

## Lipidomics in biomedical research – practical considerations

Tuulia Hyötyläinen<sup>1</sup>, Linda Ahonen<sup>2</sup>, Päivi Pöhö<sup>3</sup>, Matej Orešič<sup>4</sup>

<sup>1</sup>Department of Chemistry, Örebro University, 702 81 Örebro, Sweden.

<sup>2</sup>Steno Diabetes Center A/S, DK-2820 Gentofte, Denmark.

<sup>3</sup>Faculty of Pharmacy, University of Helsinki, FI-00014 Helsinki, Finland

<sup>4</sup>Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, FI-20520 Turku, Finland.

Correspondence:

Matej Orešič, Turku Centre for Biotechnology, Tykistokatu 6, FI-20520 Turku, Finland.

Tel.: +358 44 972 6094, Email: [matej.oresic@utu.fi](mailto:matej.oresic@utu.fi)

Character count: 793 (abstract), 17,789 (manuscript)

## **ABSTRACT**

Lipids have many central physiological roles including as structural components of cell membranes, energy storage sources and intermediates in signaling pathways. Lipid-related disturbances are known to underlie many diseases and their co-morbidities. The emergence of lipidomics has empowered researchers to study lipid metabolism at the cellular as well as physiological levels at a greater depth than was previously possible. The key challenges ahead in the field of lipidomics in medical research lie in the development of experimental protocols and *in silico* techniques needed to study lipidomes at the systems level. Clinical questions where lipidomics may have an impact in healthcare settings also need to be identified, both from the health outcomes and health economics perspectives.

## **Keywords**

Biomarkers, data processing, lipidomics, mass spectrometry, metabolomics, quality control

## **Highlights**

- Lipidomics is a key platform for the study of lipid metabolism in health and disease.
- Multiple suitable techniques for the global measurement of molecular lipids exist.
- New experimental protocols are needed to study lipidomes at physiological level.
- Better *in silico* tools are needed to study lipids at the systems level.

## **1. INTRODUCTION**

Over the past decade, lipidomics emerged as a key approach aimed at elucidating lipid-related pathways and mechanisms underlying the cellular function. Additionally, due to the undeniable role of lipids in the maintenance of system homeostasis at the cellular as well as physiological levels, lipidomics also emerged as a powerful tool for identifying the key processes and related molecular signatures associated with progression to and protection from many diseases [1, 2]. Biological importance of lipids is not only due to their roles in biochemical and signaling pathways, but also due to them being part of lipid-containing ensembles such as cell membranes and lipid particles (e.g., lipoproteins and exosomes). In order to understand the roles of lipids in specific pathologies, it is therefore essential to adopt a systems approach which accounts for spatial and dynamic complexity of lipid metabolism [3].

Much of the debate concerning lipidomics applications in health and disease has been about the advantages or disadvantages of specific technologies for the measurement of lipids. However, the current methods such as based on shotgun mass spectrometry (MS) and liquid chromatography (LC)-MS, each with its own advantages and disadvantages, are all applicable for the studies of biological systems including in clinical settings, as long as they are properly optimized and their pro-s and con-s are understood.

Here we focus on selected practical considerations when choosing and developing lipidomics methods as well as highlight key considerations and challenges.

## **2. SAMPLE PREPARATION FOR LIPIDOMIC ANALYSIS**

Regardless of the analytical method used for lipidomics data acquisition, lipids need to be extracted from the biological samples. Sample preparation has a major impact both on the quality as well as on the sample throughput in lipidomics [4]. Different types of sample matrices may also require different types of sample preparation protocols.

There are several studies comparing different sample preparation methods, yet the results have not always been consistent [5-8]. Criteria for the selection of the most suitable sample preparation protocol should include assessment of the lipid coverage and recovery, simplicity of the procedure, protein removal efficiency, repeatability of lipids measurements, costs and feasibility for automation. Among the most commonly used lipid extraction methods, the protocols based on chloroform:methanol (CHCl<sub>3</sub>:MeOH) [9, 10] can be problematic with robotic pipetting systems, as the extract is in the lower phase, and the collection of the extract through the aqueous phase and through the protein layer may cause contamination and clogging. The methyl-*tert*-butyl ether (MTBE):MeOH protocol [11], on the other hand, has a high sample-to-solvent ratio, making it more challenging to use with 96-well plates [5].

