Liquid Chromatography Mass Spectrometry (LC-MS) based Metabolomics

Data and Preprocessing

Adrie D. Dane

Department of Epidemiology and Data Science (EDS)

Bioinformatics Laboratory

Amsterdam UMC University Medical Centers | EDS EPIDEMIOLOGY AND DATA SCIENCE | BIOINFORMATICS LABORATORY

www.amsterdamumc.nl | bioinformaticslaboratory.eu | cfmetabolomics.nl

About / Acknowledgements

Prof. dr. Antoine van Kampen

Professor

a.h.vankampen@amsterdamumc.nl

++31 (0)20 566 7096

++31 (0)6-41768067

AMC | Room J1B-208

https://www.skippyweb.eu

Bioinformatics Laboratory

bioinformaticslaboratory.eu

Core Facility Metabolomics (CFM)

cfmetabolomics.nl

F.M. Vaz, PhD

Head CFM

M. van Weeghel, PhD

Manager CFM

G.S. Salomons, PhD

Head Laboratory GMZ

A.D. Dane, PhD

Bioinformatician

J.B. van Klinken, PhD

Bioinformatician

M.L. Pras-Raves, PhD

Bioinformatician

E.J.M. Wever, MSc

Bioinformatician

J. Hermans, BSc

Technician

H. van Lenthe, BSc

Technician

B.V. Schomakers, MSc

Technician

M.A.T. Vervaart, BSc

Technician

Contents

- Introduction to metabolomics
  - Challenges in measuring metabolomics
- Introduction to Liquid Chromatography Mass Spectrometry (LC-MS)
- LC-MS Based metabolomics
  - Data
  - Preprocessing

RAW LC-MS Data

Metabolite ID	Metabolite	Sample 1	Sample 2	Sample 3
1	glucose	5000	4500	1000
2	unknown	8000	10000	8000
3	unknown	1205	3000	8000

Metabolomics

Definitions

Metabolism*

- All biochemical processes that take place in the cells of organisms.

Metabolite

- Any organic molecule in the body/organism with a molar mass < 1500 Da

Metabolome

- The complete collection of all metabolites in a given cell, organ, tissue or organism

Metabolomics

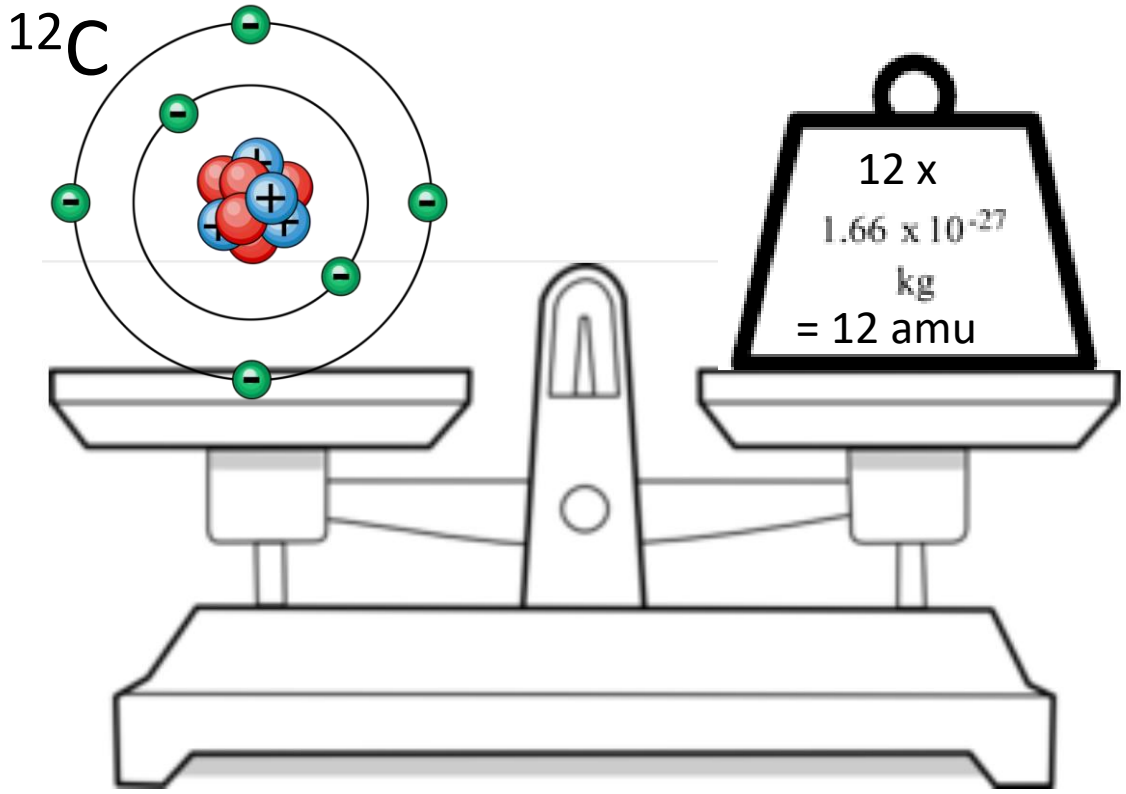
- A field of life science research that uses high throughput technologies to identify and characterize ***all*** metabolites in a given cell, organ, tissue or organism

*https://nl.wikipedia.org/wiki/Stofwisseling

Dalton (*Da*)

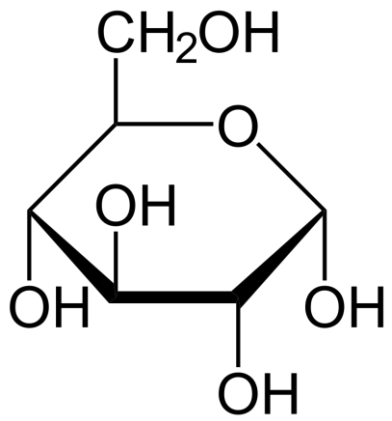
Unified atomic mass unit (*amu, u*)

- One twelfth of the mass of an unbound neutral atom of carbon-12 in its nuclear and electronic ground state
- Approximately the mass of a nucleon (proton or neutron)



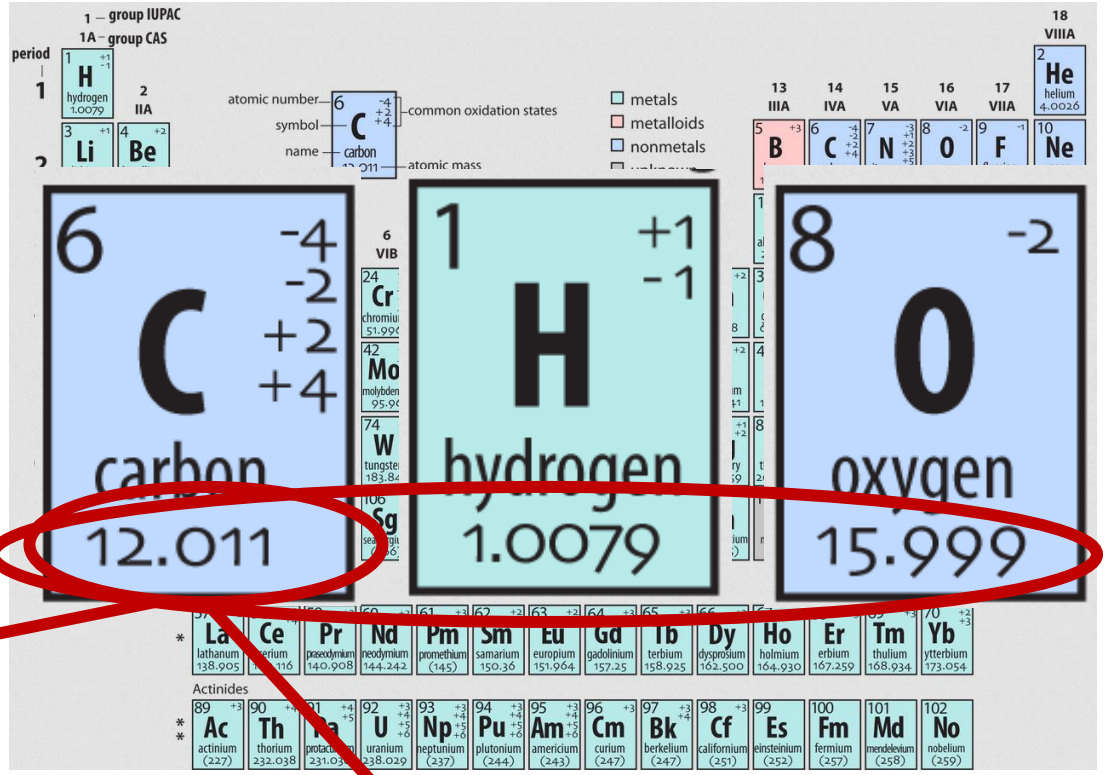
Molar Mass

Glucose



$C_6H_{12}O_6$

6x**12** + 12x**1** + 6x**16** = 180 Da



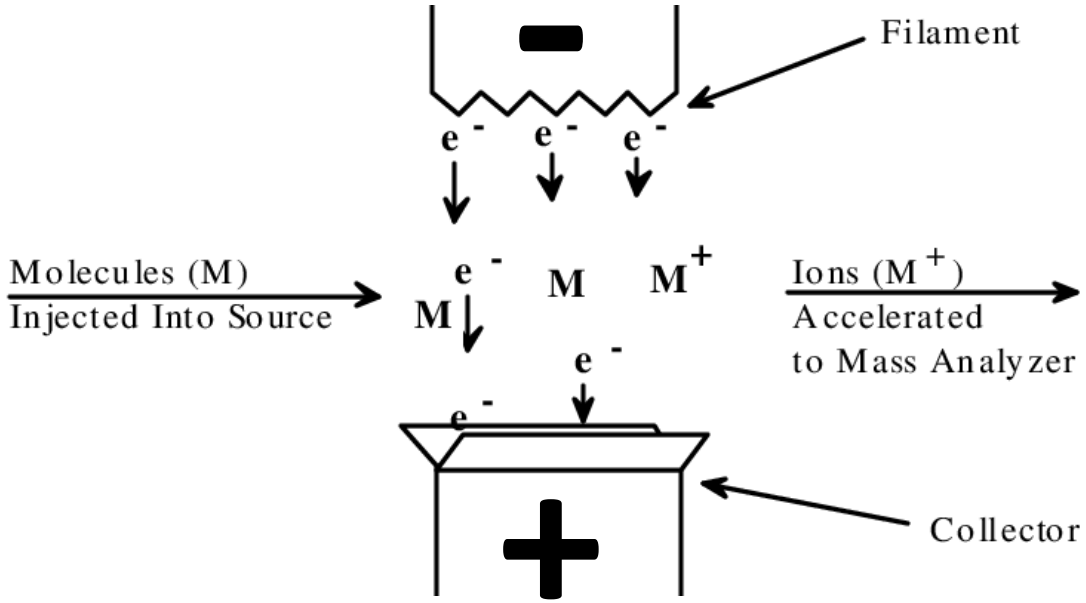
Molar Mass 180,15588 g/mol

Isotopes:

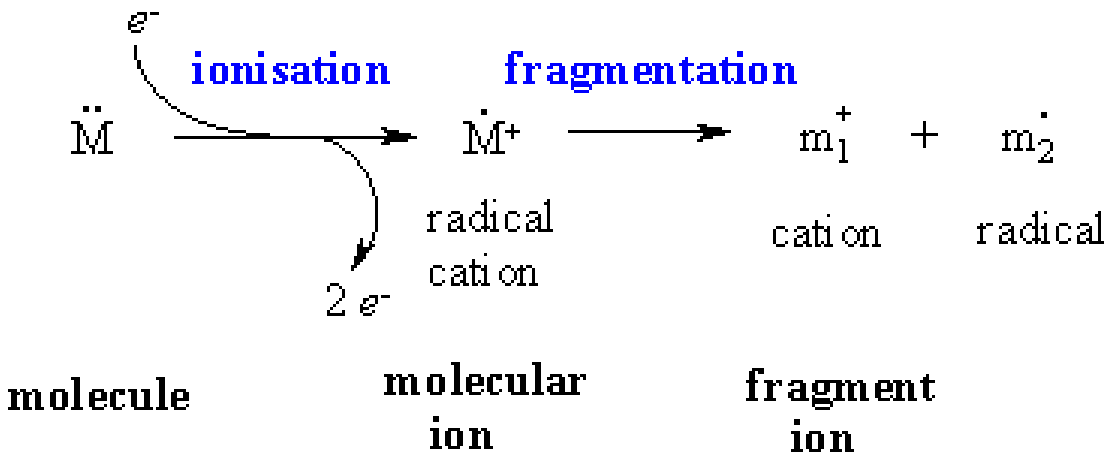
^{12}C	12Da	98.93%
^{13}C	13Da	1.1%

Mass Spectrometry

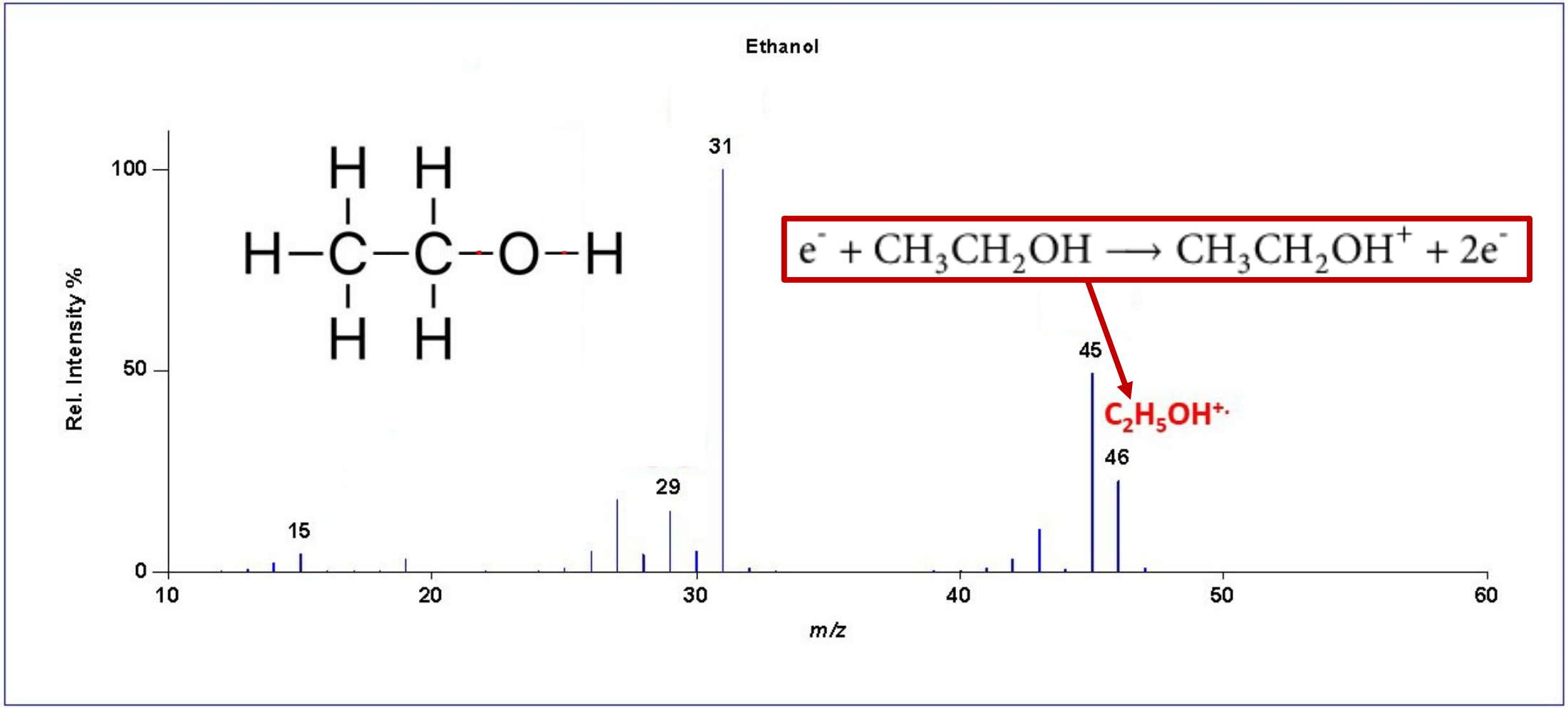
- Molar masses can be determined by mass spectrometry
- In order to “weigh” molecules they must be in gas phase and charged



Electron ionization (EI)

