu H. Metabolomic study of mouse embryonic fibroblast cells in response to autophagy based on high resolution gas chromatography–mass spectrometry. *Int J Mass Spectrom*. (2018) 434:215–21. doi: 10.1016/j.ijms.2018.09.010
- Tienstra M, Mol HGJ. Application of gas chromatography coupled to quadrupole-orbitrap mass spectrometry for pesticide residue analysis in cereals and feed ingredients. *J AOAC Int*. (2018) 101:342–51. doi: 10.5740/jaoacint.17-0408
- Postigo C, Cojocariu CI, Richardson SD, Silcock PJ, Barcelo D. Characterization of iodinated disinfection by-products in chlorinated and chloraminated waters using Orbitrap based gas chromatography-mass spectrometry. *Anal Bioanal Chem*. (2016) 408:3401–11. doi: 10.1007/s00216-016-9435-x

DATA AVAILABILITY STATEMENT

The RECETOX Exposome HR-[EI+]-MS library and accompanying compound database are available for download at: <https://doi.org/10.5281/zenodo.4471217>.

AUTHOR CONTRIBUTIONS

EJP devised the concept, generated the individual spectral files, compiled the spectral library and compound database, drafted the original, undertook editing, and review of the manuscript. JP and KC generated the individual spectral files, curated the spectral library, and reviewed the manuscript. PK and GC oversaw the chemical management, generated the individual spectral files, and reviewed the manuscript. CV and ŠK curated the compound database and reviewed the manuscript. JK secured the funding, edited, and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

EJP acknowledges support from the Czech Operational Programme Research, Development and Education—Project Postdoc@MUNI (CZ.02.2.69/0.0/0.0/16_027/0008360). All authors acknowledge the RECETOX research infrastructure supported by the Ministry of Education, Youth and Sports of the Czech Republic (LM2018121) and funding from the Ministry of Education, Youth and Sports of the Czech Republic—Project Cetocoen EXCELLENCE (CZ.02.1.01/0.0/0.0/17_043/0009632).

ACKNOWLEDGMENTS

The authors would like to thank Jakub Martiník for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2021.622558/full#supplementary-material>