Sensitivity of MS instruments has improved considerably over the past years, to the point that this is no longer a major issue as compared to the issues related to sampling and sample preparation. Much of the analytical variation and loss of lipids occurs at these initial stages of lipidomic analysis, which is particularly true when studying lipid mediators and oxidized lipids. When analyzing the human samples (biopsies, blood), where time of sample exposure at room temperature after the sampling often cannot be fully controlled for, there is a risk that the measured concentrations of lipids may not truly reflect their actual endogenous levels but

are instead the artifacts of sampling. More systematic studies are needed that address the effect of sampling (and sample types) on lipid concentrations [12], and ultimately consensus should be sought about the optimal sampling procedures for specific sample types and lipid classes.

### 3. QUANTIFICATION OF LIPIDS – DIFFERENT APPROACHES TO LIPIDOMICS

The term ‘quantitative analysis’ is often used rather loosely in life sciences. Terms such as ‘quantitation’ and ‘absolute quantitation’ have been applied in the fields of proteomics and lipidomics, even though the analyses applied were strictly speaking semi-quantitative. For an analytical chemist, quantitative analysis means that *each* compound analyzed is quantified against a structurally equivalent compound which is used as an internal standard (i.e., isotope labeled standards each corresponding to each of the individual molecules measured), and calibration curves for each separate compound over the entire concentration range measured. For quantitative analysis, FDA guidelines state that calibration curve should be generated for each analyte in the sample [13]. Therefore, the global lipidomic analyses do not fulfill the criteria of being fully quantitative analysis (neither the LC-MS nor the shotgun based methods).

In the field of metabolomics, it is still a common practice to use normalization procedures which are not based on internal standards, e.g., by using the sum of total signal. This approach is, however, not recommended for lipidomics and it is also problematic for metabolomics applications in general. As discussed in detail earlier [14], the total amount of lipids in tissues and biofluids may vary considerably across different study groups. Normalization based on data-derived measures such as total signal may thus inevitably skew the results.

The use of lipid-class-specific standards for normalization and quantitation is not fully accurate either, because different lipids typically have quantitatively different MS responses. For example, for phospholipids, the MS response for both saturated and unsaturated lipids decreases with increasing acyl-chain length. In addition, the differences in overall lipid concentrations in the study samples further enhance this phenomenon. With increasing number of standard compounds for the analysis of lipids being commercially available, the accuracy of quantification is bound to improve considerably in the future, although strictly speaking the global lipidomics approaches are likely to remain semi-quantitative. Systematic studies are needed which will examine the effect of sample matrix, e.g., across a wide range of physiological conditions of interest, on measured lipid levels.

Both the shotgun lipidomics and LC-MS approach have their own advantages and disadvantages [4, 15]. In the shotgun approach, the main advantage is that the sample is infused to the MS at a constant concentration and thus, the inevitable matrix effect (mainly ion suppression) is relatively constant, and internal standard normalization may work well if the lipid profiles between the study groups do not vary too considerably. In LC-MS, lipids are separated prior to the introduction to MS, thus reducing the matrix effects and improving the sensitivity, however, because the internal standards do not elute at the same time as most of the lipids, the normalization may work very well for part of the compounds and less so for the others. This is, however, a slightly simplistic view of the matrix effects. In shotgun MS, the total lipid concentration affects the ionization process. At high(er) concentrations, lipids can aggregate, and the aggregates cannot be efficiently ionized or reliably detected. Thus, in the