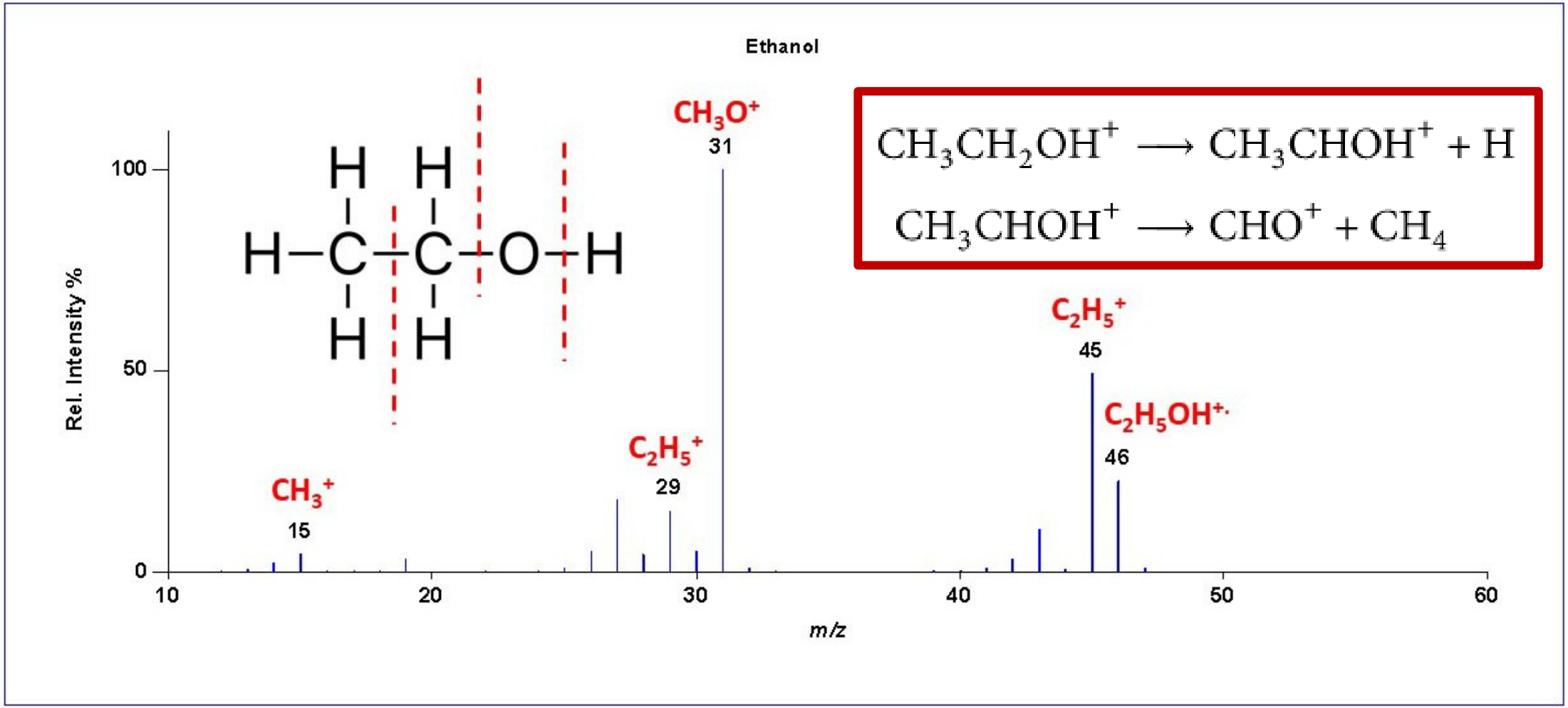


Mass Spectrum of Ethanol



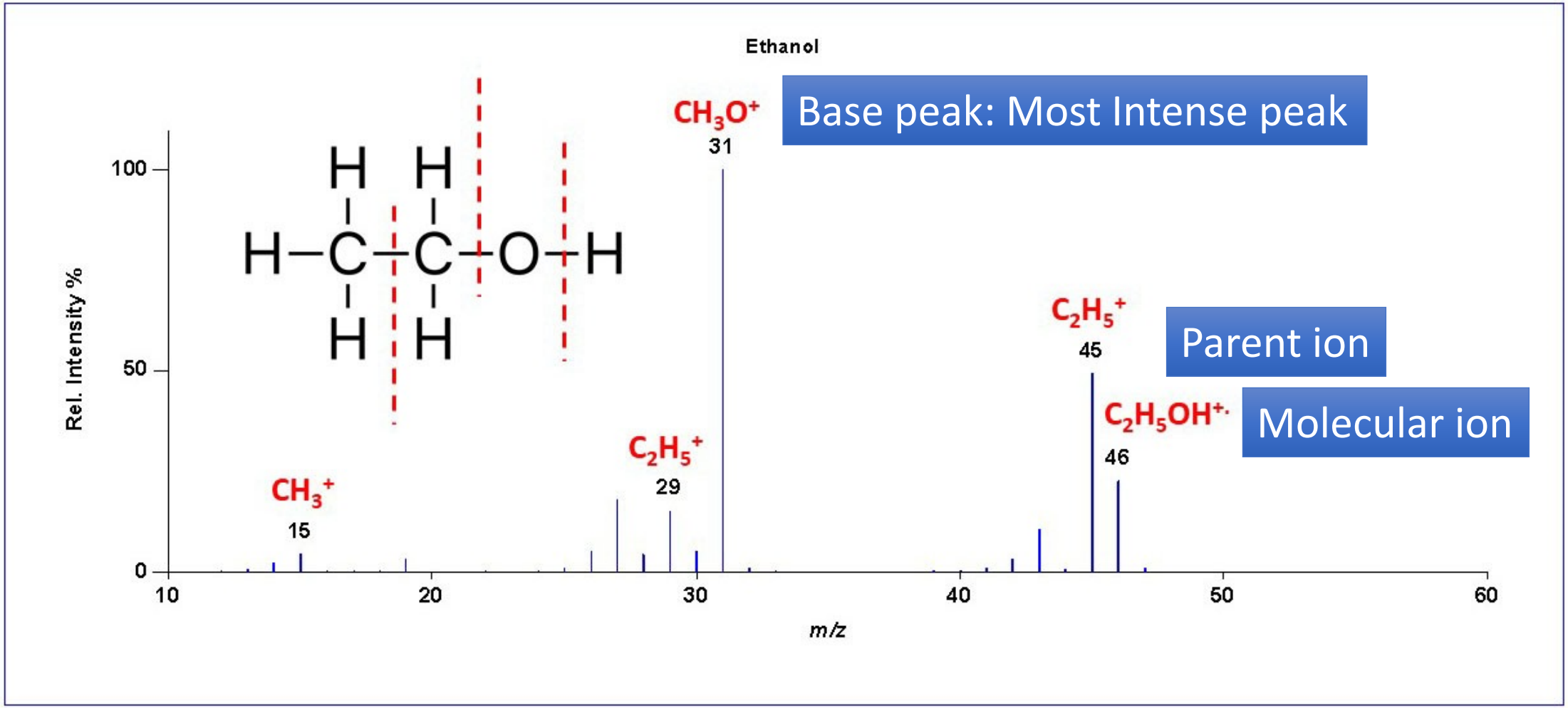
https://byo.com/article/ferment-aroma-loss/

Mass Spectrum of Ethanol



https://byo.com/article/ferment-aroma-loss/

Mass Spectrum of Ethanol

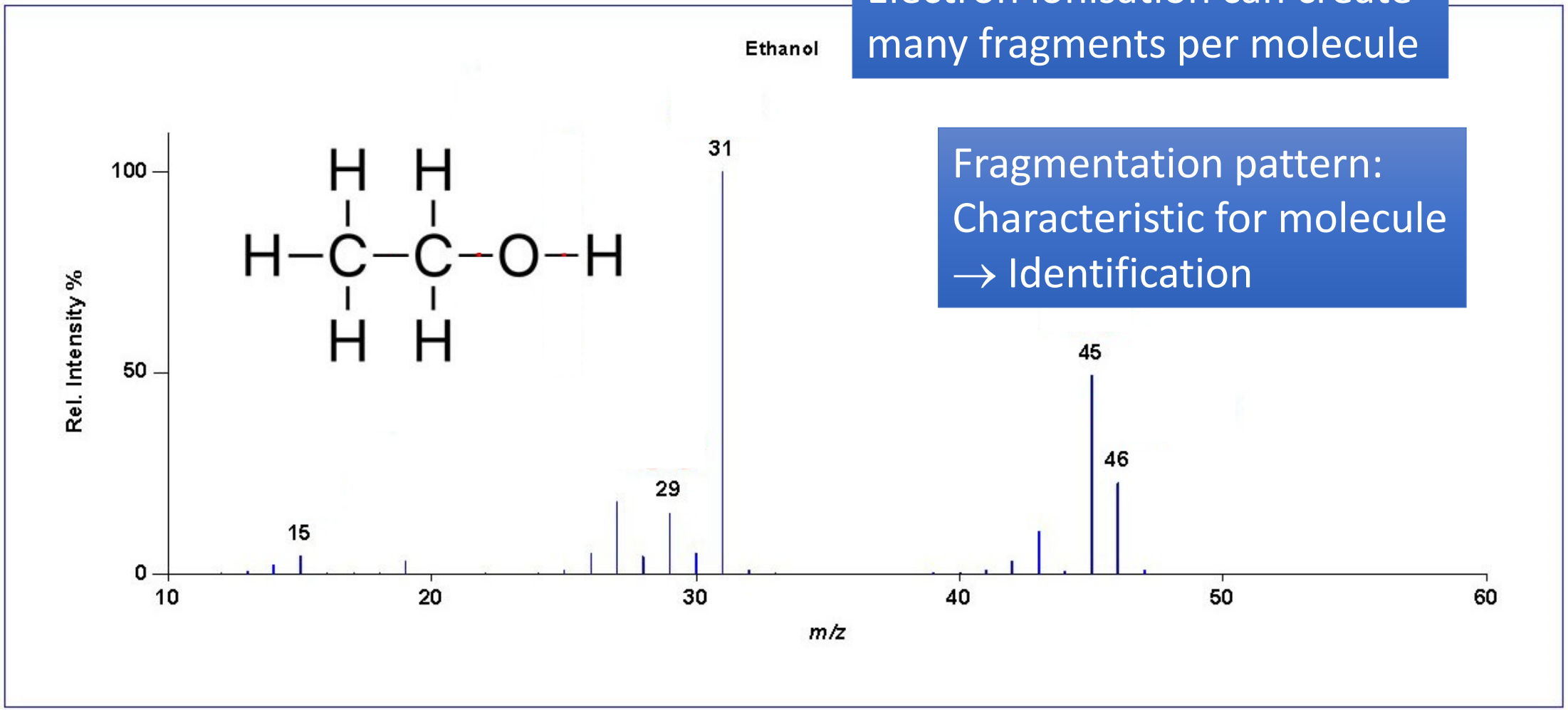


https://byo.com/article/ferment-aroma-loss/

Mass Spectrum of Ethanol

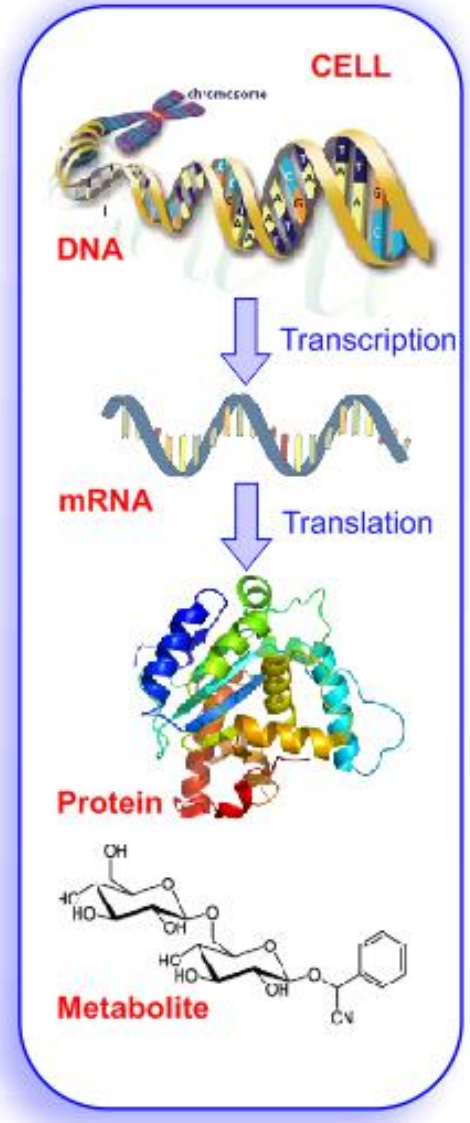
Hard ionisation:
Electron ionisation can create many fragments per molecule

Fragmentation pattern:
Characteristic for molecule
→ Identification



https://byo.com/article/ferment-aroma-loss/

Molar Masses and Omics



- Genomics
- Transcriptonics
- Proteomics
- Metabolomics

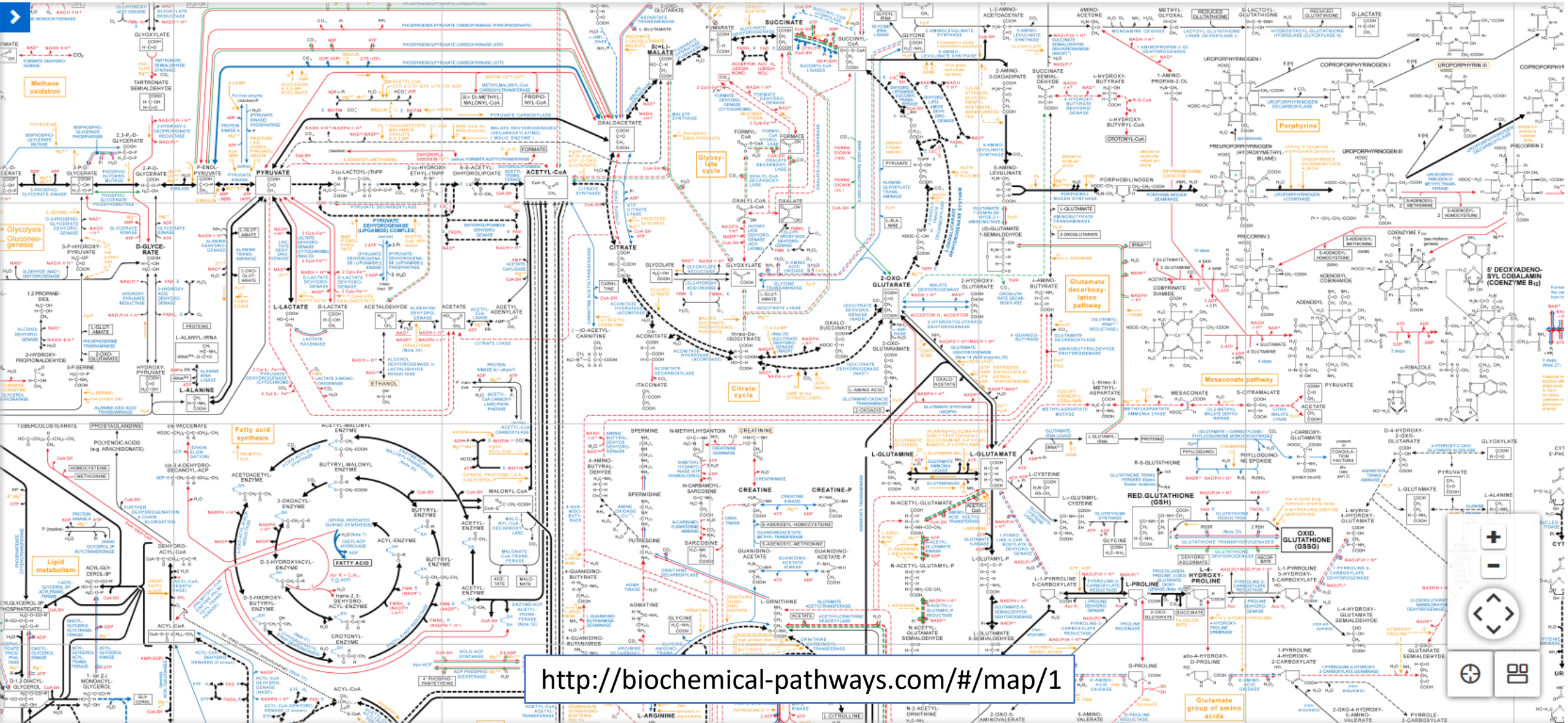
Bacteria DNA:	~10^7-10^8 Da	Chromosomes
Human DNA:	~10^{10}-10^{12} Da	
Human Genome*:	2.2×10^{12} Da = 3.652×10^{-15} kg	
transfer RNA*:	2.5×10^4 Da	RNA
28S rRNA*:	1.6×10^6 Da	
BPTI:	6512 Da	Proteins
IgG:	150,000 Da	
IgM:	900,000 Da	
	1500 Da	
Glucose:	180 Da	Metabolites
Cholesterol:	380 Da	
Cardiolipin:	1466 Da	

*https://sfvideo.blob.core.windows.net/sitefinity/docs/default-source/biotech-basics/molecular-facts-and-figures.pdf

