Lipidomics: levels of determination, data processing and interpretation

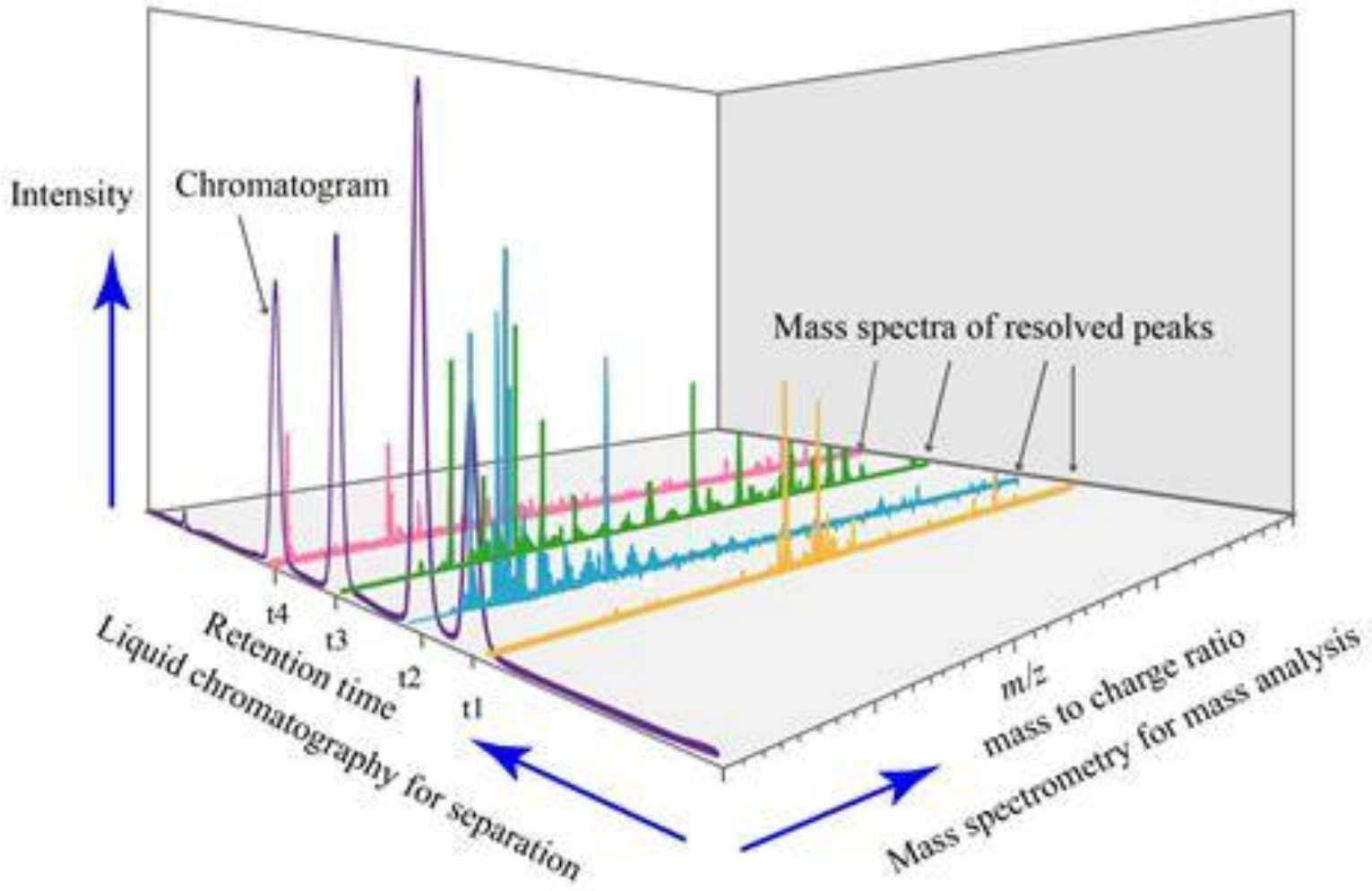
Metabolismo y Enfermedades Metabólicas

Máster en Investigación Biomédica

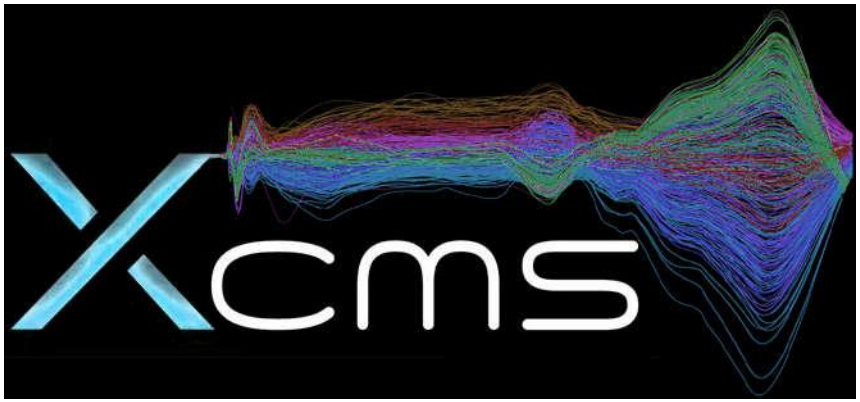
Universidad de Valladolid

Dr. David Balgoma
Instituto de Biomedicina y Genética Molecular (IBGM)
Universidad de Valladolid

