Journal of proteome research

Technical Note

Cite This: J. Proteome Res. 2019, 18, 797−802

pubs.acs.org/jpr

CEU Mass Mediator 3.0: A Metabolite Annotation Tool

Alberto Gil-de-la-Fuente,*[,†,‡] Joanna Godzien,[‡] Sergio Saugar,[†] Rodrigo Garcia-Carmona,[†] Hasan Badran,[§] David S. Wishart,[§,∥,⊥] Coral Barbas,[‡] and Abraham Otero[†,‡]

[†]Department of Information Technology, Escuela Politécnica Superior, Universidad San Pablo-CEU, CEU Universities, Campus Montepríncipe, Boadilla del Monte, Madrid 28668, Spain

[‡]Centre for Metabolomics and Bioanalysis (CEMBIO), Facultad de Farmacia, Universidad San Pablo-CEU, CEU Universities, Campus Montepríncipe, Boadilla del Monte, Madrid 28668, Spain

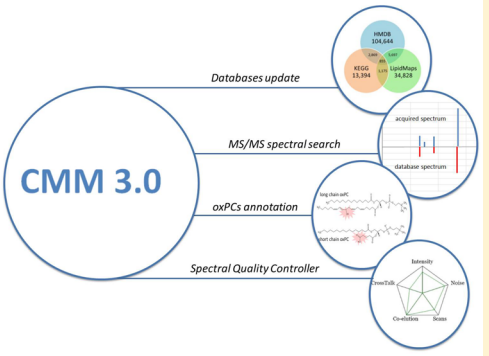
[§]Department of Biological Sciences University of Alberta, Edmonton, Alberta T6G 2E9, Canada

[∥]Department of Computing Science, University of Alberta, Edmonton, Alberta T6G 2E8, Canada

[⊥]Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, Edmonton, Alberta T6G 2N8, Canada

Supporting Information

ABSTRACT: CEU Mass Mediator (CMM, http://ceumass.eps.uspceu.es) is an online tool that has evolved from a simple interface to query different metabolomic databases (CMM 1.0) to a tool that unifies the compounds from these databases and, using an expert system with knowledge about the experimental setup and the compounds properties, filters and scores the query results (CMM 2.0). Since this last major revision, CMM has continued to grow, expanding the knowledge base of its expert system and including new services to support researchers in the metabolite annotation and identification process. The information from external databases has been refreshed, and an in-house library with oxidized lipids not present in other sources has been added. This has increased the number of experimental metabolites up 332,665 and the number of predicted metabolites to 681,198. Furthermore, new taxonomy and ontology metadata have been included. CMM has expanded its functionalities with a service for the annotation of oxidized glycerophosphocholines, a service for spectral comparison from MS^2 data, and a spectral quality-assessment service to determine the reliability of a spectrum for compound identification purposes. To facilitate the collaboration and integration of CMM with external tools and metabolomic platforms, a RESTful API has been created, and it has already been integrated into the HMDB (Human Metabolome Database). This paper will present the novel functionalities incorporated into version 3.0 of CMM.



KEYWORDS: *metabolomics, annotation, identification, knowledge representation, mass spectrometry, databases, REST, web services, software tool*

INTRODUCTION

Compound annotation and identification remains one of the major bottlenecks in untargeted metabolomics.[1−3] Mass spectrometry (MS) is the dominant platform in metabolomics and lipidomics[4] due to its high sensitivity.[5] MS is commonly coupled to different separation techniques. Among them, high-performance liquid chromatography (HPLC) is the most frequently used.[4]

LC−MS can be used in hyphenated setups (LC−MS/MS) for obtaining the fragmentation pattern of the analyzed compounds or samples. MS/MS or MS^n analyses provide structural information based on the fragmentation pattern. This fragmentation pattern can be compared against reference MS/MS spectra present in metabolomic databases, providing evidence pointing to different possible annotations.

In parallel with the development of metabolomics, there has been an increase in the number and a growth in the size of metabolomic databases.[6−15] Although some compounds are present in most databases, there is not a complete overlap among them, forcing the researcher to use different databases and then manually unifying the results. This task is highly time-consuming and prone to errors.