9. Hayward DG, Archer JC, Andrews S, Fairchild RD, Gentry J, Jenkins R, et al. Application of a high-resolution quadrupole/orbital trapping mass spectrometer coupled to a gas chromatograph for the determination of persistent organic pollutants in cow's and human milk. *J Agric Food Chem.* (2018) 66:11823–9. doi: 10.1021/acs.jafc.8b03721
10. Mol HGJ, Tienstra M, Zomer P. Evaluation of gas chromatography – electron ionization – full scan high resolution orbitrap mass spectrometry for pesticide residue analysis. *Anal Chim Acta.* (2016) 935:161–72. doi: 10.1016/j.aca.2016.06.017
11. Abushareeda W, Tienstra M, Lommen A, Blokland M, Sterk S, Kraiem S, et al. Comparison of gas chromatography/quadrupole time-of-flight and quadrupole orbitrap mass spectrometry in anti-doping analysis: I. Detection of anabolic-androgenic steroids. *Rapid Commun Mass Spectrom.* (2018) 32:2055–64. doi: 10.1002/rcm.8281
12. Brockbals L, Habicht M, Hajdas I, Galassi FM, RühlShared Last Authorship. FJ, Kraemer T. Untargeted metabolomics-like screening approach for chemical characterization and differentiation of canopic jar and mummy samples from ancient Egypt using GC-high resolution MS. *Analyst.* (2018) 143:4503–12. doi: 10.1039/c8an01288a
13. Peterson AC, Hauschild JB, Quarmby ST, Krumwiede D, Lange O, Lemke RAS, et al. Development of a GC/quadrupole-orbitrap mass spectrometer, part I: design and characterization. *Anal Chem.* (2014) 86:10036–43. doi: 10.1021/ac5014767
14. Hollender J, Schymanski EL, Singer HP, Ferguson PL. Nontarget screening with high resolution mass spectrometry in the environment: ready to go? *Environ Sci Technol.* (2017) 51:11505–12. doi: 10.1021/acs.est.7b02184
15. Kwiecien NW, Bailey DJ, Rush MJP, Cole JS, Ulbrich A, Hebert AS, et al. High-resolution filtering for improved small molecule identification via GC/MS. *Anal Chem.* (2015) 87:8328–35. doi: 10.1021/acs.analchem.5b01503
16. Su X, Lu W, Rabinowitz JD. Metabolite spectral accuracy on orbitraps. *Anal Chem.* (2017) 89:5940–8. doi: 10.1021/acs.analchem.7b00396
17. Baumeister TUH, Ueberschaar N, Pohnert G. Gas-phase chemistry in the GC orbitrap mass spectrometer. *J Am Soc Mass Spectrom.* (2019) 30:573–80. doi: 10.1007/s13361-018-2117-5
18. Margolin Eren KJ, Elkabets O, Amirav A. A comparison of electron ionization mass spectra obtained at 70 eV, low electron energies and with cold EI and their NIST library identification probabilities. *J Mass Spectrom.* (2020) 55:e4646. doi: 10.1002/jms.4646
19. McEachran AD, Balabin I, Cathey T, Transue TR, Al-Ghoul H, Grulke C, et al. Linking *in silico* MS/MS spectra with chemistry data to improve identification of unknowns. *Sci Data.* (2019) 6:141. doi: 10.1038/s41597-019-0145-z
20. Hites RA, Jobst KJ. Response to “Letter to the Editor: Optimism for Nontarget Analysis in Environmental Chemistry.” *Environ Sci Technol.* (2019) 53:5531–3. doi: 10.1021/acs.est.9b02473
21. van Den Dool H, Kratz PD. A generalization of the retention index system including linear temperature programmed gas–liquid partition chromatography. *J Chromatogr A.* (1963) 11:463–71. doi: 10.1016/S0021-9673(01)80947-X
22. Chambers MC, MacLean B, Burke R, Amodei D, Ruderman DL, Neumann S, et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol.* (2012) 30:918–20. doi: 10.1038/nbt.2377
23. Holman JD, Tabb DL, Mallick P. Employing proteowizard to convert raw mass spectrometry data. *Curr Protoc Bioinformatics.* (2014) 46:13.24.1–9. doi: 10.1002/0471250953.bi1324s46
24. Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, et al. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods.* (2015) 12:523–6. doi: 10.1038/nmeth.3393
25. Lai Z, Tsugawa H, Wohlgemuth G, Mehta S, Mueller M, Zheng Y, et al. Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics. *Nat Methods.* (2018) 15:53–6. doi: 10.1038/nmeth.4512
26. Ausloos P, Clifton CL, Lias SG, Mikaya AI, Stein SE, Tchekhovskoi DV, et al. The critical evaluation of a comprehensive mass spectral library. *J Am Soc Mass Spectrom.* (1999) 10:287–99. doi: 10.1016/S1044-0305(98)00159-7
27. Tsugawa H, Kind T, Nakabayashi R, Yukihira D, Tanaka W, Cajka T, et al. Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. *Anal Chem.* (2016) 88:7946–58. doi: 10.1021/acs.analchem.6b00770
28. Stein S. Mass spectral reference libraries: an ever-expanding resource for chemical identification. *Anal Chem.* (2012) 84:7274–82. doi: 10.1021/ac301205z
29. Stein SE, Scott DR. Optimization and testing of mass spectral library search algorithms for compound identification. *J Am Soc Mass Spectrom.* (1994) 5:859–66. doi: 10.1016/1044-0305(94)87009-8
30. NIST/EPA/NIH Mass Spectral Library (NIST 14). *NIST Standard Reference Database 1A.* National Institute of Standards and Technology (NIST) (2004). Available online at: <http://www.nist.gov/srd/>
31. Tsugawa H. *CompMS Metabolomics MSP Spectra Kit.* Available online at: <http://prime.psc.riken.jp/compms/msdial/main.html#MSP>
32. Scientific Working Group for the Analysis of Seized Drugs. *SWGDRUG Mass Spectral Library.* Available online at: <https://swgdrug.org/ms.htm> (accessed October 11, 2020).
33. Wallace WE, Ji W, Tchekhovskoi DV, Phinney KW, Stein SE. Mass spectral library quality assurance by inter-library comparison. *J Am Soc Mass Spectrom.* (2017) 28:733–8. doi: 10.1007/s13361-016-1589-4
34. Cayman Chemical. *Cayman Spectral Library.* Available online at: <https://www.caymanchem.com/forensics/publications/csl> (accessed October 11, 2020).
35. Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, Bergmuller E, et al. GMD@CSB.DB: the Golm metabolome database. *Bioinformatics.* (2005) 21:1635–8. doi: 10.1093/bioinformatics/bti236
36. Babushok VI, Linstrom PJ, Reed JJ, Zenkevich IG, Brown RL, Mallard WG, et al. Development of a database of gas chromatographic retention properties of organic compounds. *J Chromatogr A.* (2007) 1157:414–21. doi: 10.1016/j.chroma.2007.05.044
37. NIST. *NIST 17 Libraries and Software.* Maryland. (2017). doi: 10.18434/T4H594
38. Wohlgemuth G, Haldiya PK, Willighagen E, Kind T, Fiehn O. The chemical translation service—a web-based tool to improve standardization of metabolomic reports. *Bioinformatics.* (2010) 26:2647–8. doi: 10.1093/bioinformatics/btq476
39. United States Environmental Protection Agency. *CompTox Chemicals Dashboard.* Available online at: <https://comptox.epa.gov/dashboard> (accessed October 11, 2020).
40. *ACD/ChemSketch, version 2018.2.1* Toronto, ON: Advanced Chemistry Development, Inc. (2018). Available online at: www.acdlabs.com
41. Sander T, Freyss J, Von Korff M, Rufener C. DataWarrior: an open-source program for chemistry aware data visualization and analysis. *J Chem Inf Model.* (2015) 55:460–73. doi: 10.1021/ci500588j
42. Chen Y, Stork C, Hirte S, Kirchmair J. NP-scout: machine learning approach for the quantification and visualization of the natural product-likeness of small molecules. *Biomolecules.* (2019) 9:43. doi: 10.3390/biom9020043
43. Stork C, Embruch G, Šicho M, De Bruyn Kops C, Chen Y, Svozil D, Kirchmair J. NERDD: A web portal providing access to *in silico* tools for drug discovery. *Bioinformatics.* (2020) 36:1291–2. doi: 10.1093/bioinformatics/btz695
44. Djoumbou Feunang Y, Eisner R, Knox C, Chepelev L, Hastings J, Owen G, et al. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J Cheminform.* (2016) 8:1–20. doi: 10.1186/s13321-016-0174-y
45. Sievert C. *Interactive Web-Based Data Visualization With R, Plotly, and Shiny.* Florida: Chapman and Hall/CRC (2020). doi: 10.1201/9780429447273
46. Heller SR, McNaught A, Pletnev I, Stein S, Tchekhovskoi D. InChI, the IUPAC International Chemical Identifier. *J Cheminformatics.* (2015) 7:23. doi: 10.1186/s13321-015-0068-4
47. Meijer J, Lamoree M, Hamers T, Antingac J-P, Hutinet S, Debrauwer L, et al. *S71 | CECSCREEN | HBM4EU CECscreen: Screening List for Chemicals of Emerging Concern Plus Metadata and Predicted Phase 1 Metabolites.* (2020). doi: 10.5281/ZENODO.395658 (accessed October 11, 2020).
48. Stanstrup J, Broeckling C, Helmus R, Hoffmann N, Mathé E, Naake T, et al. The metaRbolomics toolbox in Bioconductor and beyond. *Metabolites.* (2019) 9:200. doi: 10.3390/metabo9100200
49. Matsuo T, Tsugawa H, Miyagawa H, Fukusaki E. Integrated strategy for unknown EI-MS identification using quality control calibration curve, multivariate analysis, EI-MS spectral database, and retention index prediction. *Anal Chem.* (2017) 89:6766–73. doi: 10.1021/acs.analchem.7b01010
50. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying small molecules via high resolution mass spectrometry: communicating