MS lipidomics: Multidimensional data



MS lipidomics: Data pre-processing

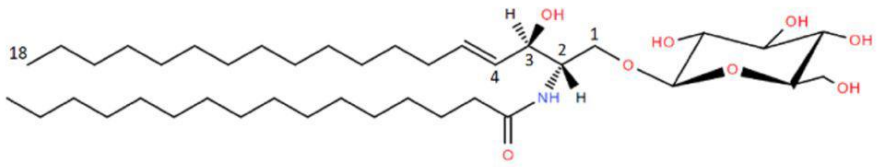


- Injection alignment
- Feature detection (RT and m/z)
- Result: Table with features and signal (arbitrary units)

			Signal of the feature in the every sample							
ID	**mzmed**	**rtmed**	1	2	3	4	5	6	8	9
Feature.1638	553.4056	363.0580	1.16	0.62	0.51	16.04	15.01	1.51	1.73	1.80
Feature.2787	664.3985	363.5940	0.85	1.47	1.29	3.52	5.51	0.72	2.10	1.80
Feature.1627	552.4026	363.6915	0.85	0.57	1.09	8.07	4.97	0.71	3.27	1.45
Feature.2781	663.8966	363.8280	1.35	1.26	0.92	7.42	7.22	1.27	2.84	1.87
Feature.2883	673.4121	363.8405	1.45	0.96	0.91	7.16	10.23	1.25	2.12	2.13
Feature.2805	665.9011	363.9800	0.52	0.76	1.77	21.50	2.76	0.43	1.54	1.09

LIPID MAPS levels of determination

Analysis	Annotation
MS1: *m/z* 700.6 [M+H]⁺ MS2: *m/z* 538.5 assumption of even hydrocarbon chains	**Hex-Cer 34:1;O2** ***Species level***
MS1: *m/z* 700.5722 [M+H]⁺ MS2: *m/z* 538.5	**Hex-Cer 34:1;O2** ***Species level***
MS1: *m/z* 700.6 [M+H]⁺ MS2: *m/z* 264.3, 538.5	**Hex-Cer 18:1;O2/16:0** ***Molecular species level***
MS1: *m/z* 700.6 [M+H]⁺ MS2: *m/z* 264.3 HILIC-Separation from isomeric Gal-Cer	**Glc-Cer 18:1;O2/16:0** ***Molecular species level***
MS1: *m/z* 700.5722 [M+H]⁺ MS2: *m/z* 252.3, 264.3, 282.3, 538.5, 682.5 HILIC-Separation from isomeric Gal-Cer Proof of DB position by OzID	**Glc-Cer 18:1(4);OH/16:0** ***Structure defined level***



Glc-Cer(1) 18:1(4E);3OH[R]/16:0[2S]

Complete Structure level

DOI 10.1194/jlr.S120001025

Fragmentation of lipids

Specifically, **depending on the lipid class,** different adducts are produced, resulting in characteristic product ions or neutral losses.

Lipid family	Adduct	Family fragment	# fragments FAs
PC (diacyl and ether)	$[M+H]^+$	m/z 184	
	$[M-CH_3]^-$ (methyl loss)		-Diacyl 2 FAs -Ether 1 FA
PE (diacyl and ether)	$[M+H]^+$	NL 141 units of m/z	
	$[M-H]^-$	m/z 196	-Diacyl 2 FAs -Ether 1 FA
PS	$[M+H]^+$	NL 185 units of m/z	
	$[M-H]^-$	NL 87 units of m/z	2 FAs
PI	$[M-H]^-$	m/z 241; also m/z 153	2 FAs
PA	$[M-H]^-$	m/z 153	2 FAs

Fragmentation of lipids

In negative mode, **fatty acids** produce specific ions depending on their number of carbons and unsaturations.

Acyl chain	Number of carbons	Number of unsaturations	m/z	Alledged lipid
16:0	16	0	255	Palmitate
18:0	18	0	283	Stearate
18:1	18	1	281	Oleate
20:4	20	4	303	Arachidonate
20:5	20	5	301	Eicosapentanoate
22:6	22	6	327	Docosahexanoate

Fragmentation of lipids

The combination of the presence of different fragments allows us to determine two levels of phospholipid identification according to LIPID MAPS:

Species Level: The phospholipid is identified by its family (polar head group in phospholipids) and the total number of carbons and unsaturations (double bonds) in the fatty acids. E.g. PE 38:4

Molecular Species Level: The phospholipid is identified by its family (polar head group in phospholipids) and the specific number of carbons. E.g. PE 18:0_20:4

Solving a puzzle

- Analysis in positive mode (low resolution)
  - + RT = 10.5 min
  - + m/z = 768.6

- Analysis in negative mode (low resolution)
  - + RT = 10.55 min
  - + m/z = 766.5

- [m/z in pos mode] – [m/z in neg mode] =
  - 2.0146 ≈ 2 * 1.0078
  - + Hydrogen H has mass 1.0078
  - + **Suggests $[M+H]^+$ and $[M-H]^-$**
  - + This is typical of **phospholipids such as ethanolamine or serine**, which present the [M+H]+ adduct in positive mode and the [M-H]- adduct in negative mode.