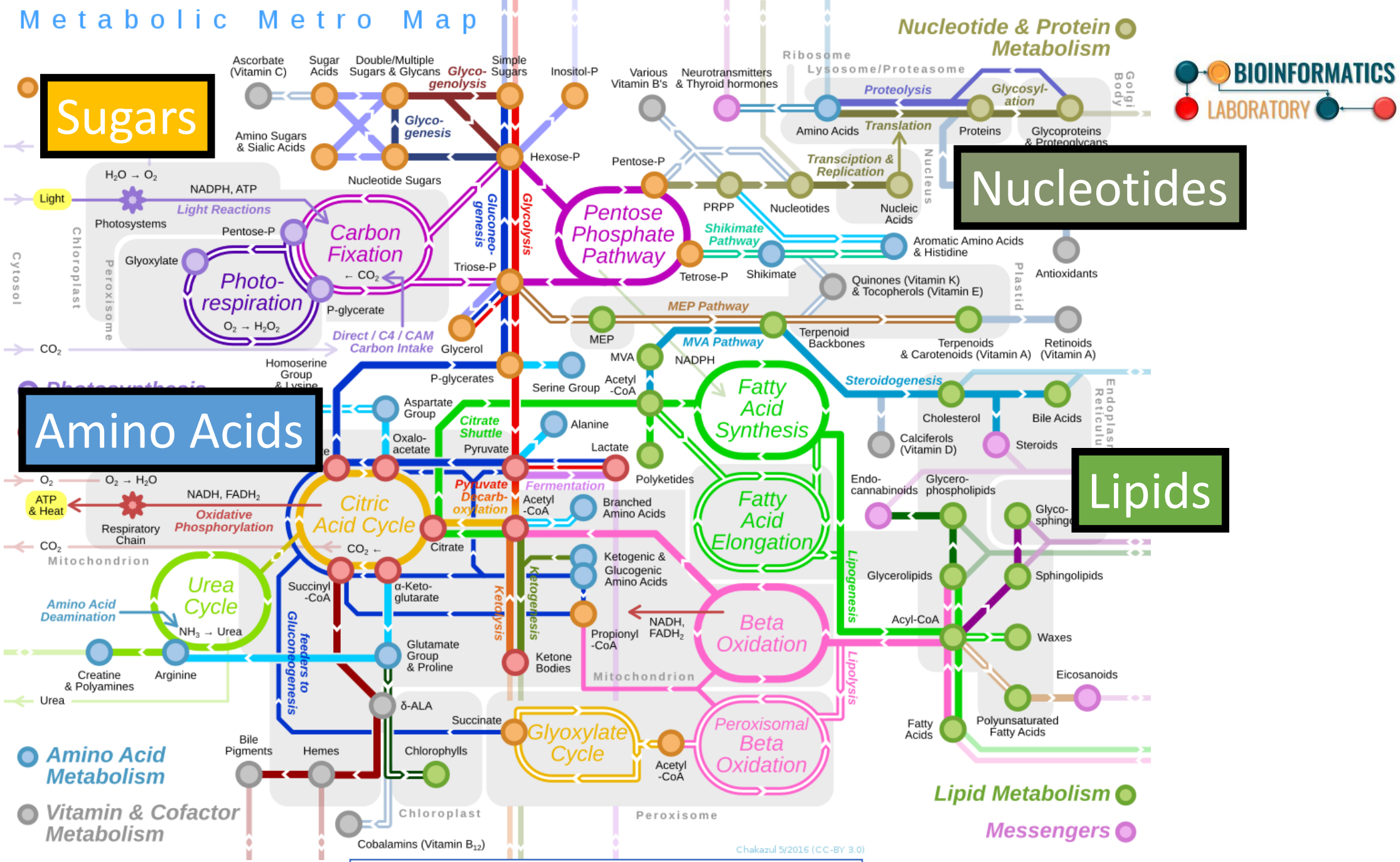
Metabolism

BIOINFORMATICS LABORATORY

Part 1: Metabolic Pathways Part 2: Cellular and Molecular Processes



http://biochemical-pathways.com/#/map/1



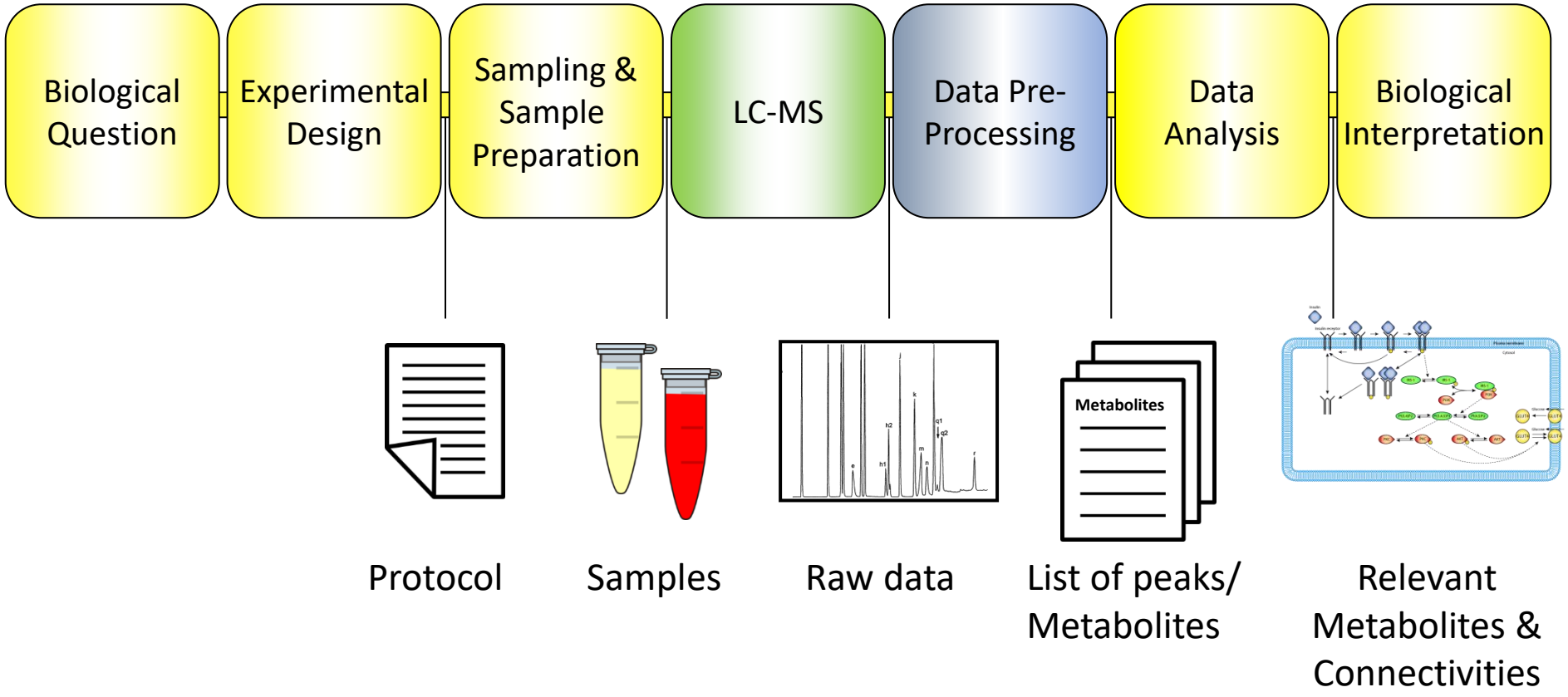
https://en.wikipedia.org/wiki/Metabolic_pathway

Applications of metabolomics

- Clinical
  - Human disease: cancer, diabetes, metabolic disorders etc.
  - Effects of pharmaceuticals or nutrition
  - Biomarker discovery

- Research and development
  - Investigate phenotypic effects of genetic manipulation or genetic differences (bacteria, yeast, plants etc.)
  - Systems biology

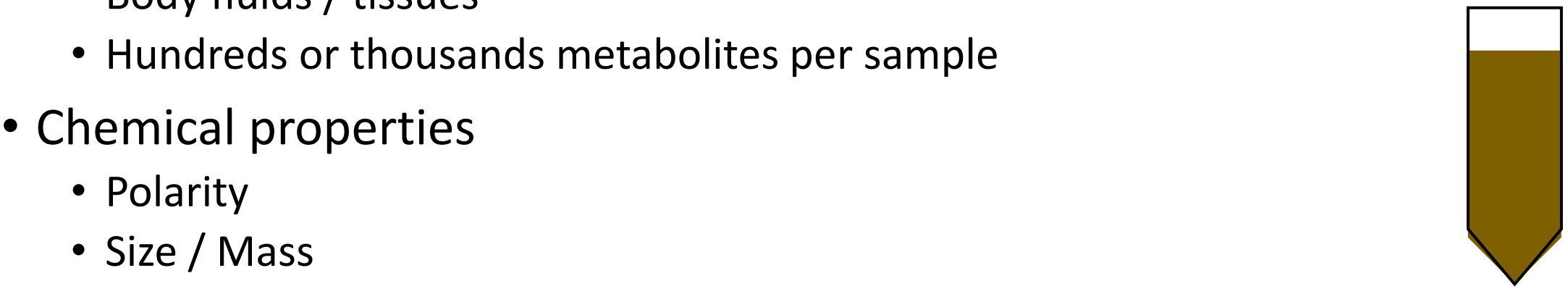
LC-MS based metabolomics workflow

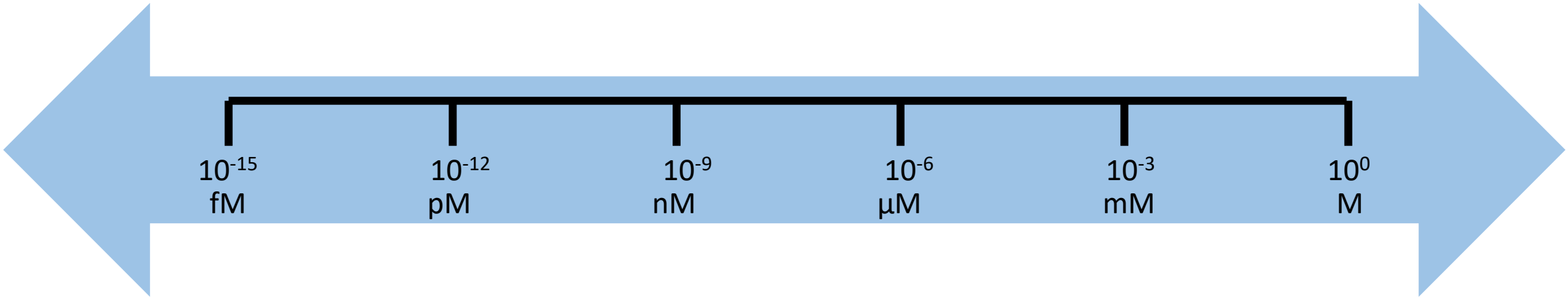


Metabolomics Challenges

- Sample Complexity
  - Body fluids / tissues
  - Hundreds or thousands metabolites per sample
- Chemical properties
  - Polarity
  - Size / Mass
- Concentrations

Biological sample





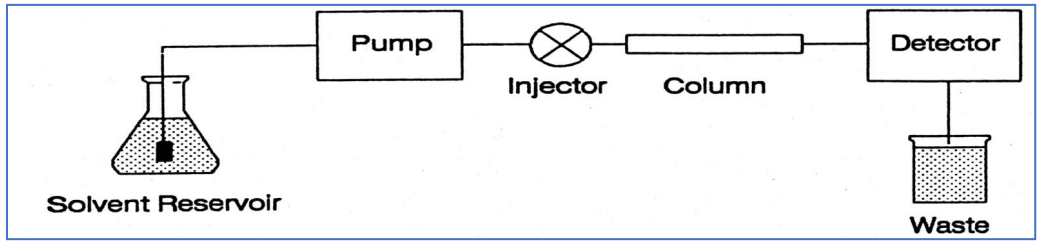
Liquid Chromatography Mass Spectrometry (LC-MS)

Introduction

Metabolomics: LC - MS

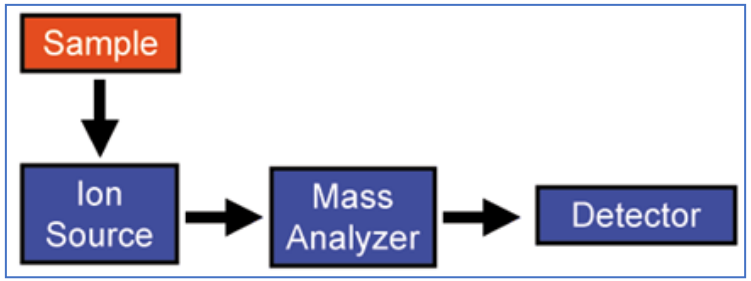
BIOINFORMATICS LABORATORY

Liquid chromatograph



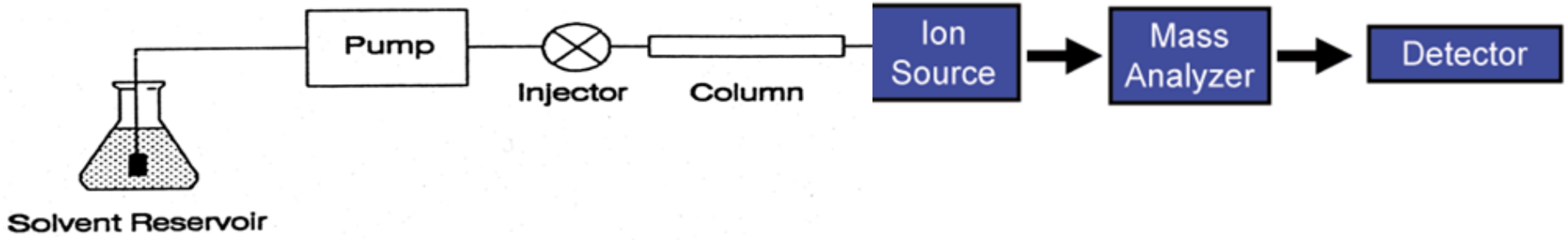
+

Mass Spectrometry

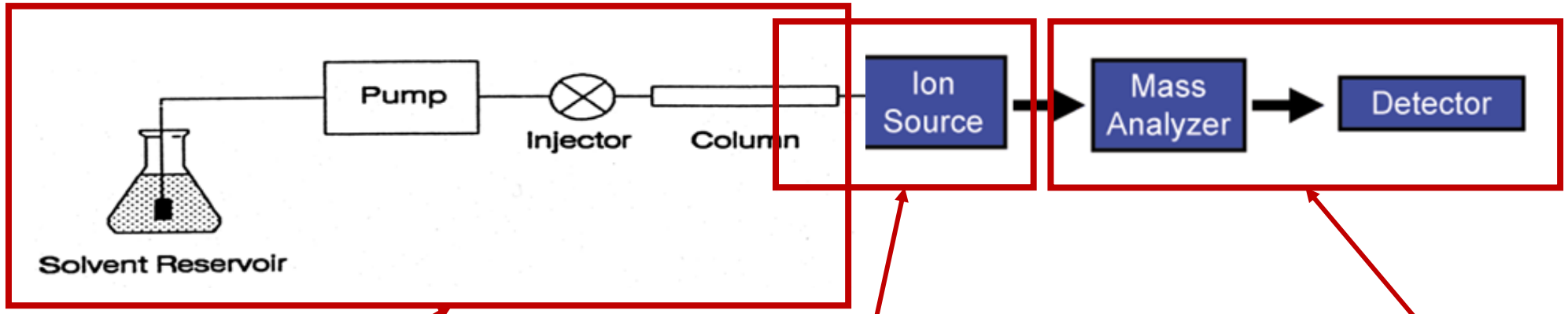


Separate analytes based on physical and/or chemical properties

Identify and characterize analytes based on their mass



Metabolomics: LC - MS



(Ultra) High Performance Liquid Chromatography (U)HPLC

Separation based on polarity

Electrospray Ionization

Generate ions in gas phase

Atmospheric Pressure

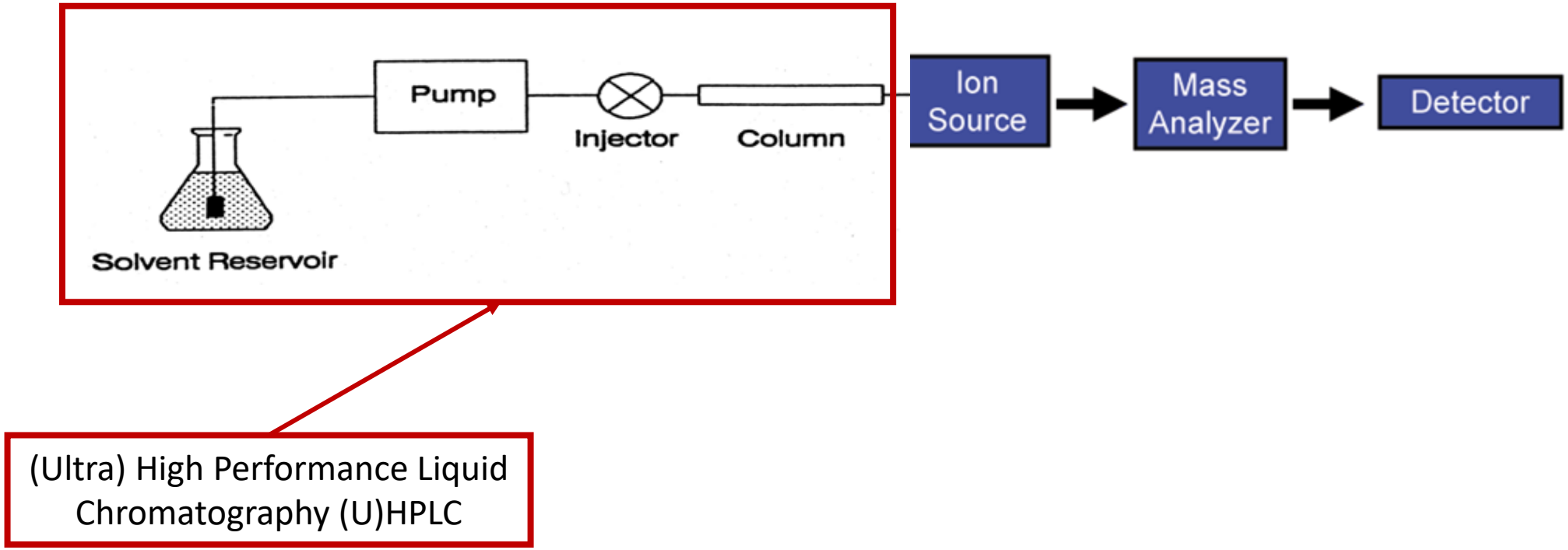
Mass Analyzers in LC-MS

- High resolution Data
- Fast data acquisition rates

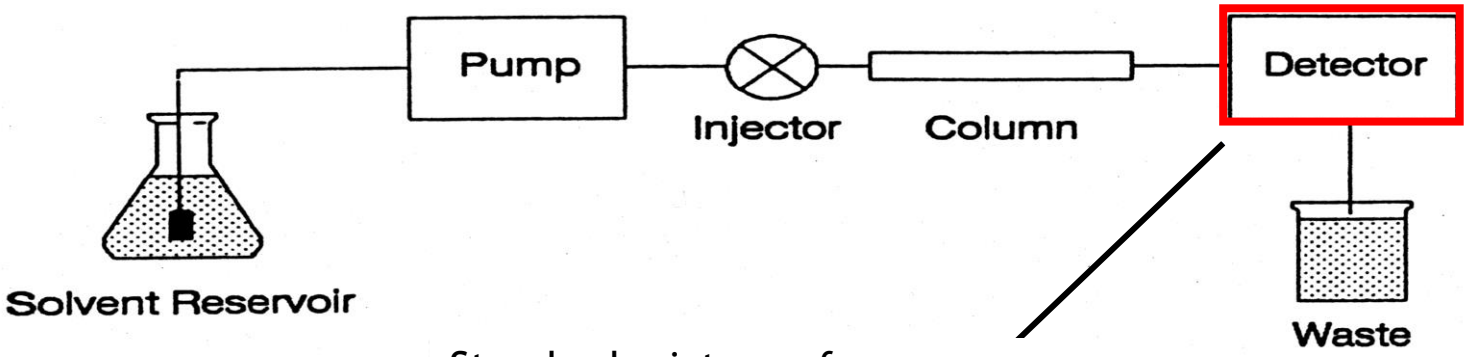
Identify and characterize metabolites based on their mass

High vacuum

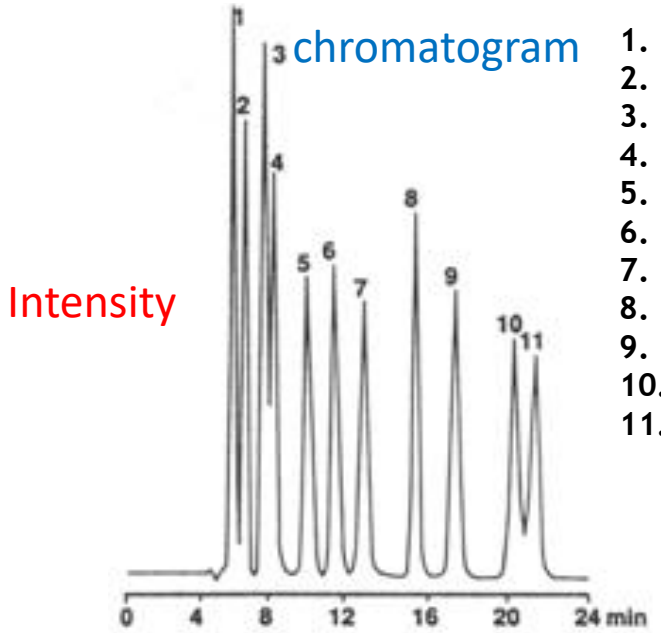
Metabolomics: LC - MS



Liquid Chromatograph: summary



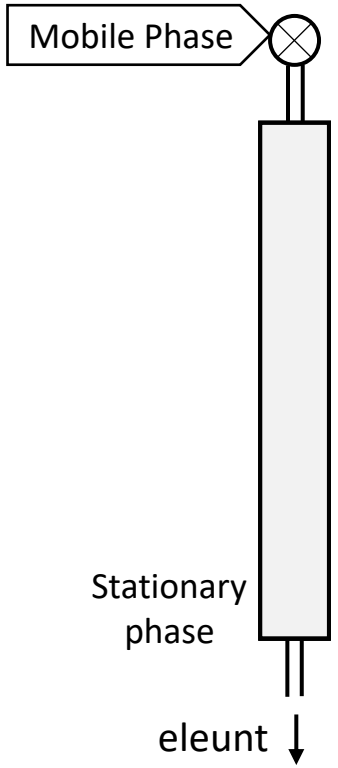
Standard mixture of sugars and alcohols



1. Stachyose
2. Raffinose
3. Sucrose
4. Lactose
5. Glucose
6. Galactose
7. Fructose
8. Erythritol
9. Mannitol
10. Galactitol
11. Sorbitol