Most of the software tools for compound annotation are based on MS/MS or MS^n information, which can provide a higher level of confidence for the annotations. However, in some experiments, the sample quantity is limited; therefore, the MS/MS analyses cannot be performed. Furthermore, in some cases, a previous filter using only MS^1 is useful, or even necessary. The time spent on compound annotation/identification can be decreased by filtering the putative annotations before using the fragmentation information with RT and m/z data.

Special Issue: Software Tools and Resources 2019

Received: September 14, 2018
Published: December 21, 2018

DOI: 10.1021/acs.jproteome.8b00720
J. Proteome Res. 2019, 18, 797−802

Table 1. Updated Confidence Levels Proposed by the Metabolomics Society

confidence level	description	matching requirements
level 0	unequivocal 3D structure, including full stereochemistry	determination of 3D structure following natural product guidelines
level 1	confident 2D structure, using reference standard or full 2D structure elucidation	at least two orthogonal characteristics, such as MS/MS fragmentation pattern, retention time (RT), or collision cross-section (CSS)
level 2	probable structure using literature data and/or fragmentation spectra and/or knowledge over the RT	at least two orthogonal characteristics matching and evidence of excluding the rest of candidates
level 3	possible structure, isomers, or class	more than one candidate; only one characteristic matched is required for supporting the proposed candidate
level 4	unknown	quantifiable feature in a sample

Author: CEU Mass Mediator (CMM) is a freely available online tool designed to support researchers in metabolite annotation tasks corresponding to the confidence levels 2 and 3 from the Metabolomics Standards Initiative (MSI) (see Table 1).[3] It allows users to execute queries over the previously unified compounds present in different databases, HMDB, KEGG, and LipidMaps, and predicted compounds from HMDB and MINE. It also collects cross-references from Metlin (due to the spectra available in this database and its wide appeal to many users) and PubChem.

CMM 1.0 was first released in 2014 as a simple service for batch queries over several databases (KEGG, LipidMaps, and Metlin).[16] CMM 2.0, released in 2017,[17] integrated compounds from the HMDB and MINE databases, bringing the total to 279,318 experimental compounds and 672,042 predicted compounds. It also featured an expert system made up of 124 rules based on knowledge obtained from researchers, which filters and scores the putative annotations before presenting them to the user.[17]

The purpose of this paper is to introduce CMM 3.0. In this major update, the information from the databases integrated in CMM has been refreshed, and information from oxidized glycerophosphocholines (oxPCs), from LipidMaps and from in-house data generated in CEMBIO, has been added. The total number of metabolites has increased to 332,665 experimental and 681,198 predicted compounds. The CMM database has also been extended with MS/MS spectra data from HMDB, taxonomic, and ontological information about the compounds (previously this information was not available) as well as additional information about metabolic pathways. CMM 3.0 also provides, for the first time, support for using MS/MS data for annotation and identification. In addition, CMM now offers support for the characterization and annotation of oxidized lipids that is based on experimental knowledge about the oxidation of glycerophosphocholines and a tool that assesses the quality of a spectrum that it is being used for metabolite identification. Finally, CMM 3.0 also provides a RESTful API that allows facile integration of CMM functionality into external tools. Using this API, we have recently integrated CMM into HMDB as a new search functionality (http://www.hmdb.ca/spectra/ms_cmm/search), and integrations with other tools are under development at this time.

CMM source code is available in GitHub (https://github.com/albertogilf/ceuMassMediator) under the GNU General Public License v3.0. CMM is a J2EE application, and it may be accessed through any web browser that supports JavaScript.

NEW FEATURES OF CMM 3.0

Because of the nature of the metabolomic software tools that use data from external sources, some improvements emerge directly from the update of the integrated sources; others benefit or depend on these updates but require additional work by the developers of the tool, whereas other improvements arise from independent work and are not related to the sources. CMM 3.0 includes improvements that fall into each of these three categories, and each enhancement will be discussed in the following sections.