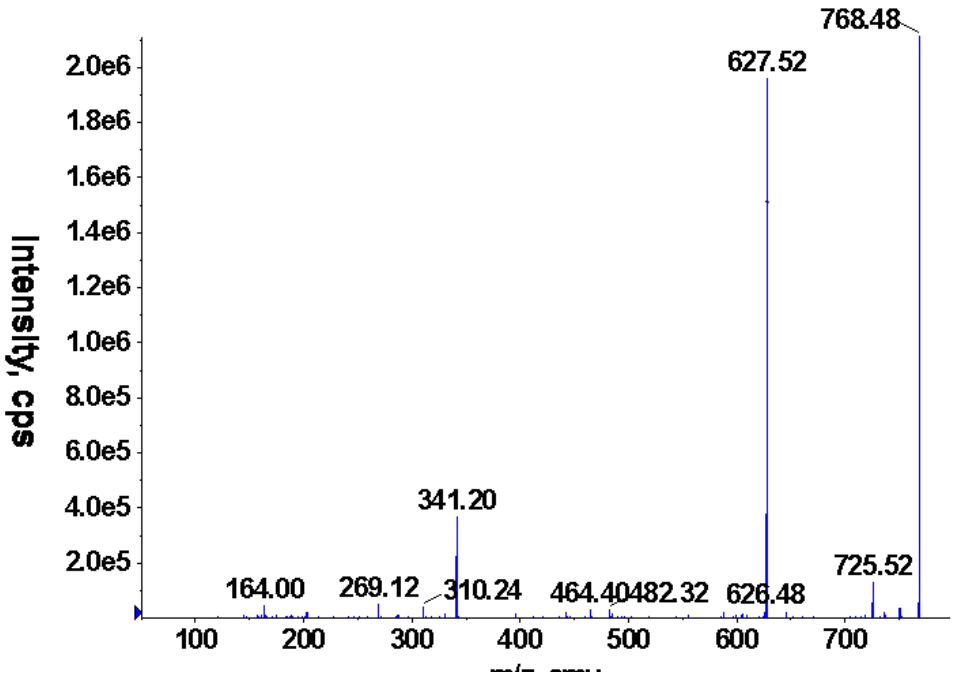
Solving a puzzle

- Search on LIPID MAPS
  - Positive m/z = 768.6
  - Negative m/z = 766.5

https://lipidmaps.org/resources/tools/bulk-structure-search/create?database=LMSD

Solving a puzzle

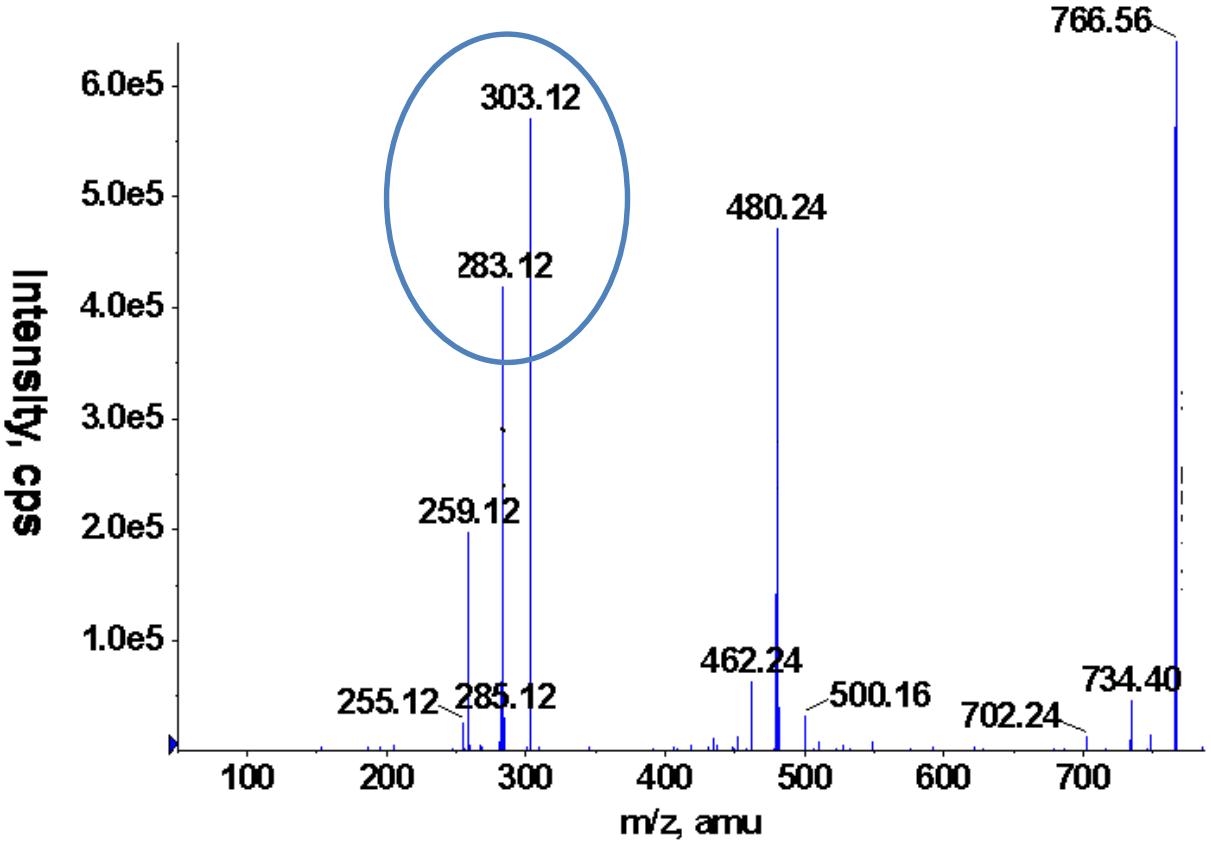
- **MS^2 in positive mode**
  - 768.48 – 627.52 = 140.96 ≈ 141