aggregation state, the ionization efficiency becomes dependent on the lipid species (the acyl-chain length, number of double bonds). The lipid concentration at which aggregation occurs depends on the solvent. For example, in CHCl<sub>3</sub>:MeOH, which is typically used for lipid extraction, the critical lipid concentration varies with the proportion of methanol, from ca. 10 pmol mL<sup>-1</sup> (CHCl<sub>3</sub>:MeOH, 1:2) to 200 pmol mL<sup>-1</sup> in (CHCl<sub>3</sub>:MeOH, 2:1). The challenge of finding the optimal concentration is that the total lipid concentrations in biological samples can vary considerably. For example, the total lipid concentration in human plasma is ca. 8000 nmol mL<sup>-1</sup>. However, some lipid concentrations can be clearly higher, e.g., in subgroups of obese individuals where the plasma triacylglycerols are elevated. Dilution of the samples is therefore typically needed in shotgun approaches. However, it is not always possible to determine the correct degree of dilution before the experiments.

#### 4. QUALITY CONTROL

Strict quality control (QC) in lipidomics is crucial, particularly when a large number of samples are analyzed. In lipidomics, repeatability is one of the most important criteria to assure the reliability of the various analytical assays. It is essential that internal standards are used for normalization, the samples are prepared and analyzed in random order, and sufficient amount of different types of quality control samples and blanks are analyzed together with the samples (**Table 1**).

Large-scale lipidomic analyses are typically performed in batches, and even under repeatable conditions, instrumental variation in MS techniques applied might arise for a number of reasons. These include for example drifts in sensitivity, ionization efficiency, and gradual

changes in column performance (in LC-MS) over the analysis periods. Batch effects can be corrected for by using different methodologies [16], and there is still no consensus which is the best method for batch effect correction. In our research, we consider that batch correction should be applied with caution. If no deviations from the accepted QC cut-off values (RSD in control samples) occur, the batch correction may not be needed. If there are some deviations and there is a clear batch pattern over the course of the analysis, correction may be however merited. The choice of the correction ultimately depends on the nature of the batch effect.

## **5. LIPIDOMICS DATA PROCESSING AND ANALYSIS**

Data processing is a crucial step, which transforms raw lipidomics data into an interpretable dataset amenable to statistical and pathway analysis. Several tools are available for lipidomics data processing, both for shotgun and LC-MS approaches [17, 18]. In addition to the solutions developed by the research community, the MS instrument vendors have also started to offer increasingly sophisticated solutions for processing and analysis of lipidomics data. The choice of the data processing solution may depend on many factors, including local set-up of the lipidomics platform, software expertise and ambition as well as familiarity with a specific methodology. It is essential that any data processing solution is thoroughly evaluated and tested as part of lipidomics method development and validation, before putting it to use. Open-source packages such as XCMS [19, 20] and MZmine [21, 22] may be initially more demanding for non-expert users, yet they may offer more flexibility as compared to alternative 'black box' software solutions.



Similarly as in metabolomics and other 'omics' approaches, high dimensionality of lipidomics data needs to be taken into account when analyzing them [17, 23]. When analyzing the data at the individual lipid level, rate of type I errors in null hypothesis testing needs to be taken into account when conducting multiple comparisons, such as by controlling for False Discovery rate [24]. However, lipids commonly exert their function as ensembles. The bioinformatics challenge is therefore how to best account for such a high degree of co-regulation [3, 25].

## 6. LIPID IDENTIFICATION

In most cases, the identification of lipids is generally not as challenging as e.g. for the polar metabolites, because lipids comprise common structural blocks that can be detected by MS/MS (or more broadly MS<sup>n</sup>) analysis. This makes lipid identification also amenable to automated computational approaches [26]. Current high-resolution MS instruments in most cases afford reliable identification of major lipids (apart from double bond and acyl chain position).

Much of the ambiguity concerning the reported lipid identities and their amounts comes from the differences in lipid annotation, and there is a clear need for consensus. The proposed shorthand nomenclature by Liebisch *et al.* provides an excellent starting point for this purpose [27]. The ambiguity in lipid annotation may also affect the reported lipid concentrations. For example, in LC-MS lipidomics, multiple lipid species may be reported containing the same carbon number and double bond count, based on their chromatographic separation. While some may prefer to determine the total amount of lipid corresponding to a specific carbon number and double bond count, alternatively one may also treat each individual lipid (corresponding

to a specific retention time) separately. The reported concentration of lipids in such ambiguous cases may therefore be considerably different, depending on the approach taken.

