Bioinformatics, YYYY, 0–0
doi: 10.1093/bioinformatics/xxxxx
Advance Access Publication Date: DD Month YYYY
Application Note

Structural Bioinformatics

LipidQuant 1.0: automated data processing in lipid class separation - mass spectrometry quantitative workflows

Denise Wolrab[1,#], Eva Cífková[1,#,§], Pavel Čáň[2], Miroslav Lísa[1,§], Ondřej Peterka[1], Michaela Chocholoušková[1], Robert Jirásko[1], Michal Holčapek[1,*]

[1]Department of Analytical Chemistry, University of Pardubice, Faculty of Chemical Technology, Pardubice, Czech Republic, [2]DEZA, a.s., Valašské Meziříčí, Czech Republic

#The first two authors should be regarded as Joint First Authors.

§Current address: Department of Chemistry, University of Hradec Králové, Rokitanského 62, 530 03 Hradec Králové, Czech Republic.

*To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Summary: We present the LipidQuant 1.0 tool for automated data processing workflows in lipidomic quantitation based on lipid class separation coupled with high-resolution mass spectrometry. Lipid class separation workflows, such as hydrophilic interaction liquid chromatography or supercritical fluid chromatography, should be preferred in lipidomic quantitation due to the coionization of lipid class internal standards with analytes from the same class. The individual steps in the LipidQuant workflow are explained, including lipid identification, quantitation, isotopic correction, and reporting results. We show the application of LipidQuant data processing to a small cohort of human serum samples.

Availability and implementation: The LipidQuant 1.0 is freely available at Zenodo https://zenodo.org/record/4905559 and https://holcapek.upce.cz/lipidquant.php.

Contact: Michal.Holcapek@upce.cz

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Lipids are biomolecules present in all cells with a large structural diversity. Lipid species can be divided into 8 main categories and numerous classes and subclasses, as introduced by Lipid MAPS (Fahy *et al*., 2009), together with a lipid species database containing more than 40,000 entries. Lipidomic analysis aims at the identification and quantitation of all lipids in biological samples using MS-based methods (Holčapek *et al.*, 2018). Lipid class separation approaches, such as hydrophilic interaction liquid chromatography (HILIC) or ultrahigh-performance supercritical fluid chromatography (UHPSFC), are based on the interaction of the lipid class head group with the stationary phase, which leads to the co-elution of all lipid species with the same head group. Consequently, only the molecular level of lipid species can be determined. The relation of the signal response of exogenous internal standards (IS) with known concentration for each lipid class, to the one of endogenous lipid species belonging to the same lipid class, allows the lipidomic quantitation. The co-ionization of analytes and IS leads to the same matrix effect, which is important for the quantitation of complex biological samples (Liebisch *et al.*, 2019).

The manual data processing for lipid identification and quantitation for numerous complex biological samples and hundreds of lipids is time consuming, hence the automation in data processing is desirable. In recent years, several open source lipidomic software packages were developed to facilitate automated data processing and ensure data processing reliability. Lipid Data Analyzer (Hartler *et al.*, 2010), LipidMatch (Koelmel *et al.*, 2019), and LipidCreator (Peng *et al.,* 2020) are well-recognized examples. All software packages have in common that they were developed to simplify the data processing of a preferred lipidomic method, mainly based on tandem mass spectrometry (MS/MS) for improved reliability of

lipid identification. The Lipid data analyzer (LDA) is a tool for automated lipid annotation of high-throughput data acquired with reversed-phase liquid chromatography hyphenated with MS/MS. LDA is independent of the type of mass spectrometer used and has an 3D algorithm embedded to confine peak boarders for *m/z* and retention time as well as the verification of peak annotations by the comparison of experimental and theoretical isotopic distributions of an analyte. Furthermore, LDA allows the normalization to IS, data visualization, and data export. LipidMatch is a tool for automated lipid identification of high resolution tandem mass spectra acquired by direct infusion or hyphenated with liquid chromatography using an embedded lipid library of more than 250,000 lipid species and *in silico* fragmentation library. LipidCreator is a tool for targeted lipidomics for fast method development, as transition lists and libraries can be quickly generated as well as collision energies can be optimized. In combination with Skyline, fast data processing can be performed, whereby lipid species can be identified with high confidence by the comparison of experimental and theoretical MS/MS spectra and quantitation by exporting the peak areas from Skyline.

LipidQuant 1.0 is a simple tool for automated data processing of peak lists acquired with lipid class separation hyphenated with high-resolution MS for lipid identification and quantitation including isotopic type II correction. The lipid identification is based on accurate masses, lipid class retention times, and co-elution of representative lipid class IS, but not on tandem mass spectra unlike to above-mentioned approaches.