- confidence. *Environ Sci Technol.* (2014) 48:2097–8. doi: 10.1021/es5002105
51. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis. *Metabolomics.* (2007) 3:211–21. doi: 10.1007/s11306-007-0082-2
52. Aksekov AA, Laponogov I, Zhang Z, Doran SLF, Belluomo I, Veselkov D, et al. Auto-deconvolution and molecular networking of gas chromatography-mass spectrometry data. *Nat Biotechnol.* (2021) 39:169–73. doi: 10.1038/s41587-020-0700-3
53. Elie N, Santerre C, Touboul D. Generation of a molecular network from electron ionization mass spectrometry data by combining MZmine2 and MetGem software. *Anal Chem.* (2019) 91:11489–92. doi: 10.1021/acs.analchem.9b02802
54. Fernie AR, Aharoni A, Willmitzer L, Stitt M, Tohge T, Kopka J, et al. Recommendations for reporting metabolite data. *Plant Cell.* (2011) 23:2477–82. doi: 10.1105/tpc.111.086272
55. Dossin E, Martin E, Diana P, Castellon A, Monge A, Pospisil P, et al. Prediction models of retention indices for increased confidence in structural elucidation during complex matrix analysis: application to gas chromatography coupled with high-resolution mass spectrometry. *Anal Chem.* (2016) 88:7539–47. doi: 10.1021/acs.analchem.6b00868
56. Schymanski EL, Williams AJ. Open science for identifying “known unknown” chemicals. *Environ Sci Technol.* (2017) 51:5357–9. doi: 10.1021/acs.est.7b01908

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Price, Palát, Coufaliková, Kukučka, Codling, Vitale, Koudelka and Klánová. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.