- **NL of 141 is specific of PE: PE confirmed!**
- **Species level: PE 38:4; in old notation PE(38:4)**

Solving a puzzle

- **MS^2 in negative mode**

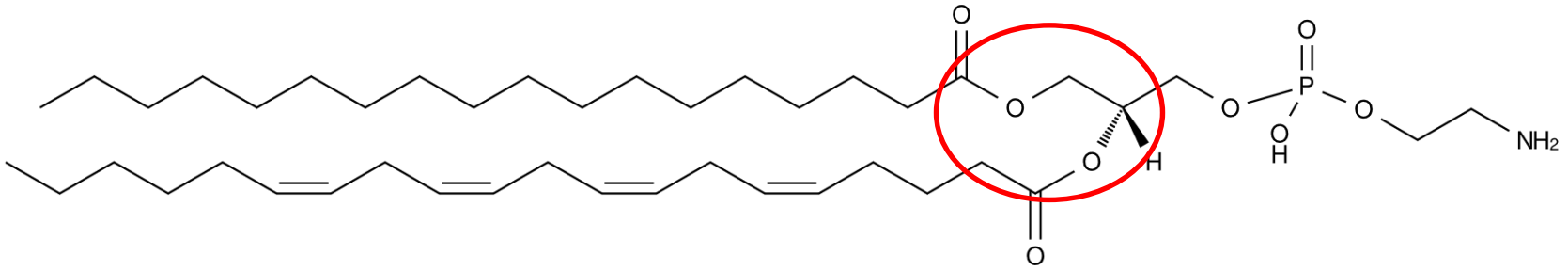


- **Molecular Species Level:**
  - **PE 18:0_20:4; PE(18:0/20:4) in old notation**

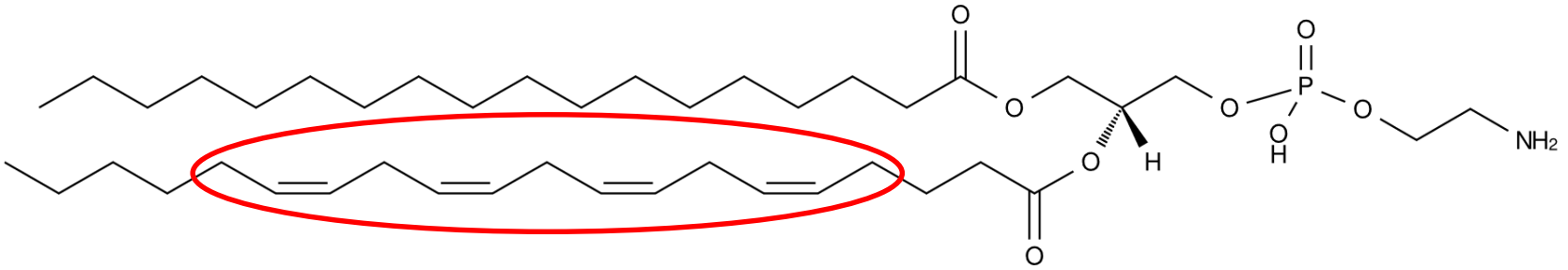
Solving a puzzle

- Full identification requires other techniques to identify

+ Regioisomer (position FAs in glycerol)



+ Position and stereochemistry of the double bonds



MS lipidomics: Types of quantification

Total Quantification

Total quantification provides an **exact measure of an analyte's concentration in a sample**, typically expressed in absolute units like micrograms or millimoles per liter.