## **7. MODELING OF LIPID METABOLISM**

Elucidation of lipid metabolism and its complex dynamics at physiological or cellular levels can only be achieved in dedicated experimental settings which are not feasible in most clinical lipidomics studies. Nevertheless, functional studies even if limited in sample size, such as by using stable isotope tracers or focusing on lipidomic profiling in isolated lipoprotein fractions [28-30], are highly important if one is to interpret lipidomic profiles in the physiological context.

One important emerging tool for modelling metabolic networks including lipid metabolism is the constraint-based modeling (CBM) approach. CBM relies on simpler physical-chemical constraints and has shown a remarkable predictive signal in the analysis of metabolic systems on a genome scale [31, 32]. Genome-scale reconstruction collects all of the relevant metabolic information of an organism and compiles it in a mathematical model, which connects genes (e.g., those expressing specific enzymes) with metabolic pathways. Constraining the genome scale network reconstruction provides the allowable network space, i.e., the space of feasible flux distributions. While genome-scale CBM is increasingly being used as an integrative modeling tool in metabolomics including in clinical setting [33], the approach has not yet been widely adopted in lipidomics. This is mainly because the current genome scale metabolic reconstructions still poorly annotate lipid metabolism at the molecular level – although progress has been made [34]. Advances in functional lipidomics and metabolic reconstructions are likely to lead to much improved metabolic models in the future.

## 8. IMPACT OF LIPIDOMICS IN CLINICAL RESEARCH

In clinical research, the main challenge is not in the technologies applied, but rather in identifying the problems where the identification of lipid-related markers may have an impact. Since lipid homeostasis is fundamental to maintain health, specific lipid disturbances may be an underlying link and shared across multiple co-morbid conditions behind specific diseases. Lipidomics may therefore be a valuable tool for identifying particularly vulnerable co-morbid sub-populations which would benefit from a specific therapy. For example, we have recently shown that those first episode psychotic patients who are at the highest risk of rapid weight gain are characterized by a specific lipidomic profile at baseline [35], similar to the one observed earlier in patients with non-alcoholic fatty liver disease (NAFLD) [36] and in those at risk of developing T2D [37]. Those patients may therefore benefit from a combined antipsychotic and anti-obesity treatment, a therapeutic strategy that has already been considered due to the significant public health burden of metabolic co-morbidities that are associated with the psychotic disorders [38].

When considering clinical applications of lipidomics, the obvious question is also if and where lipidomics can have a public health impact? While undoubtedly lipidomics research has already provided valuable and novel pathophysiological insights as well as biomarker candidates in many diseases, and thus also opened novel therapeutic avenues, translation of such research outcomes into applications in healthcare settings is an entirely different challenge. Particularly in diseases affecting large fraction of the population such as in type 2 diabetes (T2D), health economics is also an important consideration in addition to the health outcomes. Since lifestyle

plays a major role in acquiring risk of T2D, much of the relevant data for risk assessment may be easily obtained by existing routine clinical tests and by real-time monitoring of lifestyle (e.g., diet, exercise, sleep *via* mobile phone applications). If lipidomics is to deliver in this domain, the lipidomic test has to be inexpensive and its added value needs to be demonstrated in a prospective setting with respect to health economics and outcomes. The 'regulatory paths' also need to be identified in a disease-specific manner, which would clarify the steps needed to translate research findings into the applications in the clinic – which may also require developments of cost effective and quantitative assays for the analysis of specific lipids.

## 9. CONCLUDING REMARKS

Lipidomics led to many advances in the studies of health and disease over the past decade, providing valuable insights into the mechanisms and molecular signatures behind specific diseases and their co-morbidities. The field has matured to the point that the introduction of lipidomics into clinical practice is being considered and debated. While much of the discussion has so far focused on specific technologies for the measurement of lipids at the molecular level, the key challenges ahead lie elsewhere. These non-exclusively include identification of clinical questions where lipidomics may have an impact in healthcare settings as well as development of experimental protocols and *in silico* techniques needed to study lipidomes at the systems level.