Chromatography: separation of mixtures based on different affinities of substances for a stationary phase and a moving phase

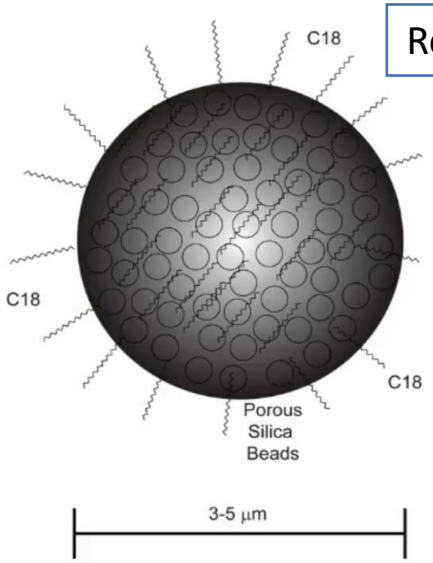
Liquid Chromatography

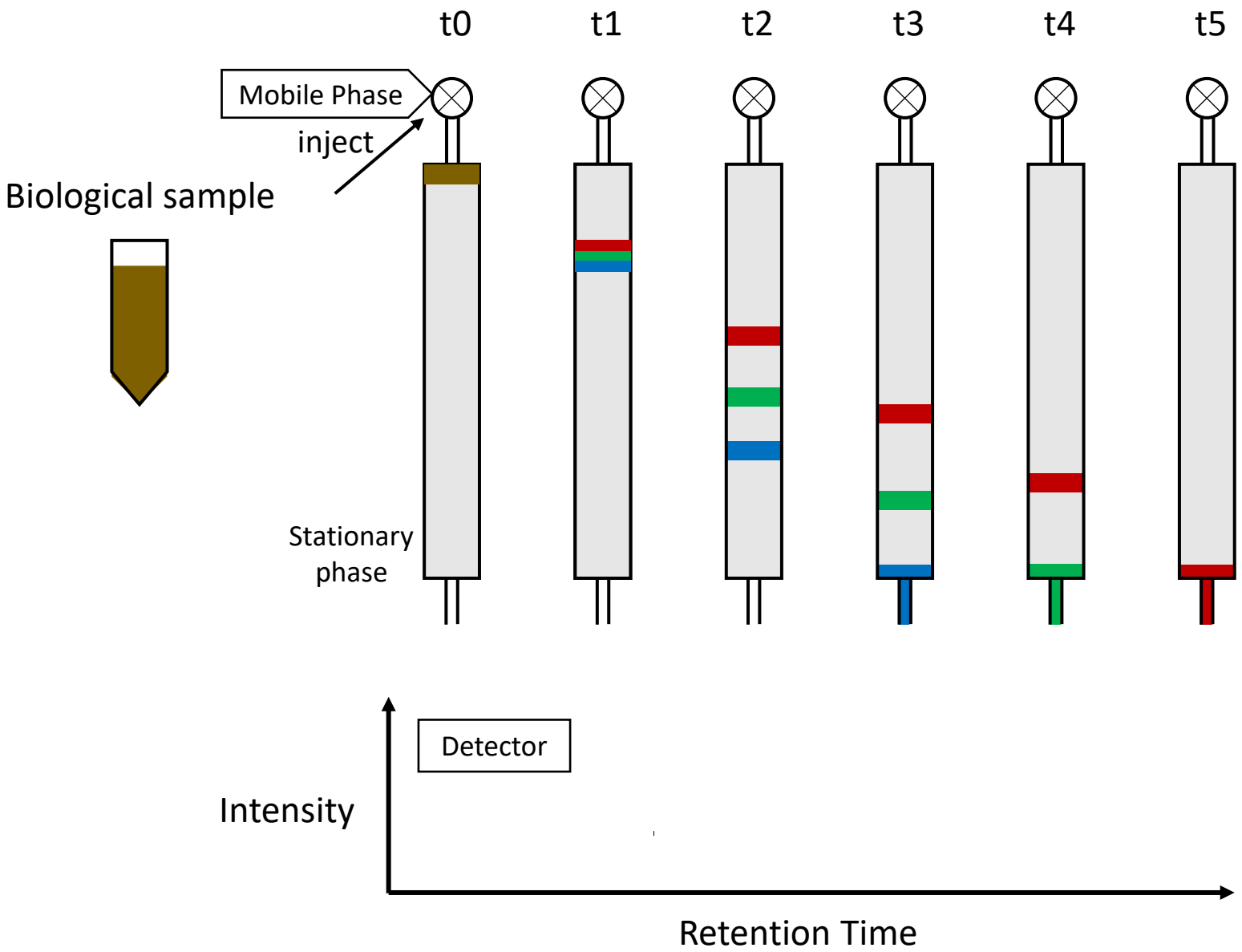


Stationary Phase: a solid material packed into a column

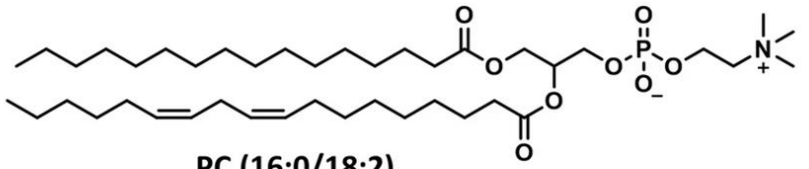
Mobile Phase: a liquid solvent

Reversed Phase



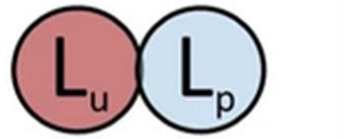

t0
t1
t2
t3
t4
t5
Mobile Phase
inject
Biological sample
Stationary phase
Detector
Intensity
Retention Time

HPLC



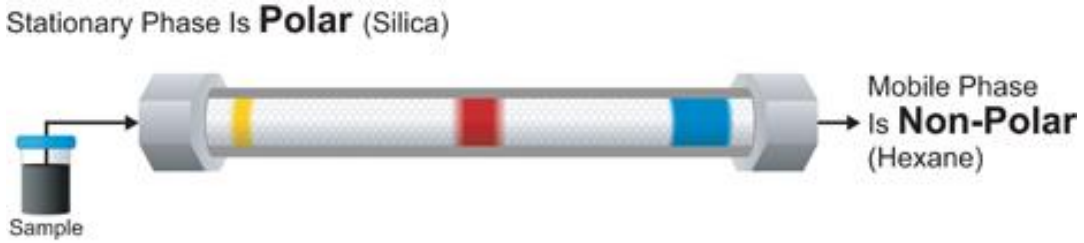
PC (16:0/18:2)

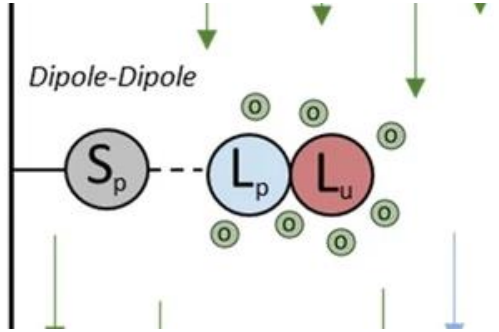
a phosphatidylcholine.



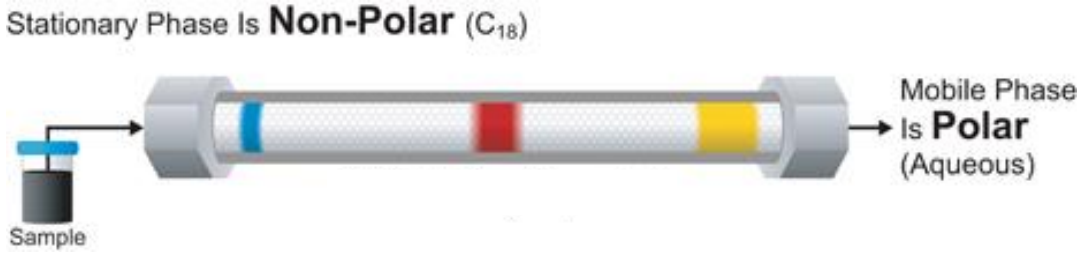
A polar tails Polar head

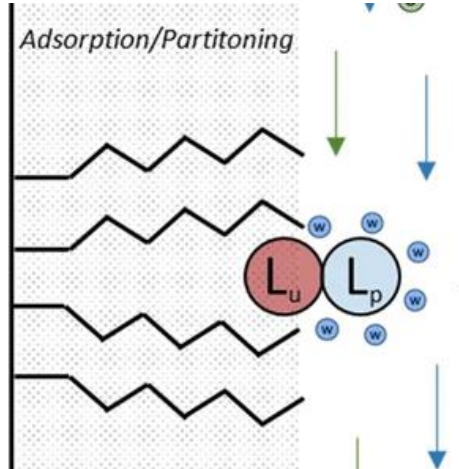
- Normal phase





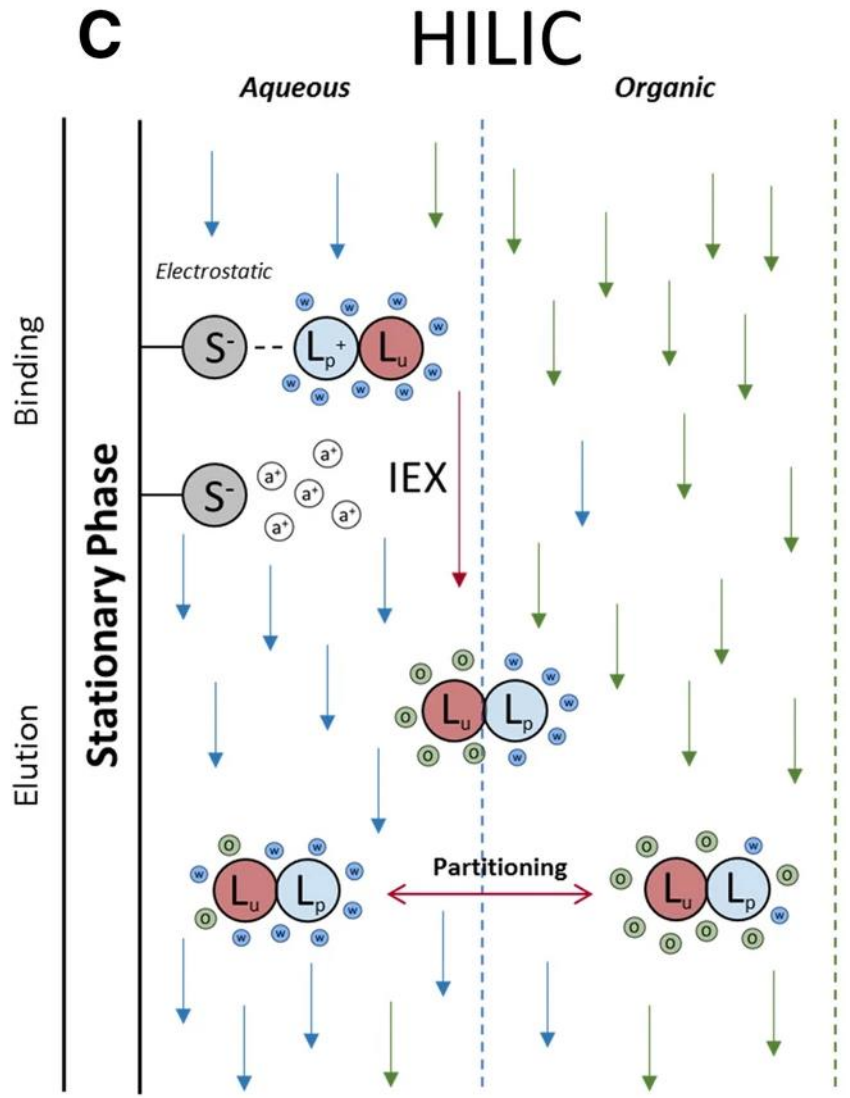
- Reversed phase





HPLC

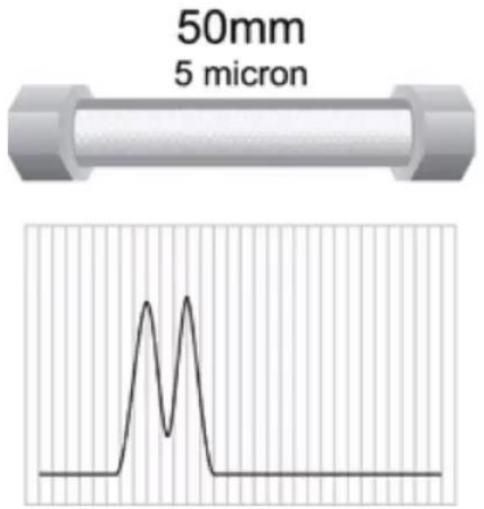
- HILIC
  - Mixed phase separates polar hydrophilic metabolites

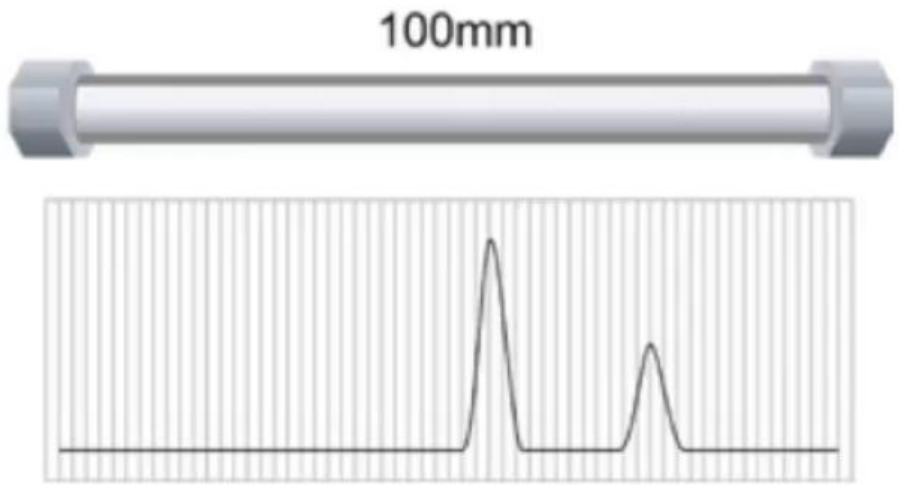


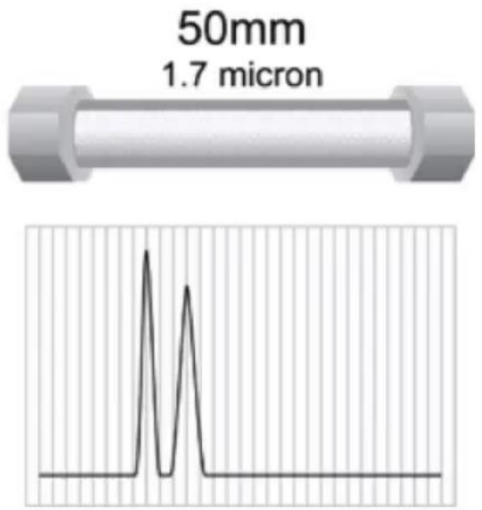
HPLC Separation

Better separation when total active surface area increases

- Longer column
  - Increased analysis time
- Smaller particles
  - More resistance: requires Higher pressure



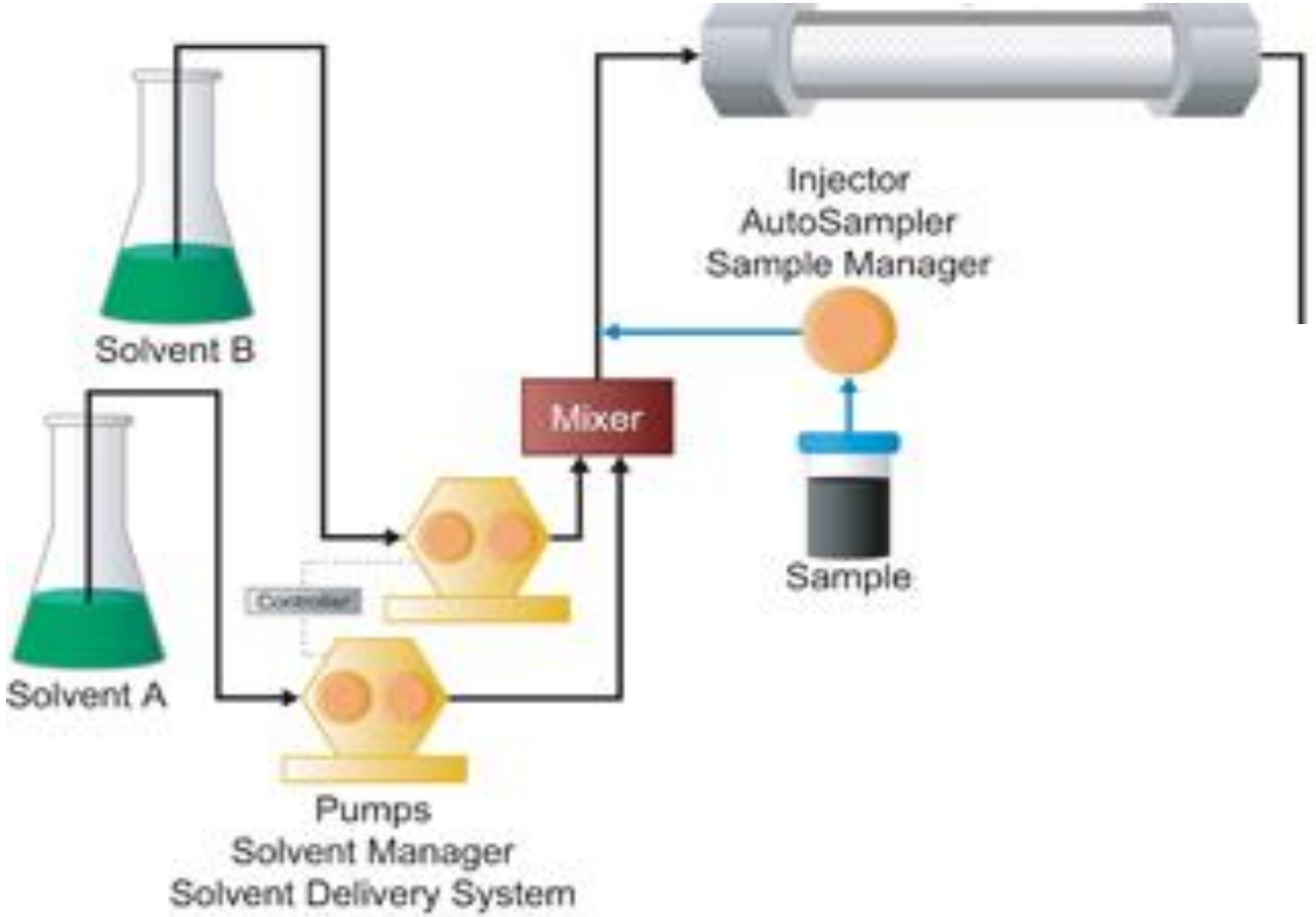




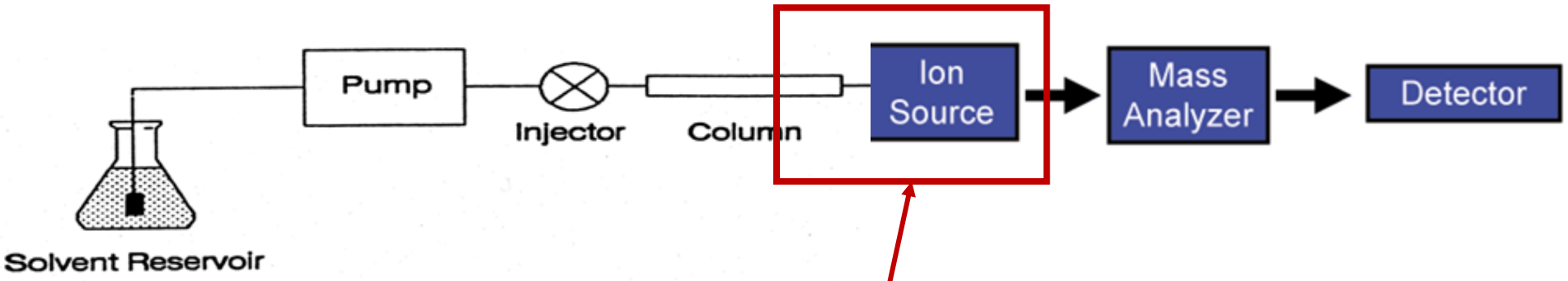
Ultra high performance liquid chromatography (UHPLC)