Database Updates

Since the release of CMM 2.0, nearly all of the external databases integrated in CMM have been updated. HMDB has released a new version, HMDB 4.0,[18] which expanded the number of metabolites from 40,153 to 114,100 (including in silico predicted compounds), as well as the information available about the compounds. HMDB has also updated the chemical taxonomy system using ClassyFire[19] for all of its compounds. This has been integrated into the CMM database and is now used by the CMM expert system to gather evidence supporting or refuting the putative annotation of lipids. The HMDB has also included information about the physiological effects of the compounds, their source (endogenous or exogenous), biological location, biological role, and the processes in which such compounds are involved. This information is very useful for the annotation of compounds. For instance, knowing that a compound has been detected in blood before and has been related to some type of kidney disease is invaluable for a metabolomics analysis that aims to study the biomarkers of a kidney disease using blood samples. This ontology information can also be used as a filter for researchers who are only interested in compounds previously detected under particular conditions. For example, some studies may have no interest in endogenous or exogenous compounds, whereas others may want to search only for compounds detected in a specific organ (bladder, kidney, liver, etc.).

LipidMaps has launched the new LipidMaps Lipidomics Gateway, where the curation of the lipids has been restarted, and more than 500 lipids have been added since the last update of CMM. Moreover, some bugs have been fixed, such as the duplication of InChIs or the correctness of the taxonomy of some compounds. The corrections made by LipidMaps to their taxonomy have resulted in an improvement of the correctness of the rules applied by the CMM expert system when gathering evidence about the putative annotations because this taxonomy is used by the CMM rules.[17]

The KEGG database has increased its number of metabolites and pathways (from 399 to 422). Pathways information is useful for the subsequent biological interpretation. The CMM pathway displayer tool uses pathway information from this resource, showing to the users the pathways where each metabolite has been detected.

DOI: 10.1021/acs.jproteome.8b00720
J. Proteome Res. 2019, 18, 797−802

CMM has been supplemented with information from oxPCs from an in-house library, expanding the current information present in the other databases. There are 248 oxidized glycerophospholipids (oxGPs) currently in LipidMaps and 48 oxPCs. The CMM library contains 24 new oxPCs that are not present in any of the databases integrated.[20] Table 2 illustrates the information present in CMM from each source.

Table 2. Number of Compounds and Type of Information Available in the Different Versions of CMM

database	CMM 1.0	CMM 2.0	CMM 3.0
HMDB	0	74,484	114,065
	N/A	structure	structure, taxonomy, pathways, ontology
KEGG	13,526	15,909	18,293
	N/A	structure, pathways	structure, pathways
LipidMaps	37,576	40,213	42,555
	N/A	structure, taxonomy	structure, taxonomy (corrected)
MINE	0	672,042	672,042
	N/A	structure	structure
in-house	0	0	24
	N/A	N/A	structure, taxonomy

Figure 1 shows the overlap of the metabolite coverage between the different databases integrated in CMM 2.0 and

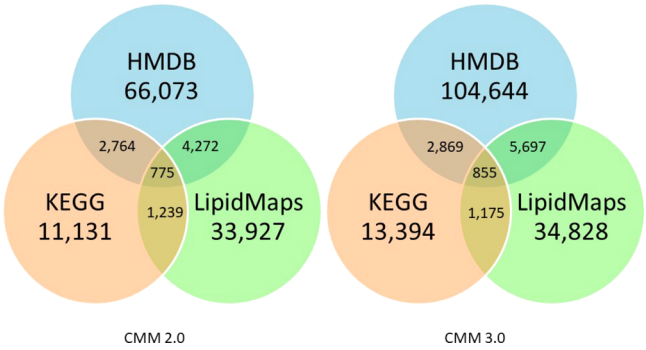


Figure 1. Venn diagram with the coverage of metabolites between different databases integrated in CMM 2.0 and CMM 3.0.

CMM 3.0. The comparison only covers those compounds for which there is sufficient information available to perform unequivocal compound unification (i.e., those with 3D structure, the InChI, the InChI key, or the canonical SMILES available). The reason why there seems to be a reduction (from 1239 to 1175 metabolites) in the overlap between LipidMaps and KEGG in CMM 3.0 (despite the addition of more compounds to both databases) is the previously discussed errata present in some LipidMaps compounds, which sometimes caused the incorrect unifications in CMM 2.0. Although the global number of metabolites increased, the overlap of metabolites among the databases is still low.