## Acknowledgments

The authors thank Thomas Sparholt Dyrland (Novo Nordisk A/S, Måløv, Denmark) for his contribution to the development of automated quality control system for lipidomics.

## REFERENCES

- [1] M. Oresic, V.A. Hänninen, A. Vidal-Puig, Lipidomics: a new window to biomedical frontiers, *Trends Biotechnol.*, 26 (2008) 647-652.
- [2] M. Oresic, Obesity and psychotic disorders: uncovering common mechanisms through metabolomics, *Dis Model Mech*, 5 (2012) 614-620.
- [3] T. Hyotylainen, M. Oresic, Systems biology strategies to study lipidomes in health and disease, *Prog Lipid Res*, 55 (2014) 43-60.
- [4] T. Hyotylainen, M. Oresic, Bioanalytical techniques in nontargeted clinical lipidomics, *Bioanalysis*, 8 (2016) 351-364.
- [5] L. Lofgren, M. Stahlman, G.B. Forsberg, S. Saarinen, R. Nilsson, G.I. Hansson, The BUME method: a novel automated chloroform-free 96-well total lipid extraction method for blood plasma, *J Lipid Res*, 53 (2012) 1690-1700.
- [6] R.M. Pellegrino, A. Di Veroli, A. Valeri, L. Goracci, G. Cruciani, LC/MS lipid profiling from human serum: a new method for global lipid extraction, *Analytical and Bioanalytical Chemistry*, 406 (2014) 7937-7948.
- [7] A. Reis, A. Rudnitskaya, G.J. Blackburn, N. Mohd Fauzi, A.R. Pitt, C.M. Spickett, A comparison of five lipid extraction solvent systems for lipidomic studies of human LDL, *J Lipid Res*, 54 (2013) 1812-1824.
- [8] M.H. Sarafian, M. Gaudin, M.R. Lewis, F.-P. Martin, E. Holmes, J.K. Nicholson, M.-E. Dumas, Objective Set of Criteria for Optimization of Sample Preparation Procedures for Ultra-High Throughput Untargeted Blood Plasma Lipid Profiling by Ultra Performance Liquid Chromatography–Mass Spectrometry, *Analytical Chemistry*, 86 (2014) 5766-5774.
- [9] J. Folch, M. Lees, G.H. Sloane Stanley, A simple method for the isolation and purification of total lipides from animal tissues, *J Biol Chem*, 226 (1957) 497-509.
- [10] E.G. Bligh, W.J. Dyer, A rapid method of total lipid extraction and purification, *Can J Biochem Physiol*, 37 (1959) 911-917.
- [11] V. Matyash, G. Liebisch, T.V. Kurzchalia, A. Shevchenko, D. Schwudke, Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics, *Journal of Lipid Research*, 49 (2008) 1137-1146.

- [12] B. Jorgenrud, S. Jantti, I. Mattila, P. Poho, K.S. Ronningen, H. Yki-Jarvinen, M. Oresic, T. Hyotylainen, The influence of sample collection methodology and sample preprocessing on the blood metabolic profile, *Bioanalysis*, 7 (2015) 991-1006.
- [13] US Food and Drug Administration, Guidance for Industry - Bioanalytical Method Validation, 2013.  
<http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm368107.pdf>
- [14] M. Sysi-Aho, M. Katajamaa, L. Yetukuri, M. Oresic, Normalization method for metabolomics data using optimal selection of multiple internal standards, *BMC Bioinformatics*, 8 (2007) 93.
- [15] K. Yang, X. Han, *Lipidomics: Techniques, Applications, and Outcomes Related to Biomedical Sciences*, *Trends Biochem Sci*, 41 (2016) 954-969.
- [16] W.B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J.D. Knowles, A. Halsall, J.N. Haselden, A.W. Nicholls, I.D. Wilson, D.B. Kell, R. Goodacre, C. Human Serum Metabolome, Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry, *Nat Protoc*, 6 (2011) 1060-1083.
- [17] M. Oresic, Informatics and computational strategies for the study of lipids, *Biochim Biophys Acta*, 1811 (2011) 991-999.
- [18] S. Castillo, P. Gopalacharyulu, L. Yetukuri, M. Orešič, Algorithms and tools for the preprocessing of LC-MS metabolomics data, *Chemometr. Intell. Lab. Syst.*, 108 (2011) 23–32.
- [19] C.A. Smith, E.J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification, *Anal Chem*, 78 (2006) 779-787.
- [20] R. Tautenhahn, G.J. Patti, E. Kalisiak, T. Miyamoto, M. Schmidt, F.Y. Lo, J. McBee, N.S. Baliga, G. Siuzdak, metaXCMS: second-order analysis of untargeted metabolomics data, *Anal Chem*, 83 (2011) 696-700.
- [21] T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC Bioinformatics*, 11 (2010) 395.