- Better separation in the same analysis time

Improved separation by using a gradient



Metabolomics: LC - MS



Electrospray Ionization

Generate ions in gas phase

Atmospheric Pressure

Electrospray ionization (ESI)

Ions are created in the solvent (mobile phase)

- Positive ion mode: protonate bases and create $[M+H]^+$

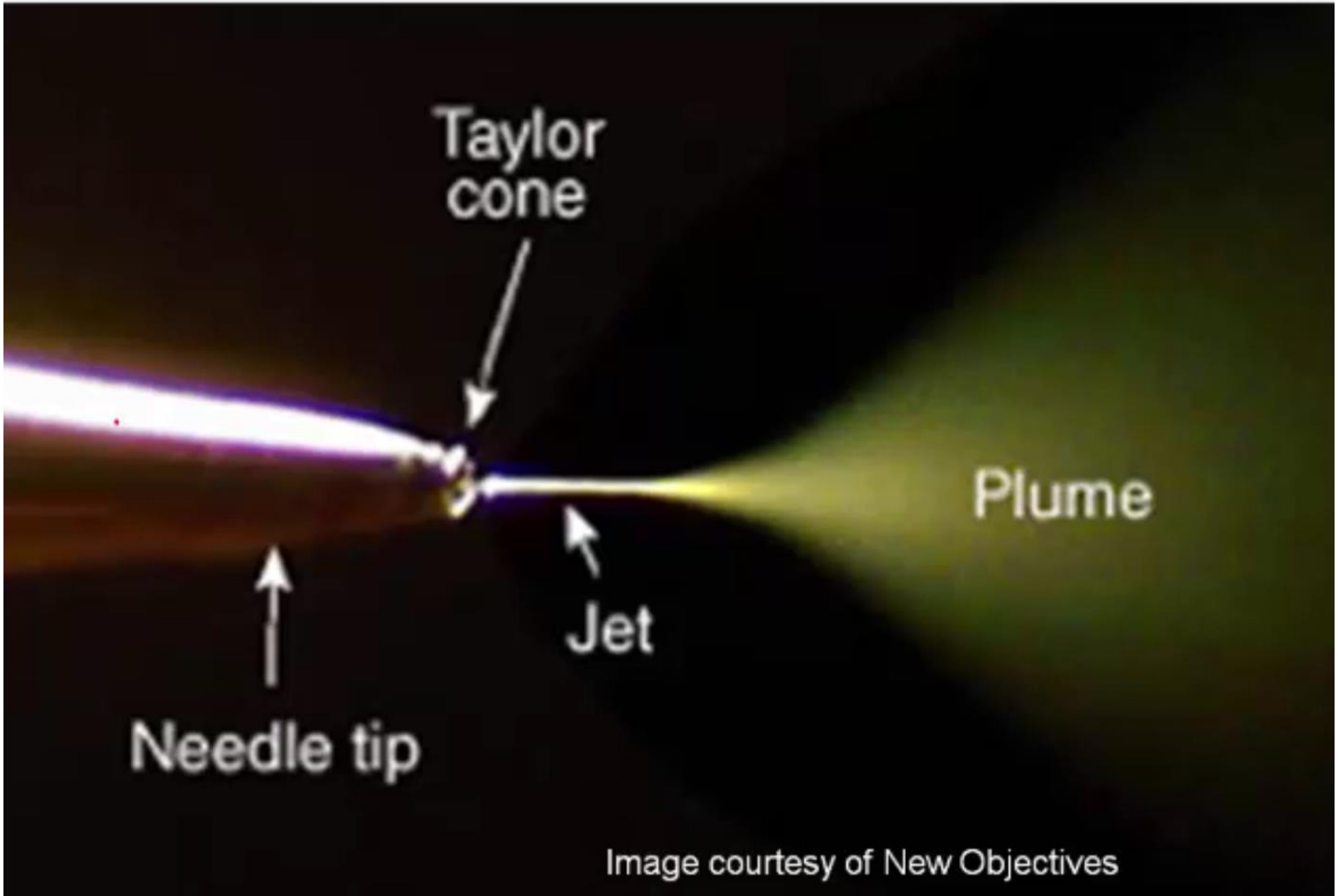
$$R\text{-}NH_2 + H_3O^+ \leftrightarrow R\text{-}NH_3^+ + H_2O$$

- Negative ion mode: deprotonate acids and create $[M-H]^-$

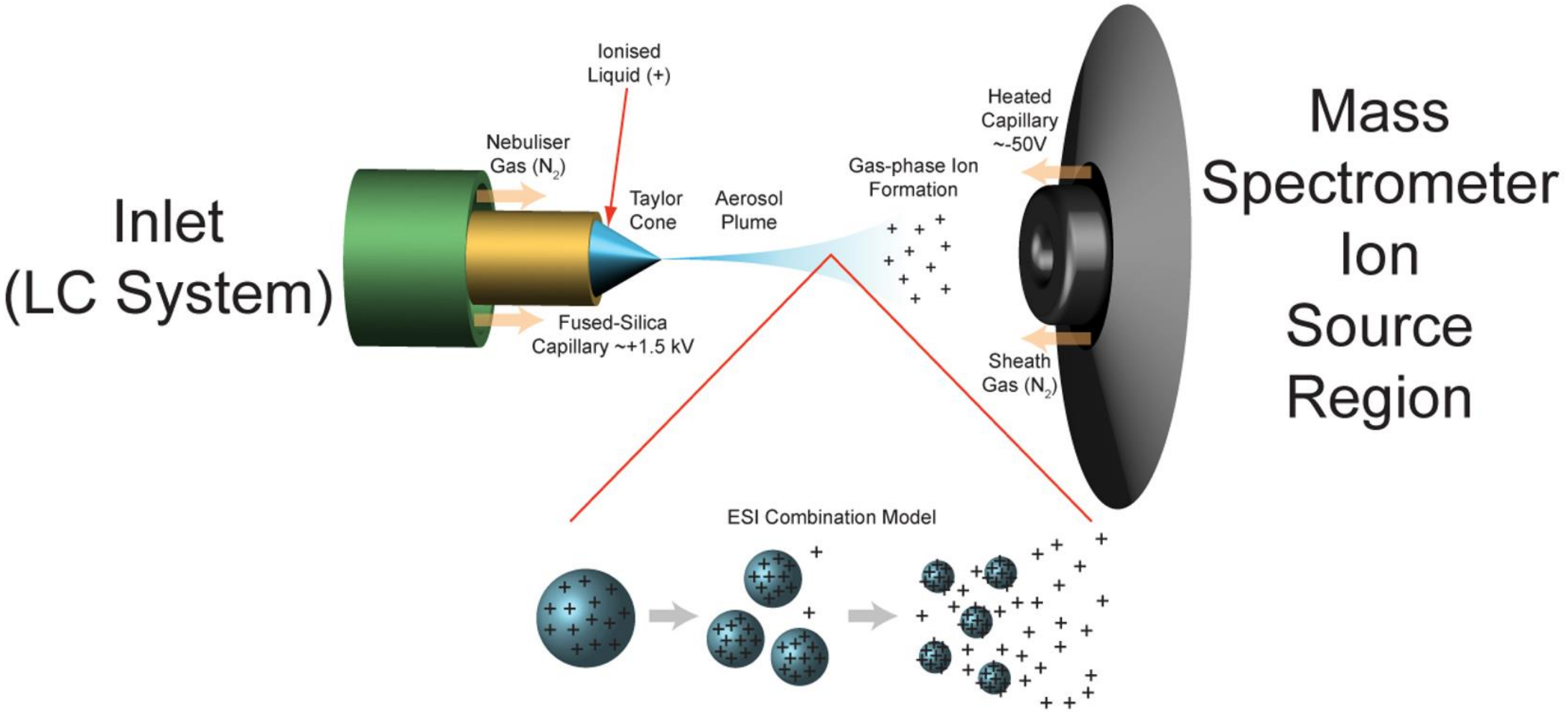
$$R\text{-}COOH + OH^- \leftrightarrow R\text{-}COO^- + H_2O$$

- Other ions may form:
  - $[M+Na]^+$, $[M+NH_4]^+$, …
  - $[M+Cl]^-$, $[M+CHO_2]^-$, $[M-2H]^{2-}$, …

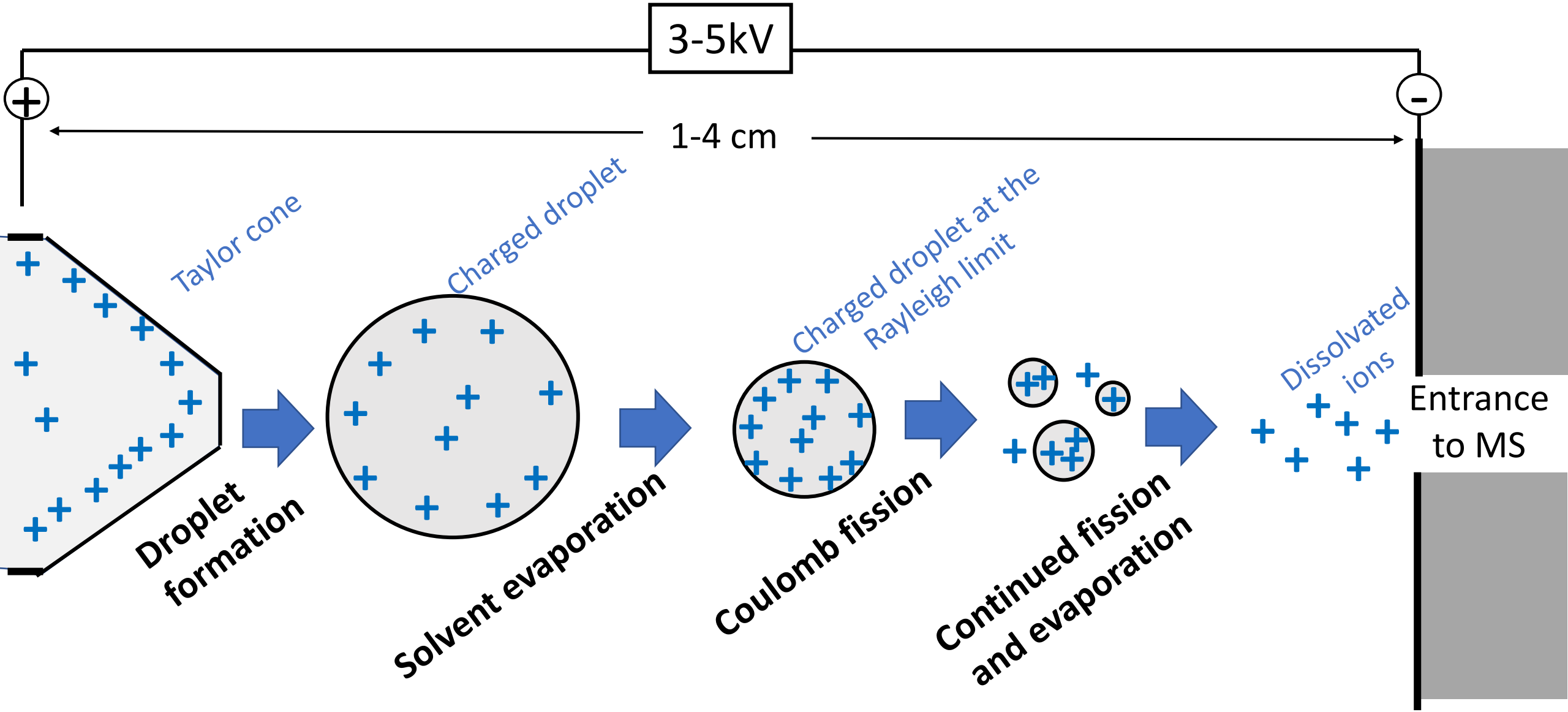
Photo of ESI Process



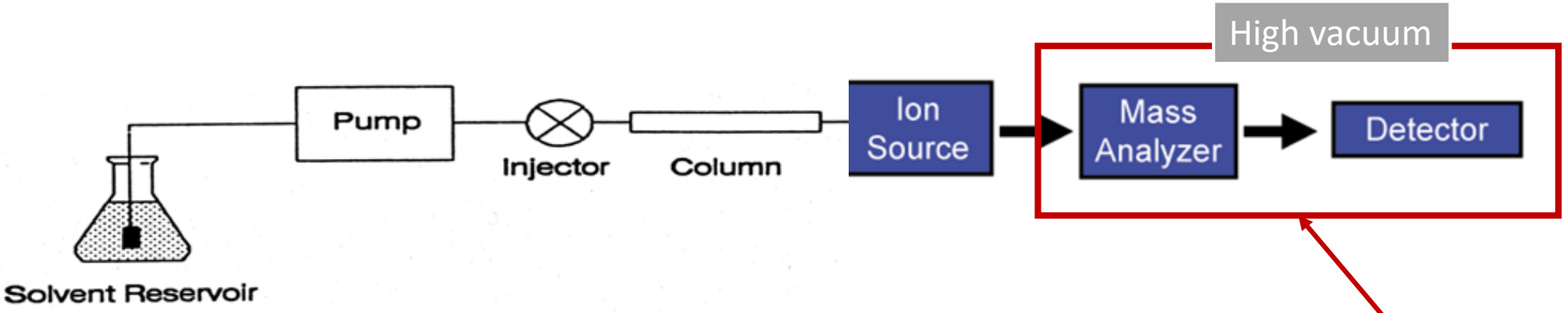
Electrospray ionization (ESI)



Electrospray ionization mechanism



Metabolomics: LC - MS



Mass Analyzers in LC-MS

- High resolution Data
- Fast data acquisition rates

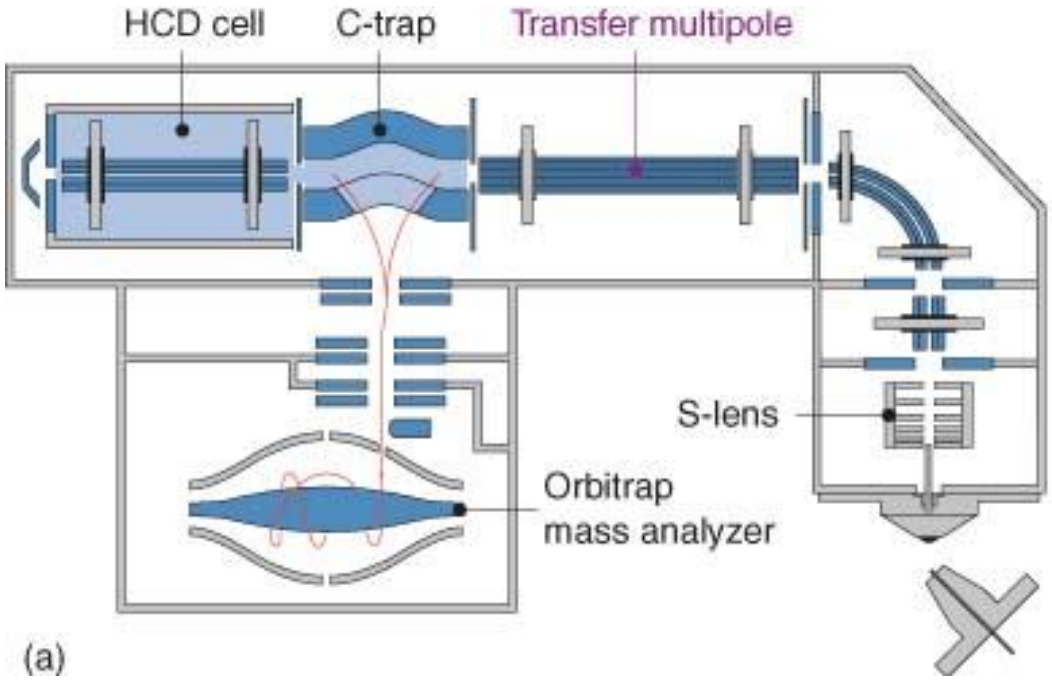
Identify and characterize metabolites based on their mass