Annotation of Oxidized Lipids

The biological role of oxidized lipids is currently an active research topic, notably contributing to the understanding of health and disease. Within human metabolomic studies, it was observed that lipids are significantly affected by oxidative stress.[21,22] However, until now, the number of tools to support the annotation of oxidized lipids has been small.[23] Nowadays, this task usually starts with the annotation of the signals by searching for experimental *m/z* matches in the databases. However, the number of oxidized lipids currently present in such databases is low. LipidMaps, the reference database for lipids, contains only 248 oxidized lipids for all types of heads. This makes the annotation of oxidized phospholipids challenging, especially when the target compound is not present in the database, and it increases the likelihood of the compound being assigned as an unknown. There are some patterns that can provide the researcher with clues about the possibility that a feature may correspond to an oxidized lipid, for example, a lower RT for reversed-phase (RP) chromatography of the shortened chains due to the lower hydrophobicity, a high level of fragmentation with 40 eV, the ionization through deprotonation ($[M - H]^-$) and the presence of the neutral loss of water (usually detectable in positive ionization mode).[20] CMM has developed a service to annotate and characterize oxidized lipids. The full process is explained in SI1.

MS/MS Search

There is a large number of MS/MS-based annotation tools using different approaches to spectral matching or compound identification. Until now, CMM had only supported the annotation of MS^1 data. CMM 3.0 has integrated the MS/MS information present in HMDB, including experimental and predicted spectra, with the aim of providing support to compare experimental MS/MS spectra against the HMDB reference data. The large number of existing algorithms to calculate spectral similarity and their robustness leads us to select the three most popular promising ones and to perform an independent evaluation instead of proposing a new solution.[8,10,24−27] The results of the evaluation are shown in SI2.

To use the MS/MS search service of CMM, the researcher should introduce the precursor ion mass, the list of pairs of *m/z* and intensity of the product ions, and the tolerance allowed for the precursor ion and the product ions (in Da or ppm). The intensity can be normalized or absolute (CMM normalizes the values if needed). The ionization mode used and voltage applied are also necessary to restrict the search over the corresponding experimental setup. The researcher can choose if the MS/MS spectra comparison should be performed against experimental or predicted spectra. Once all of the information is submitted, the CMM comparison algorithm performs an initial filter of the putative annotations based on the precursor ion, the precursor ion tolerance, the ionization mode, the fragmentation voltage and the type of spectra. The compounds with spectra available under these conditions are then scored to determine the similarity between the input spectra and the putative annotations.

Spectral Quality Assessment for Identification Purposes

The success of an untargeted metabolomics study depends on the correctness of the identification process. Decreasing the number of unknowns and misidentifications is key to having a biologically significant finding. Nevertheless, the time to perform a study is restricted, and the low availability of reference standards for clear compound identification often hinders this task in untargeted approaches. Consequently, a high-quality MS/MS spectrum is paramount to improve the annotation rate of compounds, whereas a low-quality spectrum increases the risk of misidentification.

DOI: 10.1021/acs.jproteome.8b00720
J. Proteome Res. 2019, 18, 797−802

However, assessing spectral quality can be difficult. To provide researchers with a systematic method to evaluate the quality of a MS/MS spectrum, CMM has created a pentagonal-point evaluation system that takes into account: (1) the quality of the overall intensity, (2) the impact of the noise, (3) the number of MS/MS scans obtained, (4) the presence of different precursor ions in the collision cell at the same time, and (5) the presence of delayed ions from the previous scans, a phenomena known as cross-talk. In Figure 2, we can see a graphic representation of this evaluation system. The pentagon on the left shows the evaluation result for an excellent spectrum (green lines), whereas the pentagon in the right corresponds to an inadequate spectrum (red lines). The closer the lines are to the pentagon vertex, the better the spectrum is. A full explanation about the principles used and how the spectral quality assessment has been developed can be found in SI3.[28]

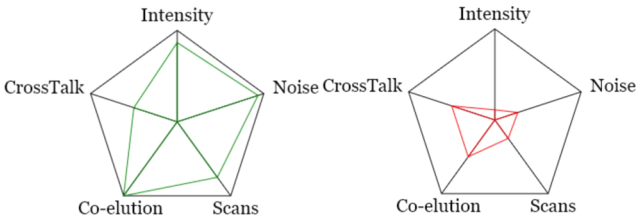


Figure 2. Graphic representation of the pentagonal-point evaluation system used in CMM 3.0 to asses the quality of the spectrum introduced by the user. On the left, a good quality spectrum. On the right, a poor quality spectrum.