This method requires **internal and external standards** with known concentrations to construct a calibration curve.

Not feasible in lipidomics because of the myriad of lipids

MS lipidomics: Types of quantification

Semi-Quantification:

Semi-quantification provides an approximate estimate of an analyte's concentration. This method uses an **internal standard or a structurally similar analyte and compares the analyte's intensity to that of the standard to estimate concentration**.

For example, in lipidomics, we assume that PC(16:0/20:4) and the internal standard PC(17:0/17:0) behave the same in extraction and MS detection

MS lipidomics: Types of quantification

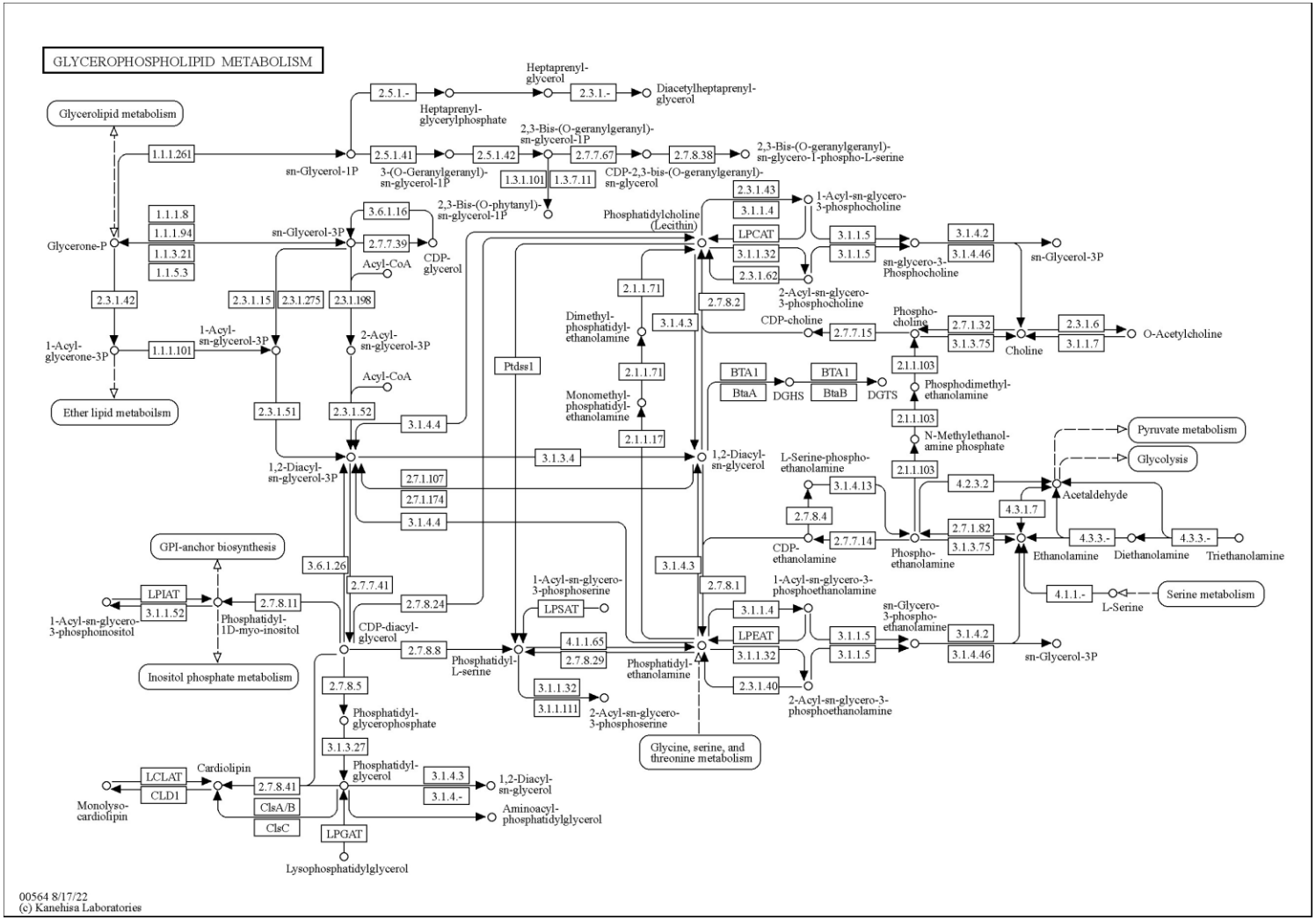
Relative Quantification

We determine the **fold change** between two or more groups by the ration of the signal in one group divided by the signal in the control group

