Profiling the Lipidome requires quality control

Harald C. Köfeler[1*], Thomas O. Eichmann[2], Robert Ahrends[3], John A. Bowden[4], Niklas Danne-Rasche[5], Maria Fedorova[6,7], William J. Griffiths[8], Xianlin Han[9], Jürgen Hartler[10], Michal Holčapek[11], Robert Jirásko[11], Jeremy P. Koelmel[12], Christer S. Ejsing[13,14], Gerhard Liebisch[15], Zhixu Ni[6,7], Valerie B. O'Donnell [16], Oswald Quehenberger[10], Dominik Schwudke[17,18,19], Andrej Shevchenko[20], Michael J.O. Wakelam[21], Markus R. Wenk[22], Denise Wolrab[11] and Kim Ekroos[23]

[1] Core Facility Mass Spectrometry and Lipidomics, ZMF, Medical University of Graz, A-8010 Graz, Austria

[2] Institute of Molecular Biosciences, University of Graz, A-8010 Graz, Austria

[3] Department for Analytical Chemistry, University of Vienna, A-1090 Vienna, Austria

[4] Department of Physiological Sciences, College of Veterinary Medicine, University of Florida, 1333 Center Drive, Gainesville, FL 32610 USA

[5] University Duisburg Essen, 45141 Essen, Germany

[6] Institute of Bioanalytical Chemistry, Faculty of Chemistry and Mineralogy, Universität Leipzig, Germany

[7] Center for Biotechnology and Biomedicine, Universität Leipzig, Germany

[8] Swansea Univerity Medical School, Singleton Park, Swansea SA2 8PP, United Kingdom

[9] Barshop Inst Longev & Aging Studies, Univ Texas Hlth Sci Ctr San Antonio, San Antonio, TX 78229 USA

[10] Department of Pharmacology, University of California, San Diego, CA 92093, USA

[11] Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Pardubice, Czech Republic

[12] Department of Environmental Health Sciences, School of Public Health, Yale University, New Haven, CT 06510, USA

[13] Department of Biochemistry and Molecular Biology, VILLUM Center for Bioanalytical Sciences, University of Southern Denmark, Odense, Denmark

[14] Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, Heidelberg, Germany

[15] Institute of Clinical Chemistry and Laboratory Medicine, University of Regensburg, Germany

[16] Systems Immunity Research Institute, Cardiff University, CF14 4XN, United Kingdom

[17] Bioanalytical Chemistry, Research Center Borstel, 23845 Borstel, Germany

[18] German Center for Infection Research, TTU TB, Borstel, Germany

[19] German Centre for Lung Research, Airway Research Center North, Borstel, Germany

[20] Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany

[21] Babraham Institute, Babraham Research Campus, Cambridge, CB22 3AT, United Kingdom

[22] Singapore Lipidomics Incubator (SLING), Department of Biochemistry, YLL School of Medicine, National University of Singapore, Singapore, Singapore

[23] Lipidomics Consulting Ltd., FI-02230 Esbo, Finland

[*] Corresponding author

A recent publication from Vasilopoulou *et al.* reports on the full lipidome profiling by a combination of trapped ion mobility spectrometry (TIMS), parallel accumulation serial fragmentation (PASEF) and nano HPLC[1]. While this is altogether an impressive technological advance having the potential to increase lipidome coverage and lower the detection limit for individual lipids, the interpretation of acquired spectra is a matter of serious concern. We

noticed that it exclusively relied on the software-assisted lipid assignments that were not confirmed by manual inspection, matching spectra and retention times with available lipid standards or computer simulation – these measures are typically employed in lipidomics to reduce the false positive rate.

A common way to examine the robustness of identifications at the lipidome scale is to test if species from the same lipid class eluted in a rationale order concordant with their proposed chemical structure. In reversed-phase liquid chromatography, the equivalent carbon number (ECN) model suggests that their retention times should increase with increasing the number of carbon atoms in fatty acid moieties and decrease with increasing the number of double bonds within a lipid class. Previous works demonstrated that this relationship shows a high degree of linearity[2-4]. However, 55 out of 171 TGs and 130 out of 301 diacyl PC species reported by Vasilopoulou *et al* do not follow the ECN model (**Figures 1a and 1b**).

Additionally, several reported lipids displayed odd elution profiles. For example, DG 32:0 spread over the elution time range from 18.9 min to 28.61 min, which is unreasonably large for the conventional mobile phase gradient employed by Vasilopoulou *et al.* This would hint that some of the later eluting DG species are, in fact, fragments produced by in source fragmentation of TG species.