2 Results

LipidQuant 1.0 (Fig. 1) is used for lipid identification and quantitation of molar concentrations on a molecular level. For data processing, it is necessary to import a summary table of all *m/z* features detected for individual lipid classes either by exporting corresponding intensities as a txt file from the vendor software or by using other peak picking software tools. It is important that the input summary table has in the first column the heading "*m/z*" followed by individual samples containing the intensities or other quantitative measures for each *m/z* feature (see example in ReadMe file).

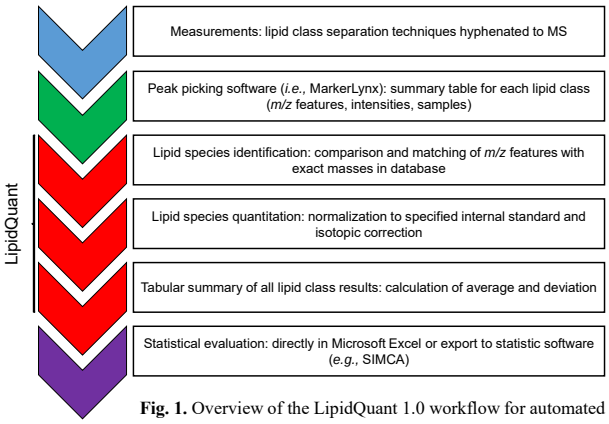

Fig. 1. Overview of the LipidQuant 1.0 workflow for automated processing of lipidomic data from lipid class separation hyphenated with high-resolution MS.

The identification is based on the comparison of the experimental *m/z* features for the lipid class with the exact *m/z* of the embedded lipid class database with a predefined mass tolerance. LipidQuant 1.0 contains 1470 lipid species entries for positive-ion mode belonging to 23 lipid classes and 1999 entries for negative-ion mode belonging to 24 lipid classes. However, LipidQuant 1.0 provides the full flexibility to modify the database, *i.e.*, by the addition of another lipid species or class. Lipid species entries are based on the molecular level, as only the sum composition of carbon atoms and double bonds in all fatty acyl chains can be assigned with lipid class separation approach.

The quantitation is achieved by dividing the signal response of the analyte by the one of the IS and multiplying with the known concentration of the IS. However, the signal response of the target lipid species can have M+2 contribution from the lipid species with one double bond more, a signal response correction has to be applied for the quantitation, called type II isotopic correction (Wang, M. *et al.,* 2017). A detailed description is provided in the Supplementary information. The use of multiple IS per lipid class allows the estimation of the quantitation error and can compensate ionization issues caused by structural differences of fatty acyl chains. LipidQuant 1.0 allows the simultaneous application of up to 3 IS for quantitation.

A detailed description of all functionalities of LipidQuant 1.0 for lipid identification and quantitation is provided in the Supplementary material and ReadMe file, as well as the application of the approach for the lipidomic quantitation of 8 lipid classes (CE, TG, DG, MG, Cer, PC, LPC, and SM) in human serum samples from healthy volunteers for UHPSFC/MS data in the positive ion mode. Results showed gender-related differences in serum lipidome, visualized by statistical analysis tools, such as the upregulation of some TG in males or some PC in females.

3 Conclusions

LipidQuant 1.0 is a freely available script written in Visual basic for application (VBA) programming language developed for lipidomic identification and molar quantitation, which can run on every computer with installed Microsoft Excel. All parameters are adjustable, such as the choice of IS, the mass tolerance window, and the extent of embedded lipid database. LipidQuant is vendor-independent, because it works with *m/z* features and their intensities in txt tables obtained by any peak picking software. The simplicity of lipid class separation approaches, in comparison to more common lipid species separation approaches typically based on reversed-phase liquid chromatography, makes the whole workflow including automated processing with LipidQuant 1.0 attractive for high-throughput applications in clinical lipidomic studies.

Funding

This work has been supported by Czech Science Foundation (No. 21-20238S).

Conflict of Interest: none declared.

References

Fahy, E. *et al.* (2009) Update of the LIPID MAPS comprehensive classification system for lipids. *J. Lipid Res.*, **50**, S9-S14.

Hartler, J. *et al.* (2010) Lipid Data Analyzer: unattended identification and quantitation of lipids in LC-MS data. *Bioinformatics*, **27**, 572-577.

Holčapek, M. *et al.* (2018) Lipidomic Analysis. *Anal. Chem.*, **90**, 4249-4257.

Koelmel, J. P. *et al.* (2019) Software tool for internal standard based normalization of lipids, and effect of data-processing strategies on resulting values. Bioinformatics, **20**, 217.

Liebisch, G. *et al.* (2019) Lipidomics needs more standardization. *Nat. Metab.*, **1**, 745-747.

Peng, B. *et al.* (2020) LipidCreator workbench to probe the lipidomic landscape. *Nat. Com.*, **11**, 2057.

Wang, M. *et al.* (2017) Selection of internal standards for accurate quantification of complex lipid species in biological extracts by electrospray ionization mass spectrometry - What, how and why? *Mass Spectrom. Rev.*, **36**, 693–714