It assumes that the response is always linear

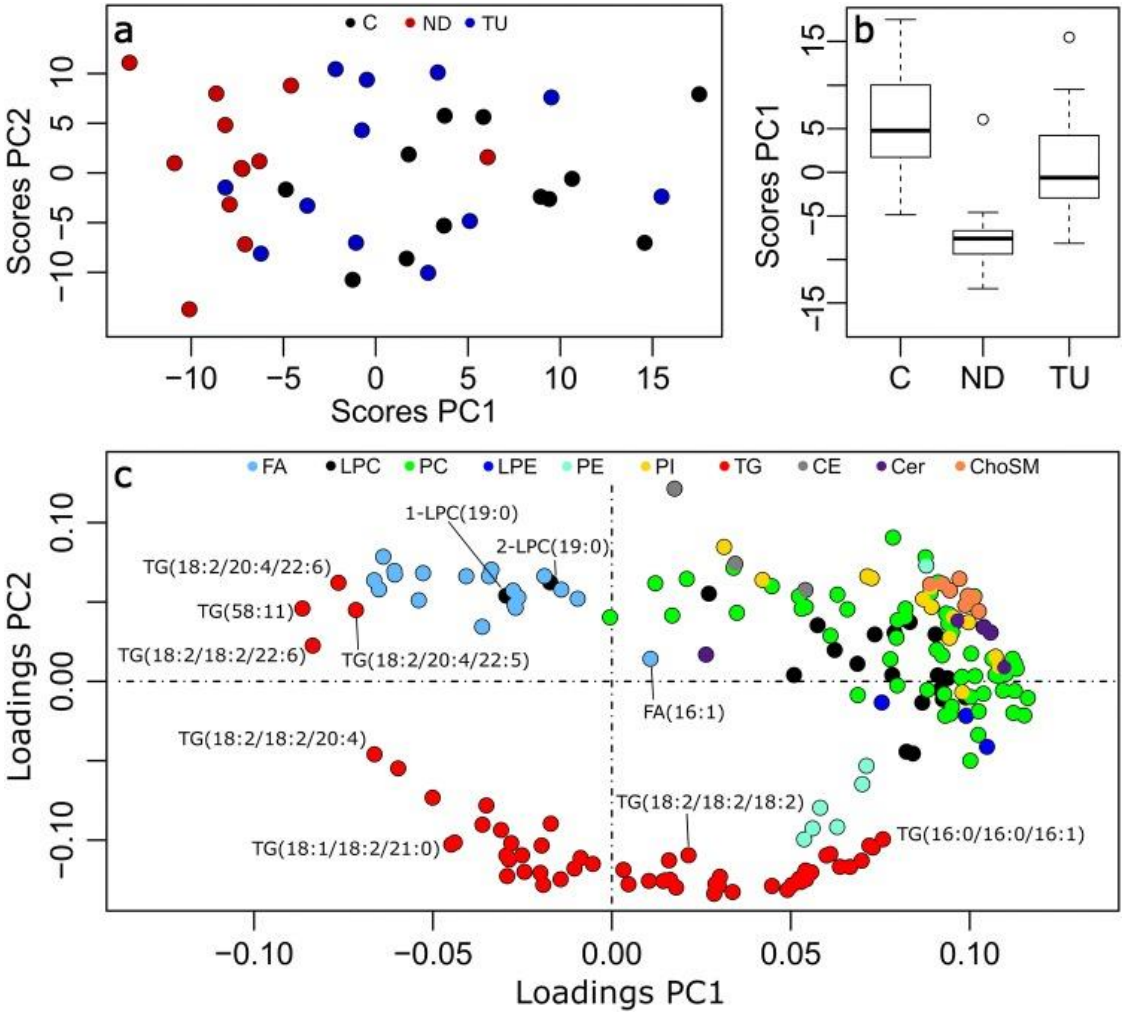
It is the most common quantification in lipidomics