In several instances, the reported lipids do not corroborate their chemical structure. For example, eight PC O-16:0_1:0 species were reported, although only two sn-1/2 isomers (PC-O 16:0/1:0 or PC-O 1:0/16:0) are possible. Alternative structures for another six assignments comprising the same moieties (for example, including branched C 16:0 fatty acid(s) / fatty alcohol(s) or even more exotic sn-2/3 isomers) conflict the basic principles of lipid biosynthesis in mammals and must be extensively validated by chemical synthesis of

relevant standards. Similarly, 5 chromatographic peaks assigned to CE 18:2 and 4 peaks of cholesteryl 11-hydroperoxy-eicosatetraenoate need further verification.

In lipidomics, it is ultimately required that the elemental composition of recognized lipids exactly match the *m/z* of their intact molecular ions. This is critical for candidate molecules possessing uncommon fatty acid moieties (e.g. comprising oxidized or odd carbon numbered fatty acid moieties) especially if no same class species with known abundant moieties (e.g. 16:0; 18:0 or 18:1) were detected. A few further examples of questionable annotation are presented in **Suppl. Table 1**. Upon low-energy CID / HCD, lipids produce relatively few abundant and informative fragments. Hence, manual inspection of MS2 spectra of seemingly unusual lipid species is required and the same criteria for spectrum to structure matching should be consistently applied for the entire dataset. In particular, the identification of phospholipids missing characteristic head group fragments (e.g., PC, SM in positive or PI in negative modes) or their losses (PE or PS in positive mode) should be disregarded[5,6]. Apparently this basic principle was not applied for PI and PS, for which half of the species were identified without matching relevant head group fragments. (**Suppl. Table 1**). In positive mode MS2 five PC ($[M+H]^+$) precursors produced no phosphocholine head group fragment (*m/z* 184.07) that ought to be exceptionally abundant in a broad range of collision energies. Furthermore, in negative mode MS2 spectra of 409 PC no informative acyl anion fragments required for their unequivocal molecular species identification were detected. The identification of only 6% of PCs (28/437) relied upon the complete set of characteristic masses (e.g. exact masses of the precursor; head group and acyl anions fragments). Of note, the phosphocholine head group fragment (m/z 184.07) is not an unequivocal identifier of PC because it can also be produced from SM whose molecular ions often overlap with isotopic clusters of PC.

Similar problems are apparent in the identification of other classes of lipids. For example, SM d16:1_25:0 indicates a very unusual combination of sphingosine backbone and N-amidated fatty acid. However, its MS2 spectrum only confirms the presence of phosphocholine head group (m/z 184.07), which leads us to believe it is more likely to be SM d18:1_23:0 - a very common mammalian sphingomyelin.

We also observed a number of instances whereby the precursors were detected in the form of unreasonable molecular adducts, e.g. $[M\text{-}CH_3]^-$ for diacyl phosphatidylinositols that are having no methyl group to loose.

Molecular adducts of intact lipid precursors were often inconsistent with the composition of the employed mobile phase containing isopropanol, methanol, water and acetonitrile doped with ammonium formate and formic acid. 31% (10/32) DG, 21% (7/33) CE, 25% (1/4) ether-LPE and 15% (11/72) of PE and ether-PE species were annotated as molecular adducts that, under these conditions, cannot be detectable. For example, nine PE and ether-PE lipids were uniquely detected as acetate $[M+OAc]^-$ adducts despite the presence of 10 mM formate. Note that even in an ammonium acetate buffer, the $[M-H]^-$, but not the $[M+OAc]^-$ is the dominant molecular form of PE. Out of 437 PC species reported by Vasilopoulou *et al.*, 36 were detected as either redundant or non-dominant adducts in negative mode.

Lipidomes (including the plasma lipidome) are conserved molecular constellations and their quantification is an important means to validate the analytical bona fide. Thus, it is difficult to reconcile the identification of very minor free sterols in plasma when free cholesterol itself or its known major metabolites were not detected. Cholesterol is the most abundant single lipid in plasma and its concentration is more than 1000-fold higher than any sterol reported by Vasilopoulou *et al*. Many sterols are present in plasma as multiple isomers,

hence, without comparing CCS, retention time and fragmentation patterns to authentic standards, identifications cannot be considered unequivocal.