- [22] M. Katajamaa, J. Miettinen, M. Oresic, MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data, *Bioinformatics*, 22 (2006) 634-636.
- [23] D. Broadhurst, D. Kell, Statistical strategies for avoiding false discoveries in metabolomics and related experiments, *Metabolomics*, 2 (2006) 171-196.
- [24] J.D. Storey, A direct approach to false discovery rates, *J. R. Stat. Soc. B*, 64 (2002) 479-498.
- [25] P.S. Niemela, S. Castillo, M. Sysi-Aho, M. Oresic, Bioinformatics and computational methods for lipidomics, *J Chromatogr B Analyt Technol Biomed Life Sci*, 877 (2009) 2855-2862.
- [26] L. Yetukuri, M. Katajamaa, G. Medina-Gomez, T. Seppanen-Laakso, A. Vidal-Puig, M. Oresic, Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis, *BMC Syst Biol*, 1 (2007) 12.
- [27] G. Liebisch, J.A. Vizcaino, H. Kofeler, M. Trotsmuller, W.J. Griffiths, G. Schmitz, F. Spener, M.J. Wakelam, Shorthand notation for lipid structures derived from mass spectrometry, *J Lipid Res*, 54 (2013) 1523-1530.
- [28] A. Kontush, M. Lhomme, M.J. Chapman, Unraveling the complexities of the HDL lipidome, *J Lipid Res*, 54 (2013) 2950-2963.
- [29] A. Kontush, M.J. Chapman, Lipidomics as a tool for the study of lipoprotein metabolism, *Curr Atheroscler Rep*, 12 (2010) 194-201.
- [30] A. Kotronen, V.R. Velagapudi, L. Yetukuri, J. Westerbacka, R. Bergholm, K. Ekroos, J. Makkonen, M.R. Taskinen, M. Oresic, H. Yki-Jarvinen, Serum saturated fatty acids containing triacylglycerols are better markers of insulin resistance than total serum triacylglycerol concentrations, *Diabetologia*, 52 (2009) 684-690.
- [31] M.A. Oberhardt, B.O. Palsson, J.A. Papin, Applications of genome-scale metabolic reconstructions, *Mol Syst Biol*, 5 (2009) 320.
- [32] J. Nielsen, Systems Biology of Metabolism: A Driver for Developing Personalized and Precision Medicine, *Cell Metab*, 25 (2017) 572-579.
- [33] T. Hyotylainen, L. Jerby, E.M. Petaja, I. Mattila, S. Jantti, P. Auvinen, A. Gastaldelli, H. Yki-Jarvinen, E. Ruppin, M. Oresic, Genome-scale study reveals reduced metabolic adaptability in patients with non-alcoholic fatty liver disease, *Nat Commun*, 7 (2016) 8994.
- [34] I. Thiele, N. Swainston, R.M. Fleming, A. Hoppe, S. Sahoo, M.K. Aurich, H. Haraldsdottir, M.L. Mo, O. Rolfsson, M.D. Stobbe, S.G. Thorleifsson, R. Agren, C. Bolling, S. Bordel, A.K. Chavali, P. Dobson, W.B. Dunn, L. Endler, D. Hala, M. Hucka, D. Hull, D. Jameson, N.