Lipidome analysis



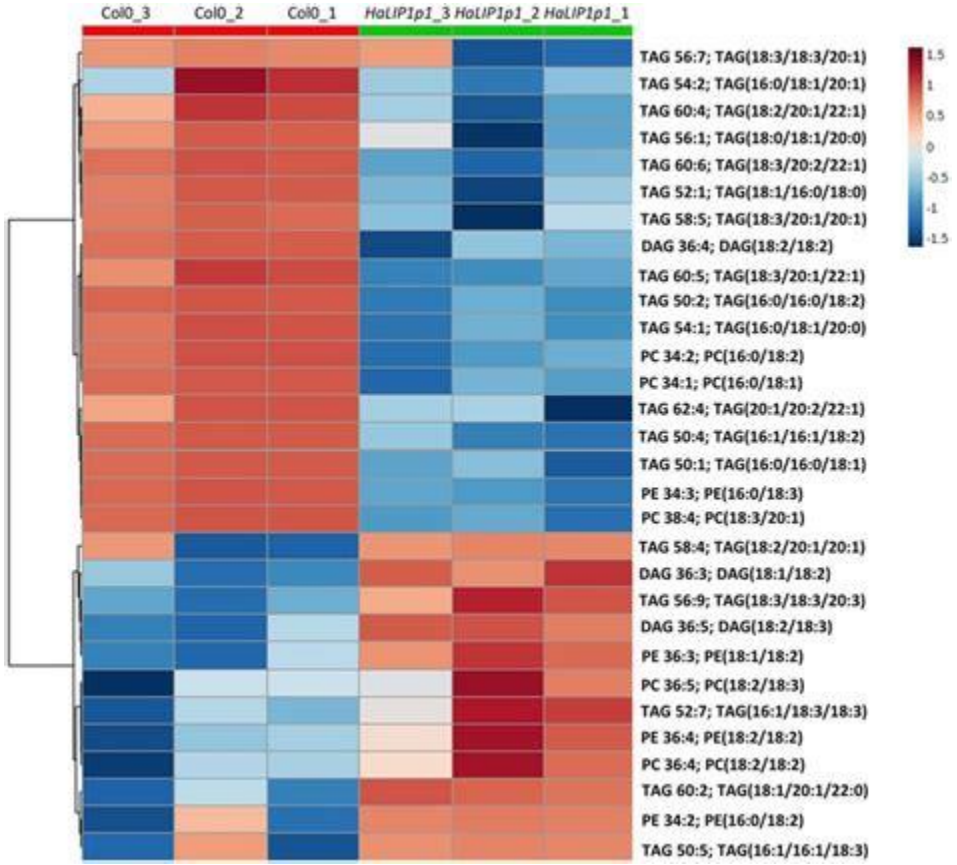
Unsupervised data analysis

Principal component analysis (PCA): scores and loadings



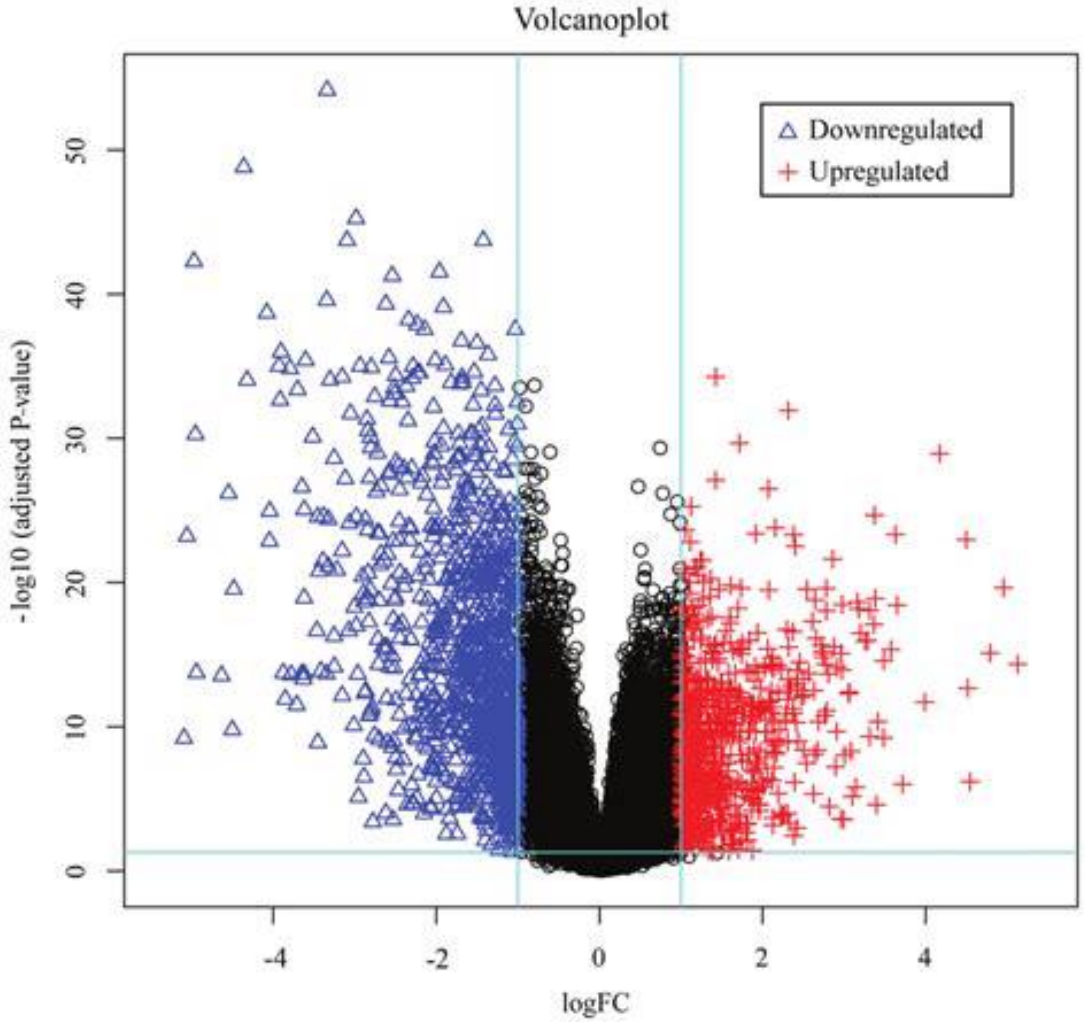
Unsupervised data analysis

Heatmaps



Unsupervised data analysis

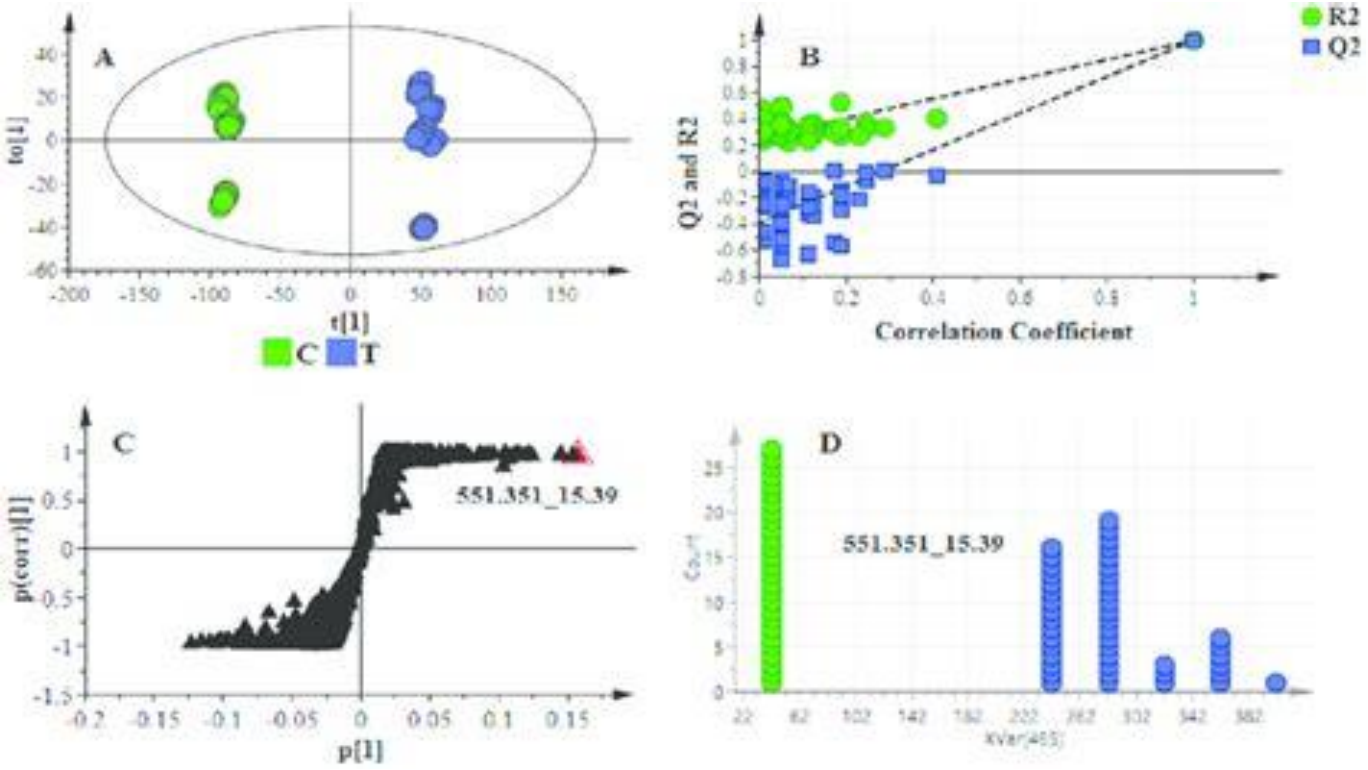
Volcano plots



Supervised data analysis

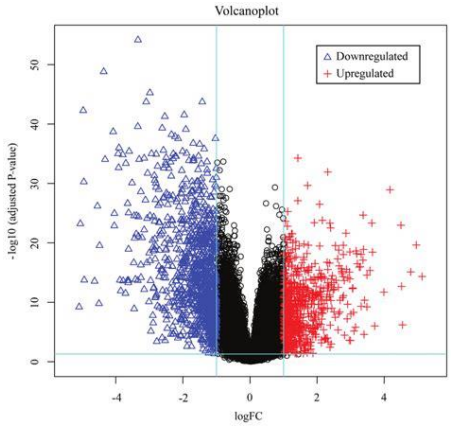
Orthogonal Partial Least Squares – Discriminant analysis (OPLS-DA)

- “Greedy” algorithm



Trends in data analysis

- **Avoiding p-values or other statistics as TRUE/FALSE decision boundaries (statistical significance)**



- **AI connecting data analysis and biological interpretation**

Seminar

- **Identification of lipids**