The list of examples of problematic identification could be extended, however we suggest that many of them could have been avoided by applying a simple identification rationale. The retention time of a proposed lipid structure should match the retention time pattern of its lipid class; the elemental composition of identified species must match the accurate masses of their precursor ions; molecular adducts of intact molecular ions should be in the dominant forms expected for the mobile phase composition; the matched fragments should be specific and corroborate the proposed structures and, last but not least, structural annotation of each species (including the identification of positional isomers) should match individual MS2 spectra rather than uniformly applied for the whole lipid class. When considering low abundance precursors or novel lipids, spectra should be subjected to manual re-inspection and, if possible, independent means of structural conformation should be sought. Although this could lower the number of reported lipid candidates, it vastly improves the data quality and integrity and ensures high biological relevance of the lipidome profile.

Lipidomics methods detecting large number of molecular species are increasingly used by the biomedical community, however data from these approaches should always be interpreted in the most rigorous fashion possible, taking into account all potential pitfalls and being extremely careful not to overinterpret findings. Helpful guidelines for interpretating and reporting the lipidome composition are provided by International Lipidomics Society (ILS), the Lipidomics Standards Initiative (LSI) and LIPID MAPS alike[7,8].

Acknowledgements

Funding to V.O.D. and M.O.J.W. from Wellcome Trust for LIPID MAPS (203014/Z/16/Z) is gratefully acknowledged. M.O.J.W. acknowledges funding from UKRI-BBSRC BBS/E/B/000C0431. Financial support for M.F. from German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept for SysMedOS project is gratefully acknowledged. J.H. gratefully acknowledges funding by a Max Kade fellowship awarded by the Austrian Academy of Sciences and H.K. gratefully acknowledges funding from the Austrian Federal Ministry of Education, Science and Research grant number BMWFW-10.420/0005-WF/V/3c/2017.

Author contributions

H.K., A.S., M.J.O.W., O.Q., V.O.D., W.J.G. and J.A.B. designed the text. H.K., R.A., N.D.R., J.A.B., T.O.E., M.F., Z.N., M.H., R.J., D.W. and J.P.K. checked the data. T.O.E., R.A., M.F., X.H., J.H., M.H., J.P.K., C.S.E., G.L., D.S., M.R.W. and K.E. improved the quality of the manuscript by critically reading and editing it.

Competing interests

The authors declare no competing interests.

References

1 Vasilopoulou, C. G. *et al.* Trapped ion mobility spectrometry and PASEF enable in-depth lipidomics from minimal sample amounts. *Nat Commun* **11**, 331, doi:10.1038/s41467-019-14044-x (2020).

2 Fauland, A. *et al.* A comprehensive method for lipid profiling by liquid chromatography-ion cyclotron resonance mass spectrometry. *J Lipid Res* **52**, 2314-2322 (2011).

3 Ovcacikova, M., Lisa, M., Cifkova, E. & Holcapek, M. Retention behavior of lipids in reversed-phase ultrahigh-performance liquid chromatography-electrospray ionization mass spectrometry. *J Chromatogr A* **1450**, 76-85, doi:10.1016/j.chroma.2016.04.082 (2016).

4 Danne-Rasche, N., Coman, C. & Ahrends, R. Nano-LC/NSI MS Refines Lipidomics by Enhancing Lipid Coverage, Measurement Sensitivity, and Linear Dynamic Range. *Anal Chem* **90**, 8093-8101, doi:10.1021/acs.analchem.8b01275 (2018).

5 Hartler, J. *et al.* Deciphering lipid structures based on platform-independent decision rules. *Nat Methods* **14**, 1171-1174, doi:10.1038/nmeth.4470 (2017).
6 Koelmel, J. P. *et al.* LipidMatch: an automated workflow for rule-based lipid identification using untargeted high-resolution tandem mass spectrometry data. *Bmc Bioinformatics* **18**, 331, doi:10.1186/s12859-017-1744-3 (2017).
7 Liebisch, G. *et al.* Shorthand notation for lipid structures derived from mass spectrometry. *J Lipid Res* **54**, 1523-1530 (2013).
8 Liebisch, G. *et al.* Lipidomics needs more standardization. *Nature Metabolism* **1**, 745-747, doi:10.1038/s42255-019-0094-z (2019).

Figure 1

Plot of retention times versus number of fatty acyl carbons for (a) triacylglycerols and (b) diacyl phosphatidylcholines in plasma. DB0 through DB8 represent the cumulative number of double bonds in the fatty acyl chains. For each individual DB assignment the relationship should be close to linear on a C18 stationary phase column.