Jamshidi, J.J. Jonsson, N. Juty, S. Keating, I. Nookaew, N. Le Novere, N. Malys, A. Mazein, J.A. Papin, N.D. Price, E. Selkov, Sr., M.I. Sigurdsson, E. Simeonidis, N. Sonnenschein, K. Smallbone, A. Sorokin, J.H. van Beek, D. Weichart, I. Goryanin, J. Nielsen, H.V. Westerhoff, D.B. Kell, P. Mendes, B.O. Palsson, A community-driven global reconstruction of human metabolism, *Nat Biotechnol*, (2013).

[35] T. Suviataival, O. Mantere, T. Kieseppa, I. Mattila, P. Poho, T. Hyotylainen, J. Suvisaari, M. Oresic, Serum metabolite profile associates with the development of metabolic co-morbidities in first-episode psychosis, *Transl Psychiatry*, 6 (2016) e951.

[36] M. Oresic, T. Hyotylainen, A. Kotronen, P. Gopalacharyulu, H. Nygren, J. Arola, S. Castillo, I. Mattila, A. Hakkarainen, R.J. Borra, M.J. Honka, A. Verrijken, S. Francque, P. Iozzo, M. Leivonen, N. Jaser, A. Juuti, T.I. Sorensen, P. Nuutila, L. Van Gaal, H. Yki-Jarvinen, Prediction of non-alcoholic fatty-liver disease and liver fat content by serum molecular lipids, *Diabetologia*, 56 (2013) 2266-2274.

[37] E.P. Rhee, S. Cheng, M.G. Larson, G.A. Walford, G.D. Lewis, E. McCabe, E. Yang, L. Farrell, C.S. Fox, C.J. O'Donnell, S.A. Carr, R.S. Vasani, J.C. Florez, C.B. Clish, T.J. Wang, R.E. Gerszten, Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans, *J Clin Invest*, 121 (2011) 1402-1411.

[38] V.A. de Silva, C. Suraweera, S.S. Ratnatunga, M. Dayabandara, N. Wanniarachchi, R. Hanwella, Metformin in prevention and treatment of antipsychotic induced weight gain: a systematic review and meta-analysis, *BMC Psychiatry*, 16 (2016) 341.

[39] O. Quehenberger, A.M. Armando, A.H. Brown, S.B. Milne, D.S. Myers, A.H. Merrill, S. Bandyopadhyay, K.N. Jones, S. Kelly, R.L. Shaner, C.M. Sullards, E. Wang, R.C. Murphy, R.M. Barkley, T.J. Leiker, C.R. Raetz, Z. Guan, G.M. Laird, D.A. Six, D.W. Russell, J.G. McDonald, S. Subramaniam, E. Fahy, E.A. Dennis, Lipidomics reveals a remarkable diversity of lipids in human plasma, *J. Lipid Res.*, 51 (2010) 3299-3305.



**Table 1.** Quality control and standard samples for lipidomics.

<b>Sample</b>	<b>Type</b>	<b>Use</b>	<b>Frequency</b>
Quality control sample	Pooled sample with an aliquot of each sample or surrogate sample with similar lipid composition than the samples	For quality monitoring, correction of possible batch effects	After every 6-12 study samples*
Certified reference material	SRM 1950 - Metabolites in Frozen Human Plasma**	Allows comparison of results in different studies	Triplicate in each batch
Blank sample	Blank sample treated in the same way as all the samples	Shows the possible contaminants from sample preparation and analysis	After every 6-12 samples
Pure standards	Standards in solvent, including ISTDs	Shows the performance of the instrument	After every 6-12 samples
Calibration samples	Six different concentrations of standards, with internal standards	For (semi)quantitation	In the beginning, middle and end of a batch in large series

\*The minimum number of QC samples to ensure proper control of the assay should be at least 5 % of the number of unknown samples or a total of six QC samples, whichever is greater (FDA).

\*\*Quantitative levels of over 500 different lipid molecular species have been reported [39].