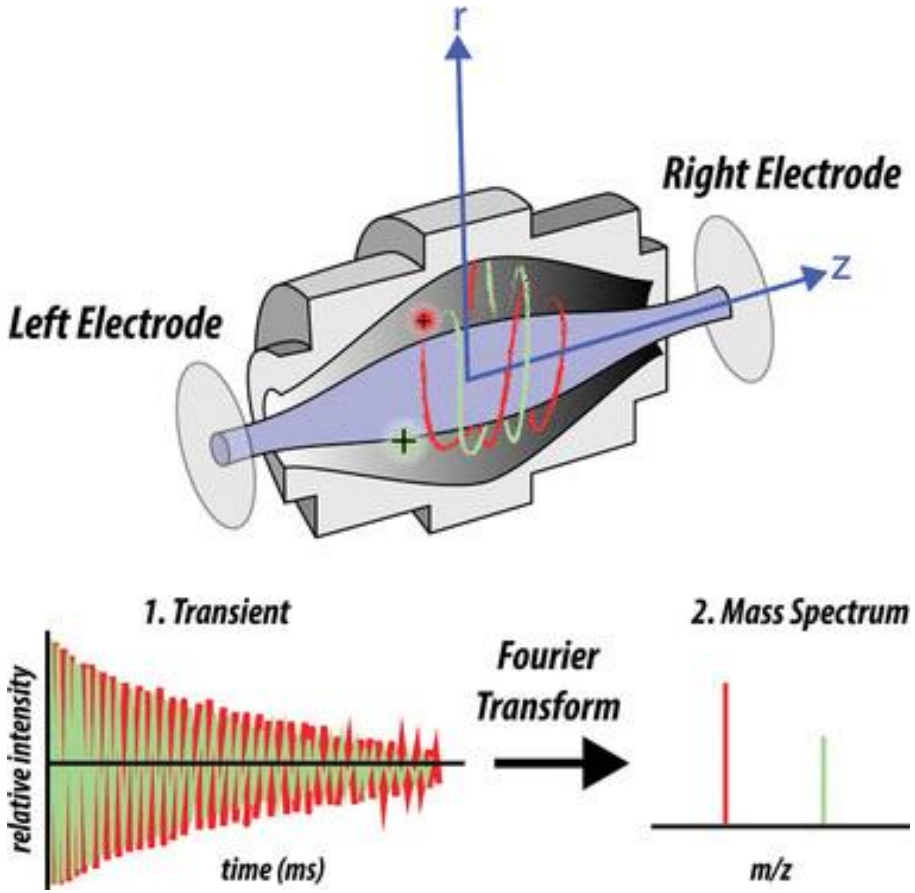
Mass Analyzers

- Once the ions are in the gas phase they can be manipulated by electric fields, magnetic fields, varying voltages and radiofrequency potentials etc….. They can be
  - Accelerated
  - Decelerated
  - Deflected
  - Reflected
  - Selected
  - Trapped
  - Brought into orbit
  - Ejected
  - Collided
  - Fragmented
  - etc….
- All mass analyzers (mass filters) separate the ions according to their *mass to charge* ratio

Mass Analyzers: Orbitrap

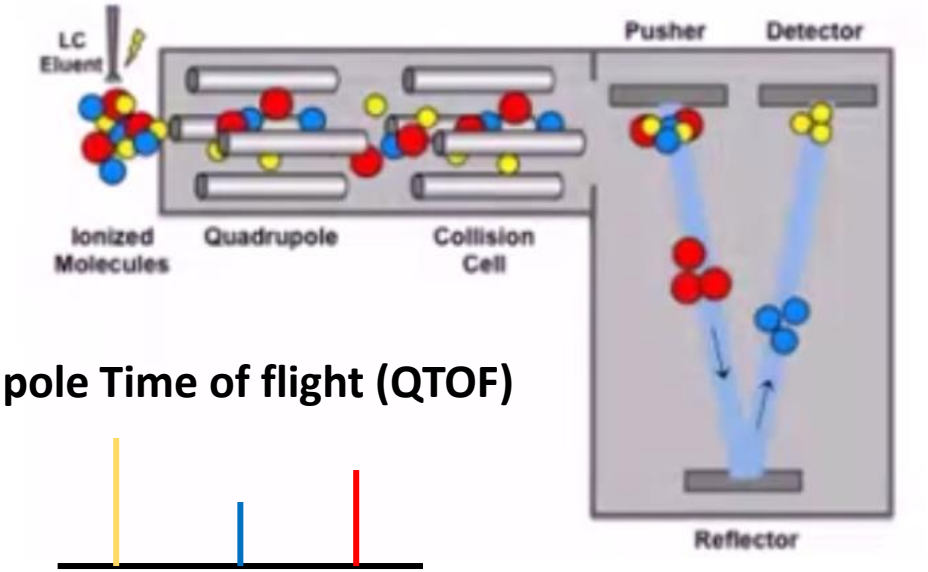
Orbitrap





https://www.researchgate.net/figure/The-Orbitrap-FT-mass-analyzer-In-the-Orbitrap-ions-oscillate-around-a-central_fig9_306542804

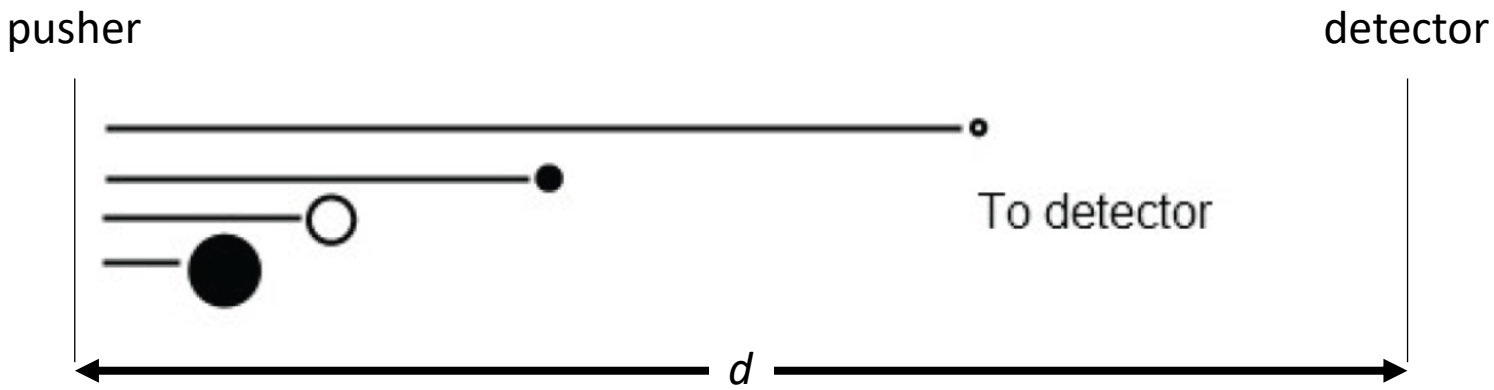
Mass Analyzers: QTOF



Quadrupole Time of flight (QTOF)

Time of Flight

- Measures the amount of time required for an ion to travel a known distance

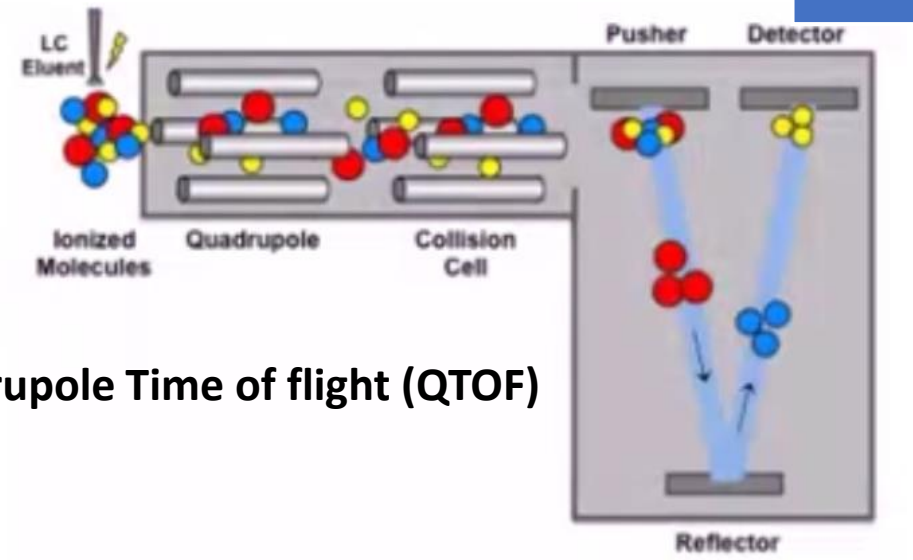


Bruker Impact II QTOF

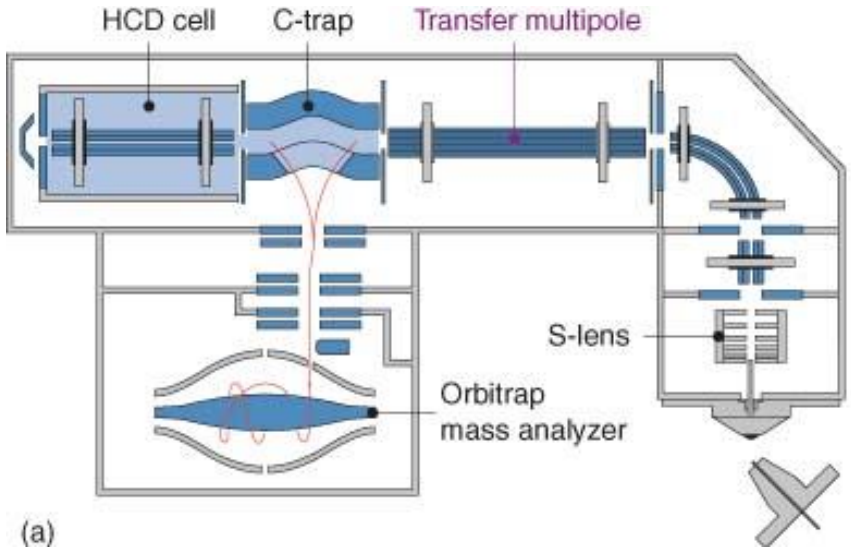
Mass Analyzers

High resolution Data

Fast data acquisition rates



Quadrupole Time of flight (QTOF)

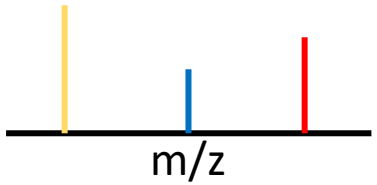


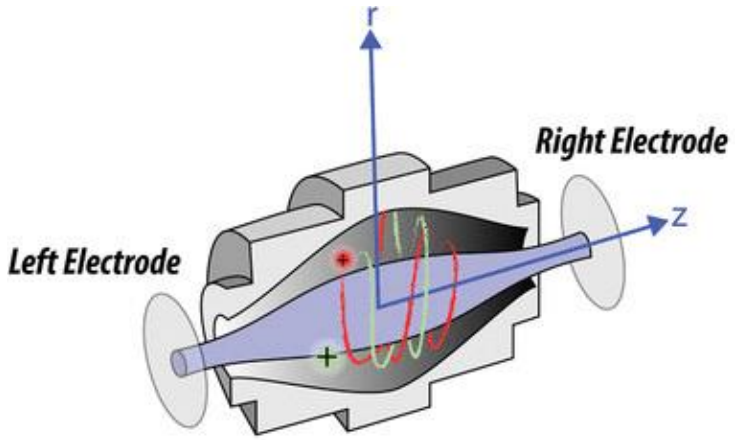
Orbitrap

Core Facility Metabolomics (CFM)
cfmetabolomics.nl

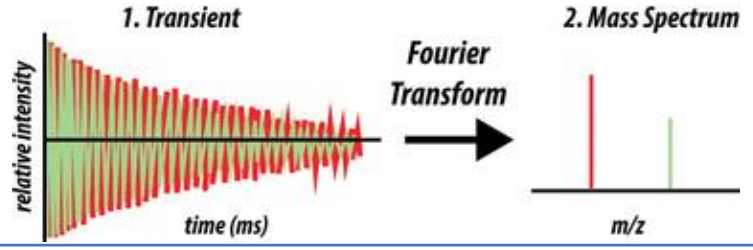
Bruker Impact II QTOF

Mass spectrum



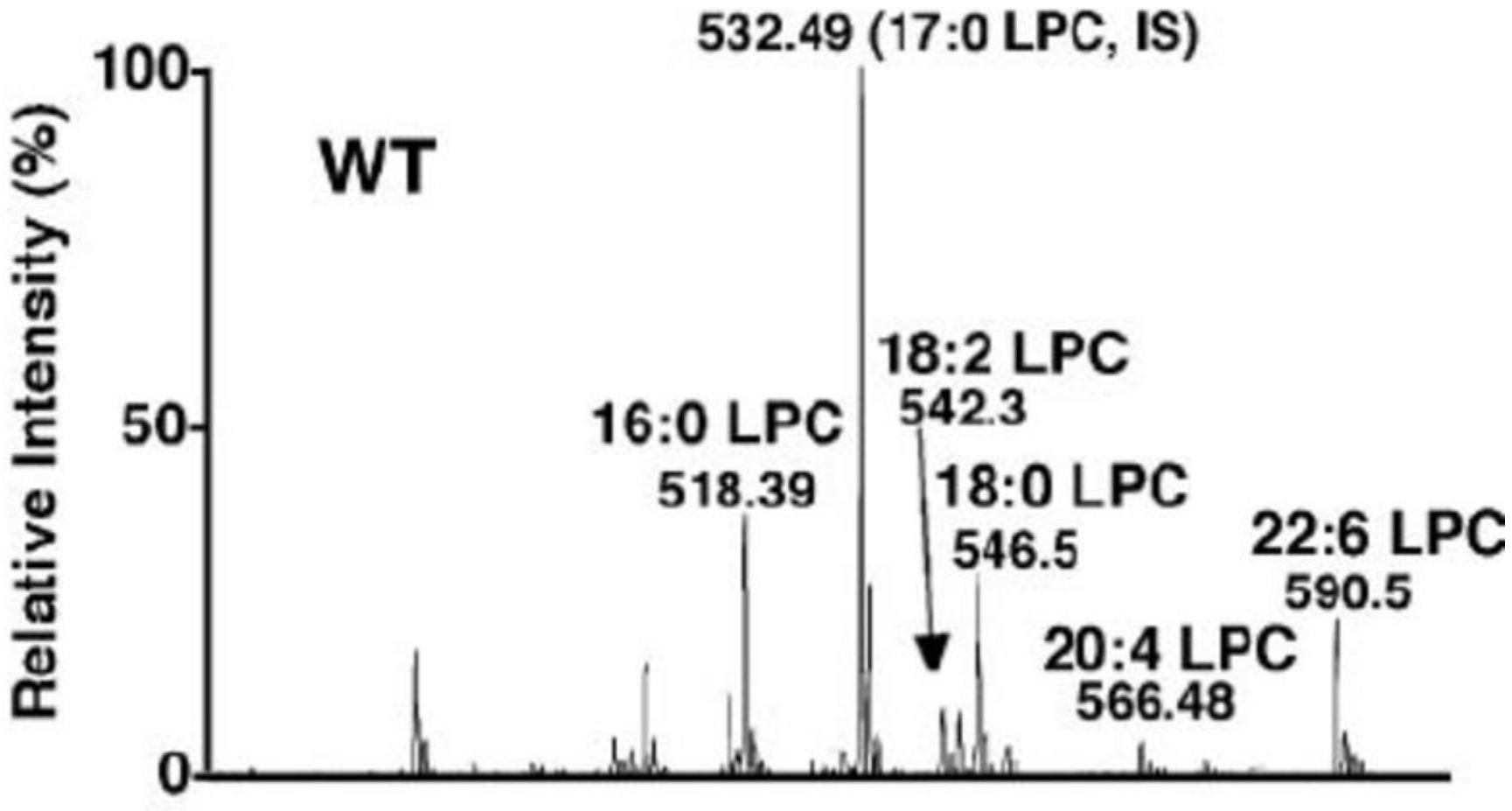


Thermo Scientific Q Exactive plus

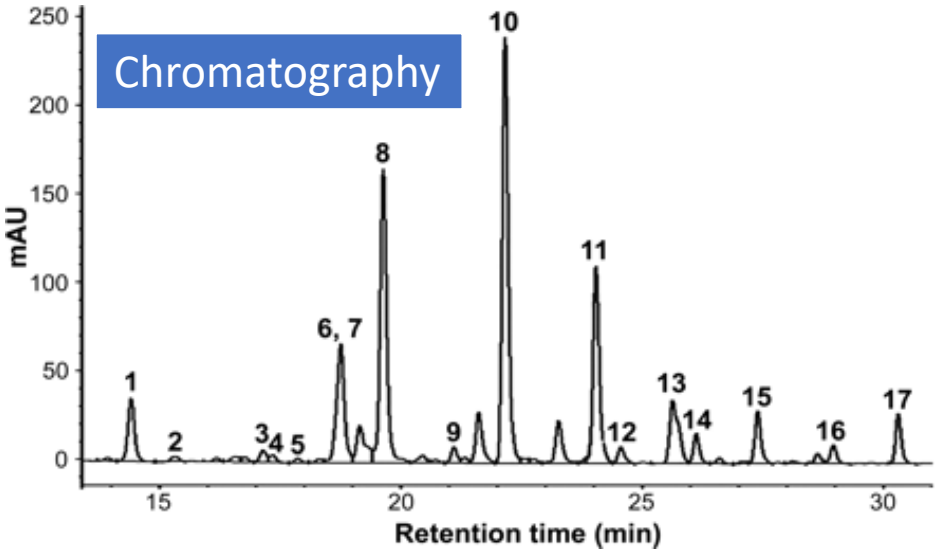


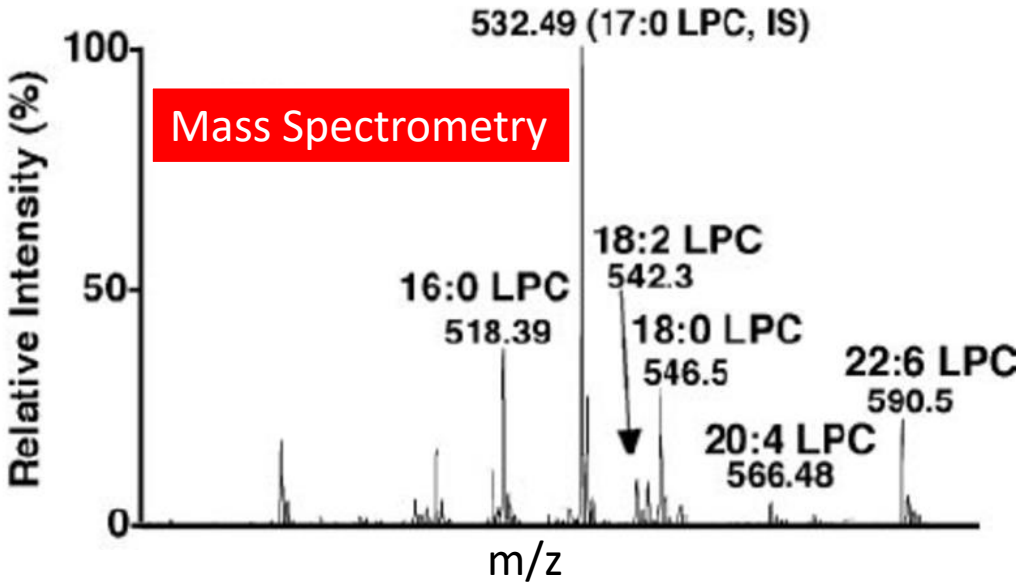
https://www.researchgate.net/figure/The-Orbitrap-FT-mass-analyzer-In-the-Orbitrap-ions-oscillate-around-a-central_fig9_306542804