RESTful API

The metabolomics field is continuously growing and so is the number of tools available to assist in metabolomic data analysis. Typically, metabolomic tools do not provide all of the functionalities that a metabolomics workflow requires. Therefore, researchers often have to use different tools to carry out their analyses and are forced to use results from one tool in another, a task that is not always trivial. To mitigate this disadvantage, several platforms that try to integrate different external tools into a single pipeline have emerged. Two of the most popular open platforms are the Workflow4Metabolomics[29] and PhenoMeNal.[30] The Elixir metabolomics community aims to share the data and the software tools with these frameworks because they have proven to be useful, and they can improve the reproducibility of the data analysis.[31]

CMM provides functionality that is not available in other metabolomics tools: its knowledge-based approach to filtering and scoring putative annotations as well as its support for the identification of oxidized lipids using experimental knowledge. The compound unification from multiple external databases also provides great value. Because of these unique features, there has been a growing number of requests to integrate CMM features in external tools through an application programming interface (API). This RESTful (REpresentational State Transfer architectural style compliant) API for CMM has been already integrated into the HMDB environment (http://www.hmdb.ca/spectra/ms_cmm/search), where users can take advantage of the CMM filtering and scoring functionality directly from the HMDB web interface. The details of this API can be found in SI4.

CONCLUSIONS

We have presented version 3.0 of CMM, a free online tool to support many of the needs of researchers in the annotation of metabolites. CMM integrates and unifies experimental and predicted metabolites from several databases, including HMDB, KEGG, LipidMaps, and MINE, allowing the user to query in all of them through a single interface. In addition, CMM uses an expert system to filter and score putative annotations, allowing researchers to focus on those annotations that are more plausible.

After the data refreshing performed in version 3.0 and the integration of an in-house library of oxidized lipids, the total number of available experimental metabolites in CMM 3.0 is 332,665 and the total number of predicted metabolites is 681,198. Taxonomy and ontology information from HMDB and LipidMaps has been added to the CMM database, being used by its expert system. Novel functionalities that have been added to CMM 3.0 include MS/MS search support, a service for the annotation of oxidized glycerophospholipids and a spectral quality-assessment tool to measure the quality of the MS/MS spectra. Furthermore, in CMM 3.0, the search services are now available through a RESTful API that has already been used to integrate CMM functionalities into the HMDB. For future work, we intend to further exploit the taxonomy and ontology information that is available now in the CMM database to enhance the filtering and scoring performed by our expert system. Some of this information is already used by the 124 rules currently available in the expert system.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.8b00720.

SI1: Annotation of oxidized lipids. SI2: Independent evaluation of MS/MS search. SI3: Spectral quality assessment development. SI4: CMM RESTful API (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: alb.gil.ce@ceindo.ceu.es. Tel: +34 913724711.

ORCID

Alberto Gil-de-la-Fuente: 0000-0002-5951-1601
Joanna Godzien: 0000-0002-9477-057X
Sergio Saugar: 0000-0002-4216-9256
Rodrigo Garcia-Carmona: 0000-0003-4427-9579
Hasan Badran: 0000-0002-8072-5233
David S. Wishart: 0000-0002-3207-2434
Coral Barbas: 0000-0003-4722-491X
Abraham Otero: 0000-0003-4568-2933

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

A.G.-F., J.G., C.B., and A.O. received funding from the Spanish Ministerio de Economía y Competitividad (grant CTQ2014-55279-R). A.G.-F., C.B., and A.O. received funding from USPCEU PCON10/2016. A.G.-F. thanks Fundación Uni-