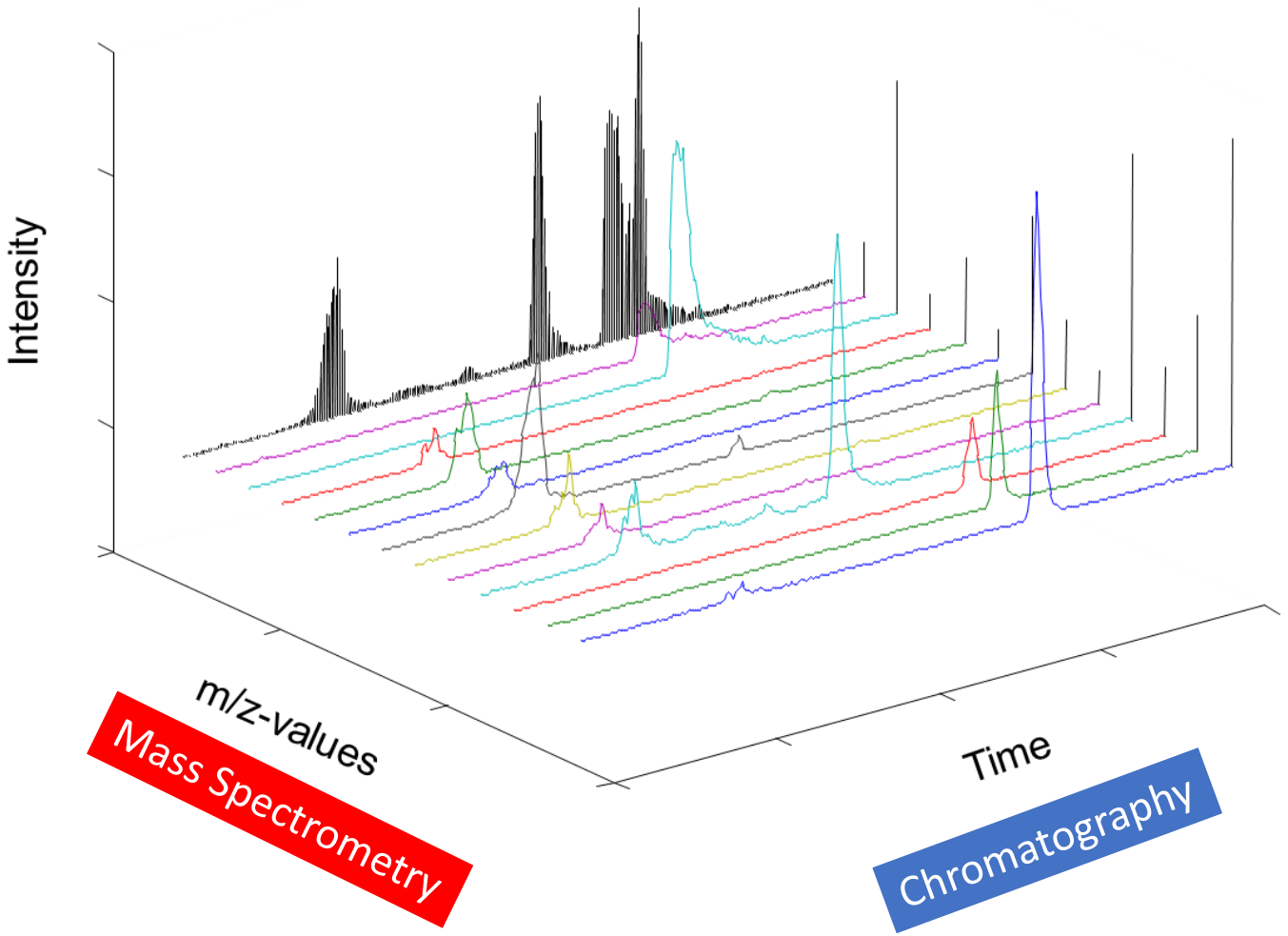
ESI MS spectrum



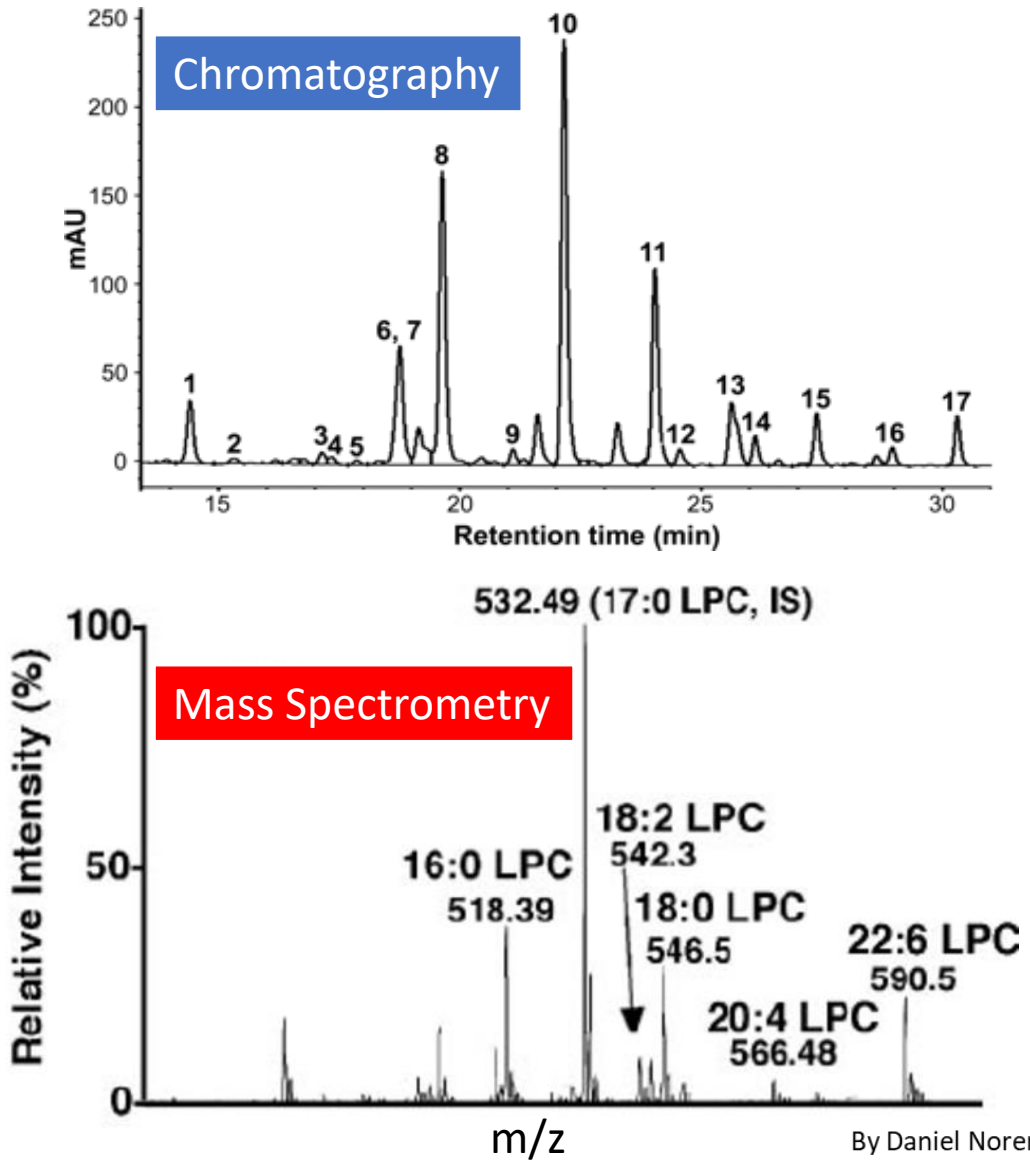
LC-MS chromatograms are shown

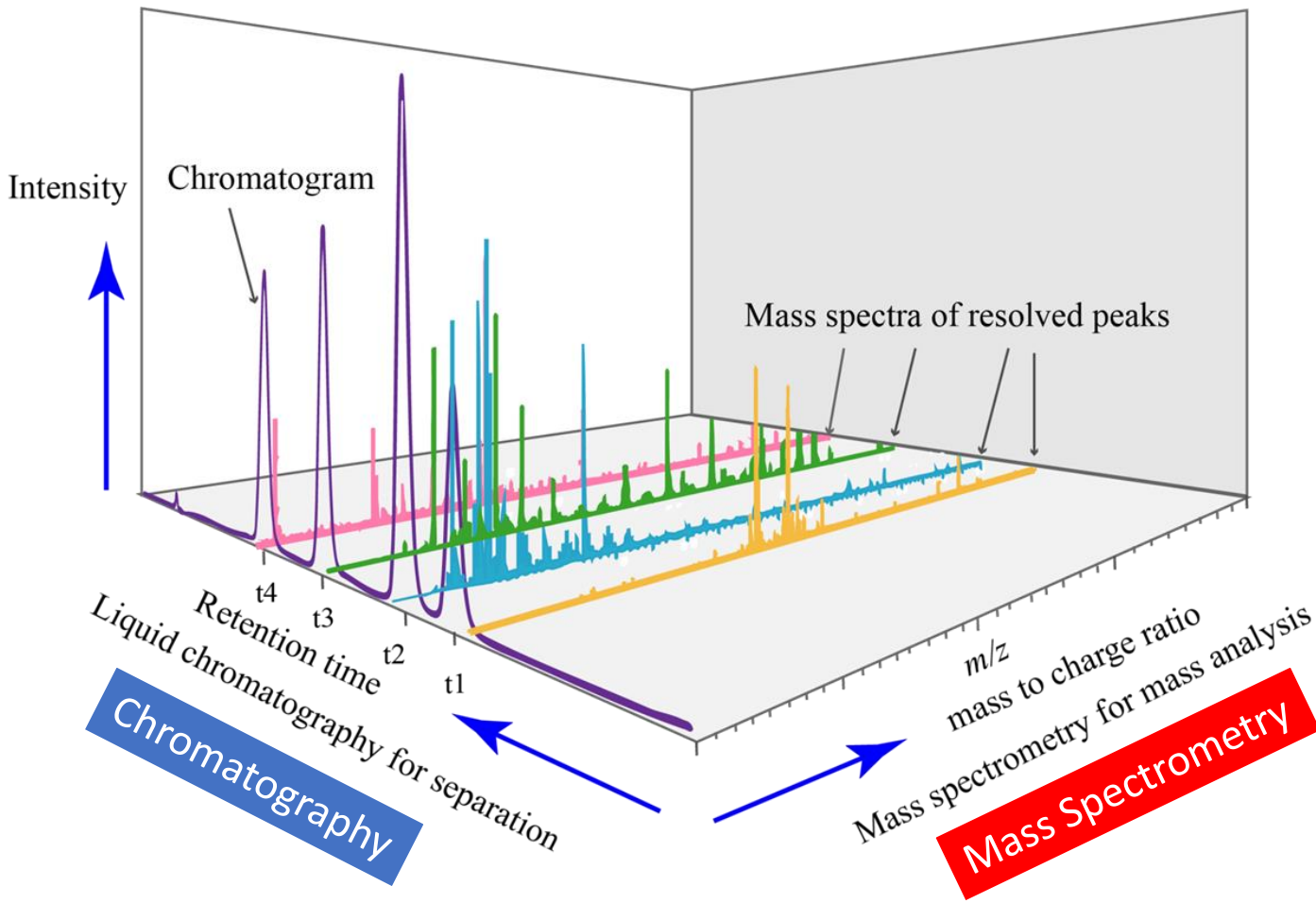






LC-MS mass spectra are shown



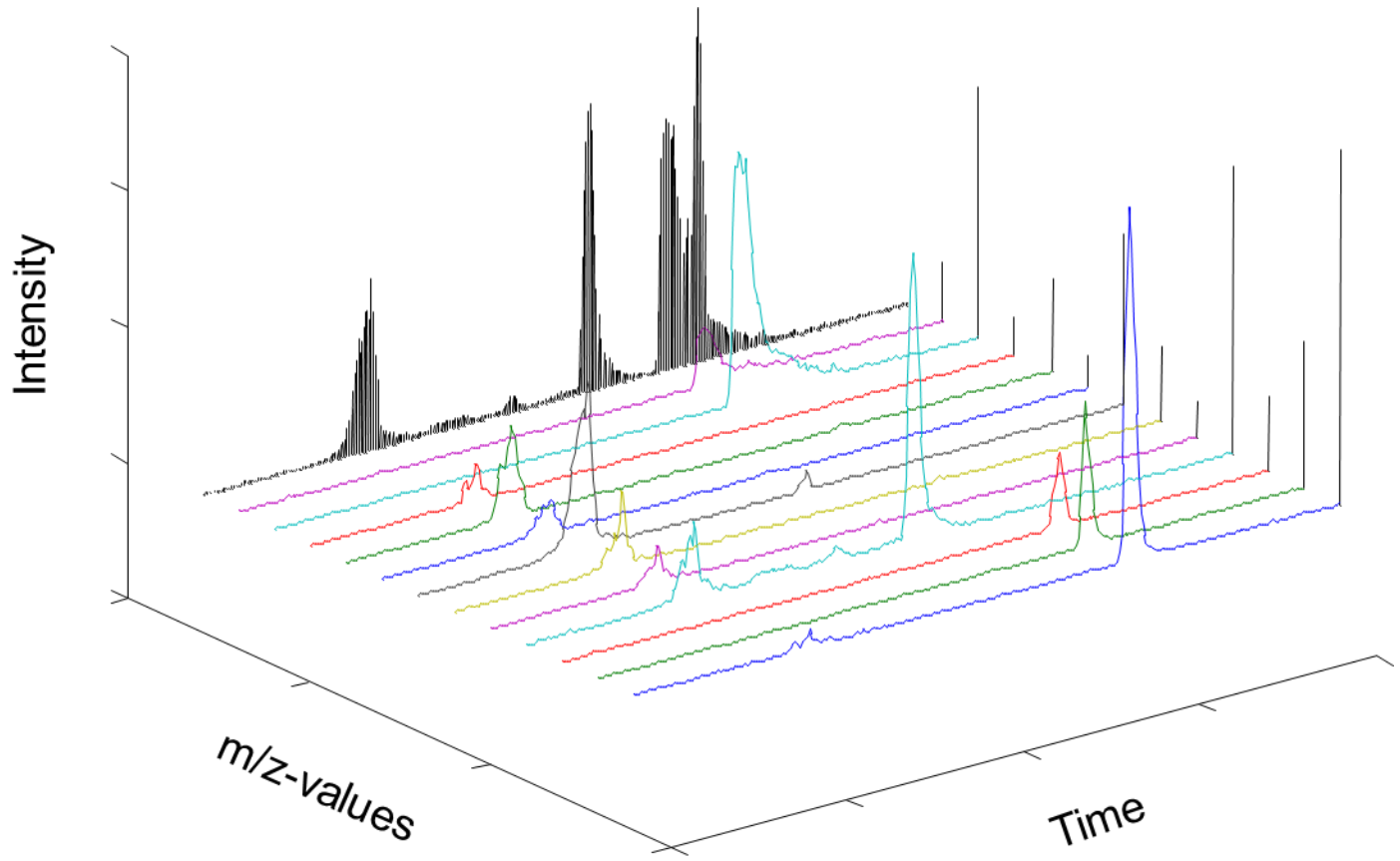


By Daniel Norena-Caro - Created in Illustrator C6, CC0, https://en.wikipedia.org/w/index.php?curid=53483722

LC-MS Based metabolomics

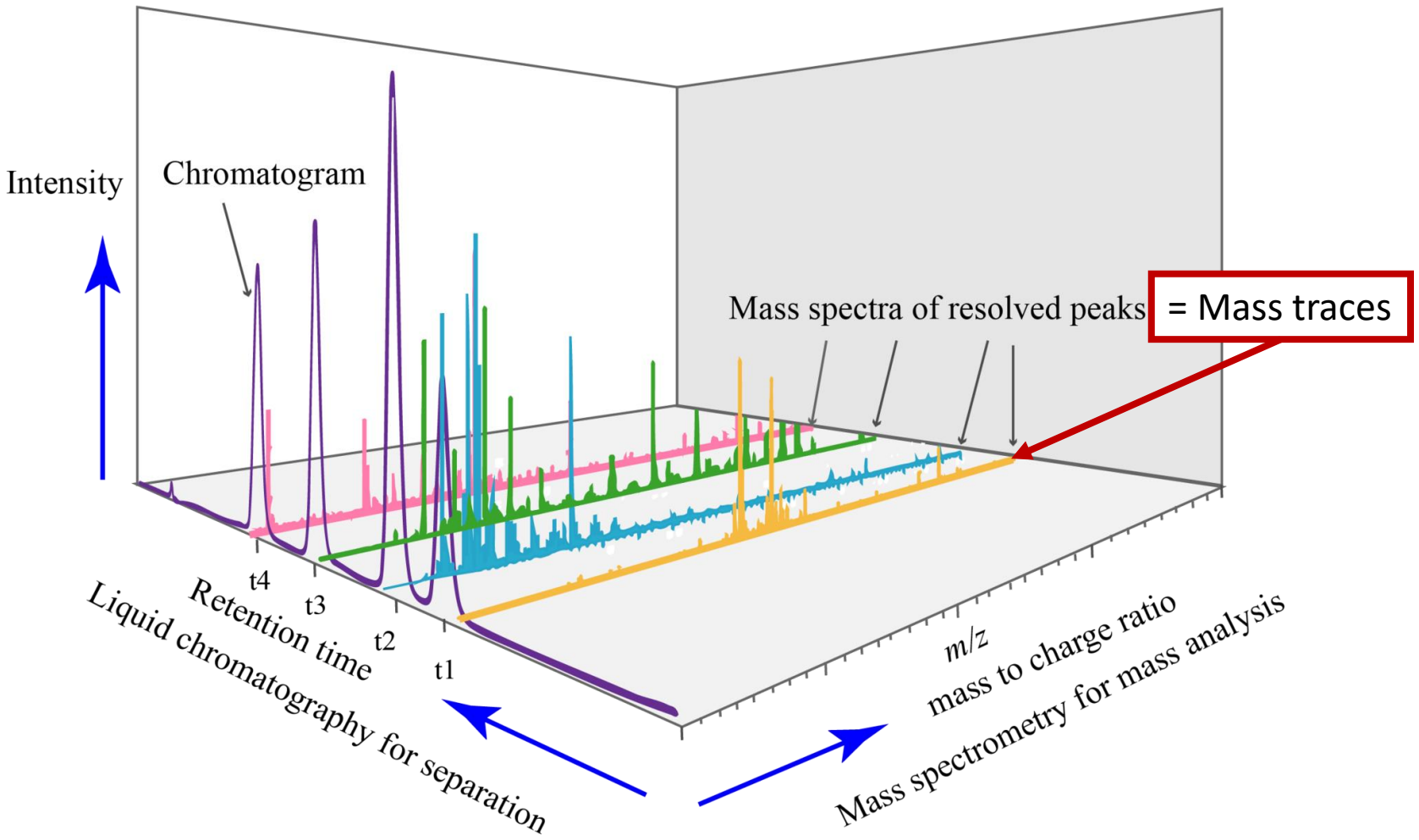
Data

LC-MS data (chromatograms are shown→ EICs)