DOI: 10.1021/acs.jproteome.8b00720
J. Proteome Res. 2019, 18, 797−802

versitaria San Pablo CEU for his Ph.D. fellowship and the scholarship from the III international mobilities for researchers funded by CEU-Banco Santander. H.B. and D.S.W. received funding from the HMDB project, which is supported by the Canadian Institutes of Health Research (award no. 111062), Alberta Innovates - Health Solutions, and by The Metabolomics Innovation Centre (TMIC), a nationally funded research and core facility that supports a wide range of cutting-edge metabolomic studies. TMIC is funded by Genome Alberta, Genome British Columbia, and Genome Canada, a not-for-profit organization that is leading Canada's national genomics strategy with $900 million in funding from the federal government. We thank all of the publicly available databases that provide the data that allow us to create software tools for research.

■ ABBREVIATIONS

CEMBIO, Centre for Metabolomics and Bioanalysis; CMM, CEU Mass Mediator; CSS, collision cross-section; HPLC, high-performance liquid chromatography; J2EE, Java 2 Platform, Enterprise Edition; MS, mass spectrometry; MSI, metabolomics standards initiative; oxGPs, oxidized glycerophospholipids; oxPCs, oxidized glycerophosphocholines; REST, representational state transfer; RESTful, representational state transfer architectural style compliant; RP, reversed-phase; RT, retention time

■ REFERENCES

(1) Gil de la Fuente, A.; Grace Armitage, E.; Otero, A.; Barbas, C.; Godzien, J. Differentiating signals to make biological sense - A guide through databases for MS-based non-targeted metabolomics. *Electrophoresis* **2017**, *38*, 2242−2256.

(2) Peisl, B. Y. L.; Schymanski, E. L.; Wilmes, P. Dark matter in host-microbiome metabolomics: Tackling the unknowns - A review. *Anal. Chim. Acta* **2018**, *1037*, 13−27.

(3) Blazenovic, I.; Kind, T.; Ji, J.; Fiehn, O. Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. *Metabolites* **2018**, *8* (2), 31.

(4) Cajka, T.; Fiehn, O. Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics. *Anal. Chem.* **2016**, *88* (1), 524−545.

(5) Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* **2007**, *26*, 51−78.

(6) Wishart, D. S. Current Progress in computational metabolomics. *Briefings Bioinf.* **2007**, *8*, 279−293.

(7) Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research (United Kingdom)* **2000**, *28*, 27−30.

(8) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. METLIN - A metabolite mass spectral database. *Ther. Drug Monit.* **2005**, *27*, 747−751.

(9) Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill, A. H. J.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W.; Subramaniam, S. LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* **2007**, *35*, D527−D532.

(10) Horai, H.; et al. MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **2010**, *45*, 703−714.

(11) Wishart, D. S.; et al. HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Res.* **2012**, *41*, D801−D807.

(12) Jeffryes, J. G.; Broadbelt, L. J.; Tyo, K. E. J.; Colastani, R. L.; Henry, C. S.; Elbadawi-Sidhu, M.; Kind, T.; Fiehn, O.; Niehaus, T. D.; Hanson, A. D. MINEs: Open access databases of computationally predicted enzyme promiscuity products for untargeted metabolomics. *J. Cheminf.* **2015**, *7*, 44 DOI: 10.1186/s13321-015-0087-1.

(13) Caspi, R.; Billington, R.; Fulcher, C. A.; Keseler, I. M.; Kothari, A.; Krummenacker, M.; Latendresse, M.; Ong, Q.; Paley, S.; Subhraveti, P.; Weaver, D. S.; Karp, P. D.; Ferrer, L.; Foerster, H.; Mueller, L. A. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **2016**, *44*, D471−D480.

(14) Sud, M.; Fahy, E.; Cotter, D.; Azam, K.; Vadivelu, I.; Subramaniam, S.; Burant, C.; Edison, A.; Fiehn, O.; Higashi, R.; Nair, K. S.; Sumner, S. Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Res.* **2016**, *44*, D463−D470.

(15) *MzCloud*, 2018. https://www.mzcloud.org/.

(16) Johnson, C. H.; Ivanisevic, J.; Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 451−459.

(17) Gil de la Fuente, A.; Godzien, J.; Fernández López, F. J.; Rupérez, F. J.; Barbas, C.; Otero, A. Knowledge-based metabolite annotation tool: CEU Mass Mediator. *J. Pharm. Biomed. Anal.* **2018**, *154*, 138−149.

(18) Wishart, D. S.; et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **2018**, *46*, D608.

(19) Djoumbou Feunang, Y.; Eisner, R.; Knox, C.; Chepelev, L.; Hastings, J.; Owen, G.; Fahy, E.; Steinbeck, C.; Subramanian, S.; Bolton, E.; Greiner, R.; Wishart, D. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J. Cheminf.* **2016**, *8*, 1.

(20) Gil de la Fuente, A.; Traldi, F.; Siroka, J.; Kretowski, A.; Ciborowski, M.; Otero, A.; Barbas, C.; Godzien, J. Characterization and annotation of oxidized glycerophosphocholines for non-targeted metabolomics with lc-qtof-ms data. *Anal. Chim. Acta* **2018**, *1037*, 358−368.

(21) Filomeni, G.; De Zio, D.; Cecconi, F. Oxidative stress and autophagy: the clash between damage and metabolic needs. *Cell Death Differ.* **2015**, *22*, 377.

(22) Houten, S. M.; Violante, S.; Ventura, F. V.; Wanders, R. J. A. The Biochemistry and Physiology of Mitochondrial Fatty Acid beta-Oxidation and Its Genetic Disorders. *Annu. Rev. Physiol.* **2016**, *78*, 23.

(23) Ni, Z.; Angelidou, G.; Hoffmann, R.; Fedorova, M. LPPtiger software for lipidome-specifc prediction and identifcation of oxidized phospholipids from LC-MS datasets. *Sci. Rep.* **2017**, *7*, 15138.

(24) Ridder, L.; van der Hooft, J. J. J.; Verhoeven, S.; de Vos, R. C. H.; van Schaik, R.; Vervoort, J. Substructure-based annotation of high-resolution multistage MSn spectral trees. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 2461−2471.

(25) 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*, W94.

(26) Huan, T.; Li, L.; Tang, C.; Li, R.; Shi, Y.; Lin, G. MyCompoundID MS/MS Search: Metabolite Identification Using a Library of Predicted Fragment-Ion-Spectra of 383,830 Possible Human Metabolites. *Anal. Chem.* **2015**, *87*, 10619.

(27) Ruttkies, C.; Wolf, S.; Neumann, S.; Schymanski, E. L.; Hollender, J. MetFrag relaunched: Incorporating strategies beyond in silico fragmentation. *J. Cheminf.* **2016**, *8* (1), 3.

(28) Traldi, F.; de la Fuente, A. G.; Kowalczyk, T.; Ciborowski, M.; Otero, A.; Barbas, C.; Godzien, J. *A Spectral Quality Assessment for Annotation in Metabolomics Studies*; Universidad San Pablo-CEU, CEU Universities, 2018. (unpublished work).

(29) Guitton, Y.; et al. Create, run, share, publish, and reference your LC−MS, FIA−MS, GC−MS, and NMR data analysis workflows with the Workflow4Metabolomics 3.0 Galaxy online infrastructure for metabolomics. *Int. J. Biochem. Cell Biol.* **2017**, *93*, 89.

(30) Meier, R.; Ruttkies, C.; Treutler, H.; Neumann, S. Bioinformatics can boost metabolomics research. *J. Biotechnol.* **2017**, *261*, 137.

(31) van Rijswijk, M.; Beirnaert, C.; Caron, C.; Cascante, M.; Dominguez, V.; Dunn, W. B.; Ebbels, T. M. D.; Giacomoni, F.;

DOI: 10.1021/acs.jproteome.8b00720
J. Proteome Res. 2019, 18, 797−802

Gonzalez-Beltran, A.; Hankemeier, T.; et al. The future of metabolomics in ELIXIR. *F1000Research* **2017**, *6*, 1649 DOI: 10.12688/f1000research.12342.1.

DOI: 10.1021/acs.jproteome.8b00720
J. Proteome Res. 2019, 18, 797−802