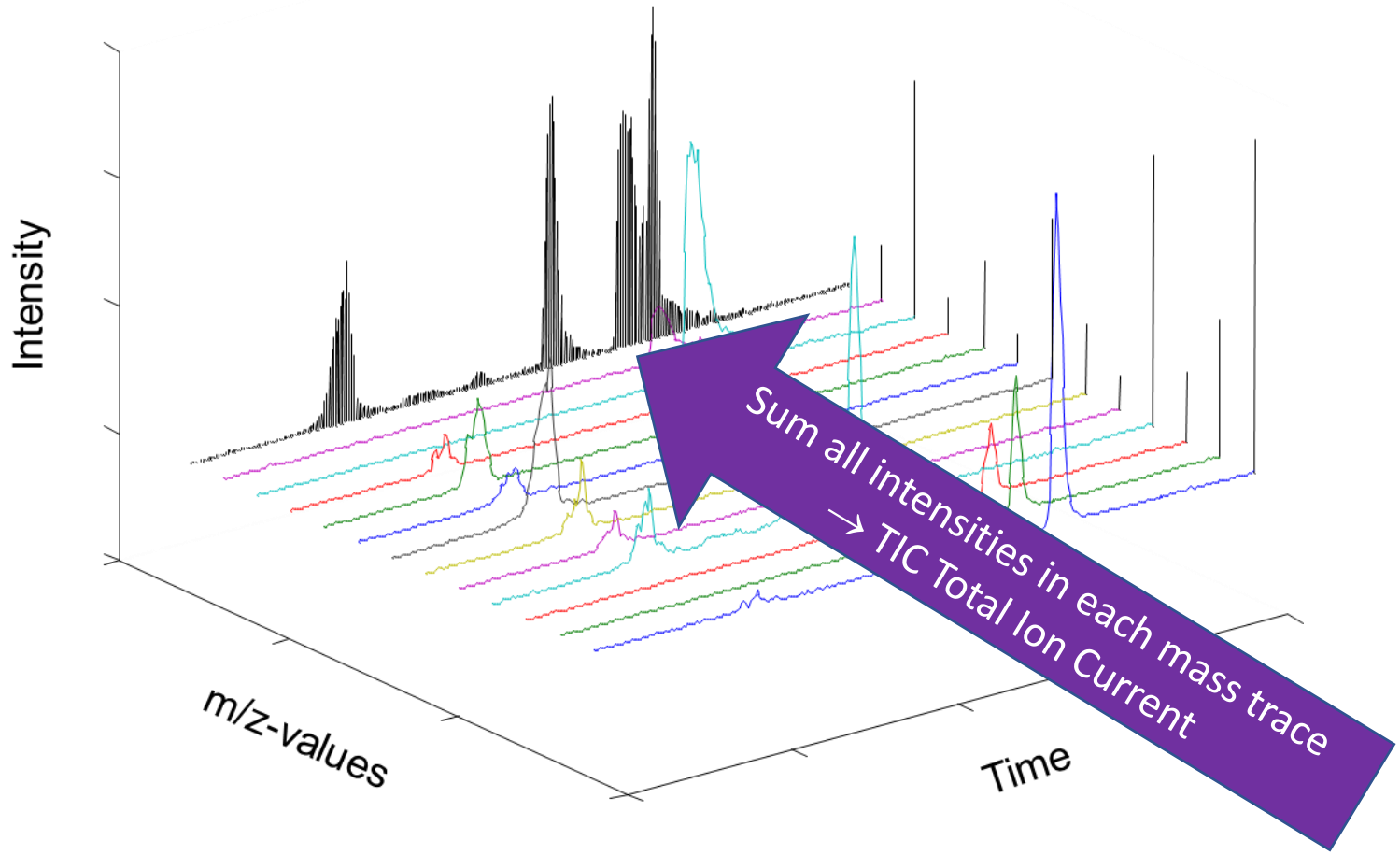
NUGO workshop 2007, TNO

LC-MS data (mass spectra are shown)



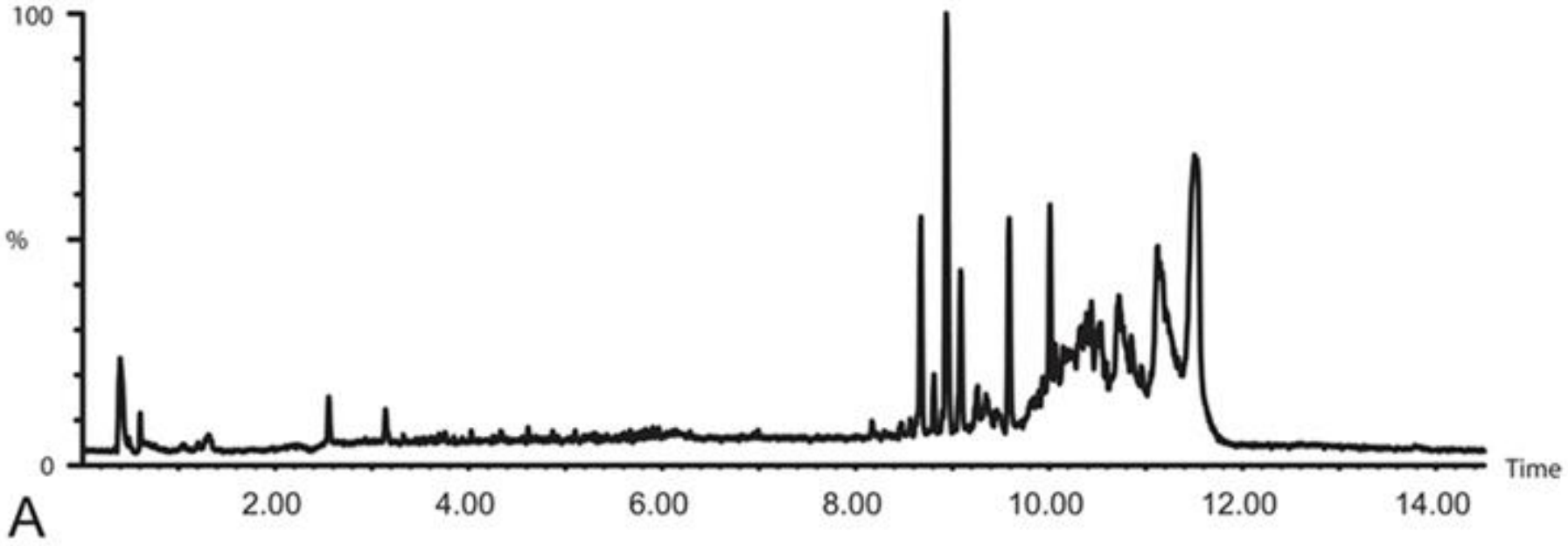
By Daniel Norena-Caro - Created in Illustrator C6, CC0, https://en.wikipedia.org/w/index.php?curid=53483722

LC-MS data (chromatograms are shown→ EICs)



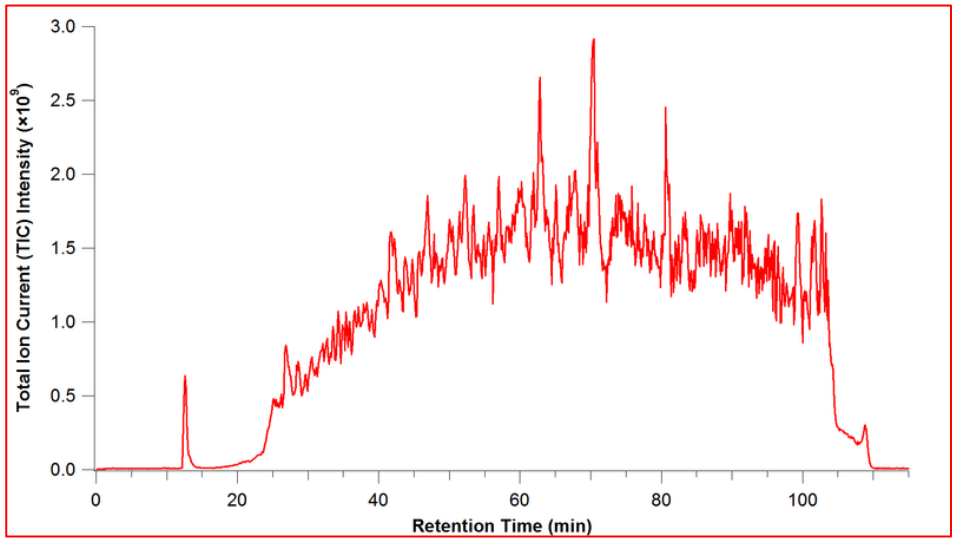
NUGO workshop 2007, TNO

Total Ion Current chromatogram (example)



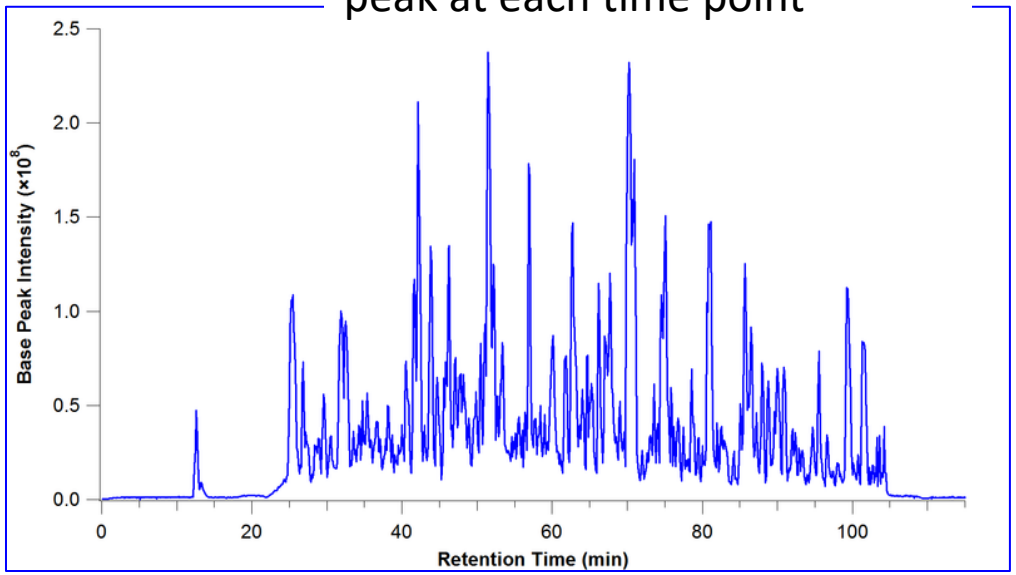
TIC and base peak chromatogram

Total Ion Current chromatogram



chromatogram of most intense peak at each time point

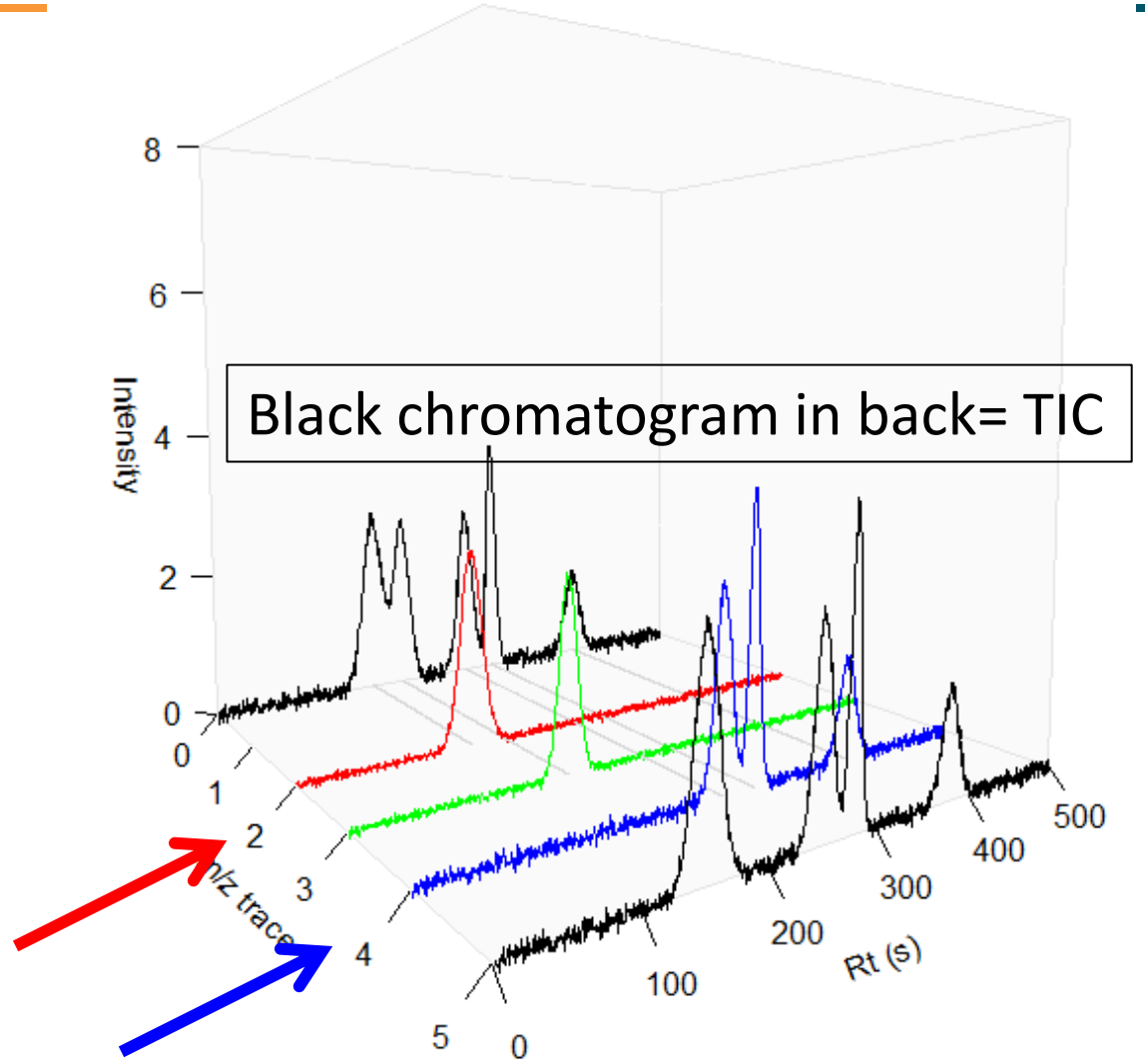
Base peak chromatogram



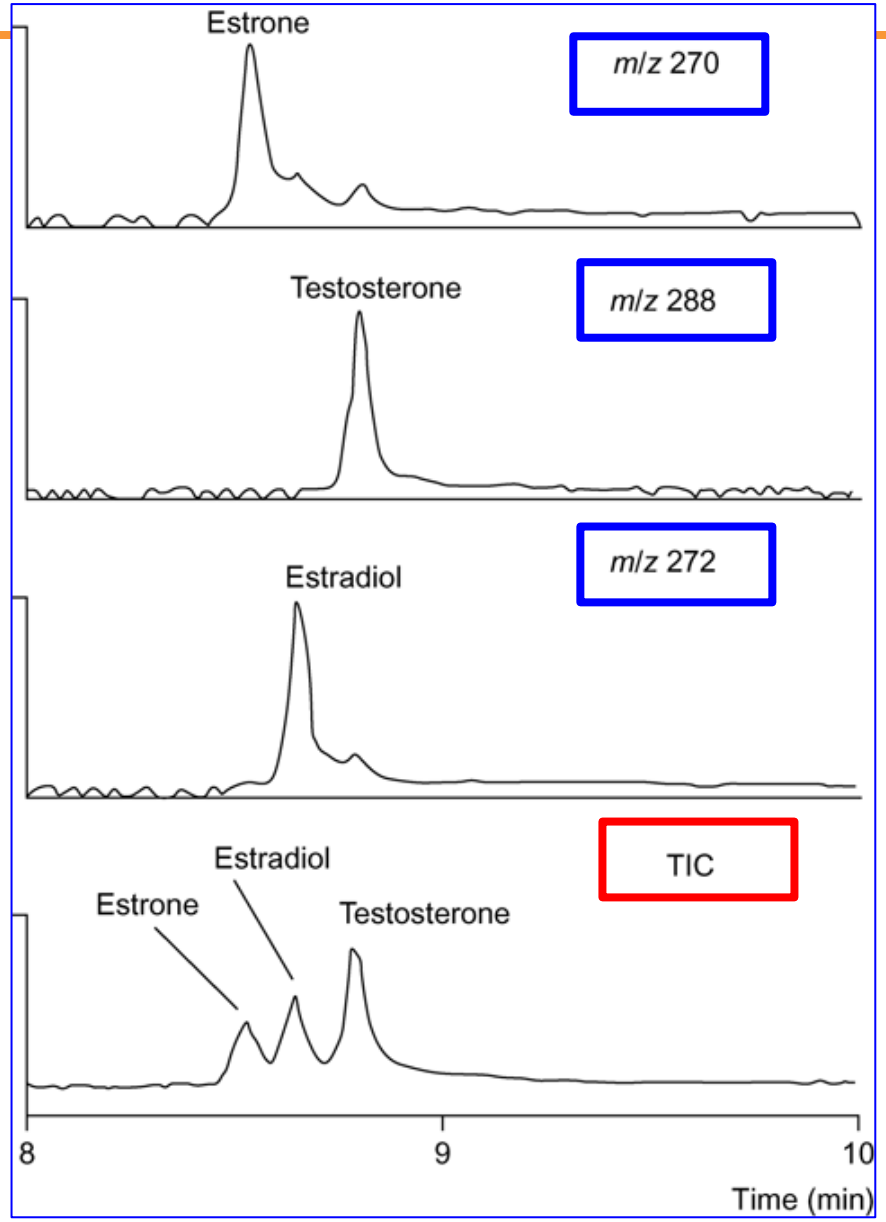
Extracted Ion Chromatogram (EIC)

Reconstruct chromatogram
for specific m/z value(s)

Example:
Use RED and BLUE m/z values



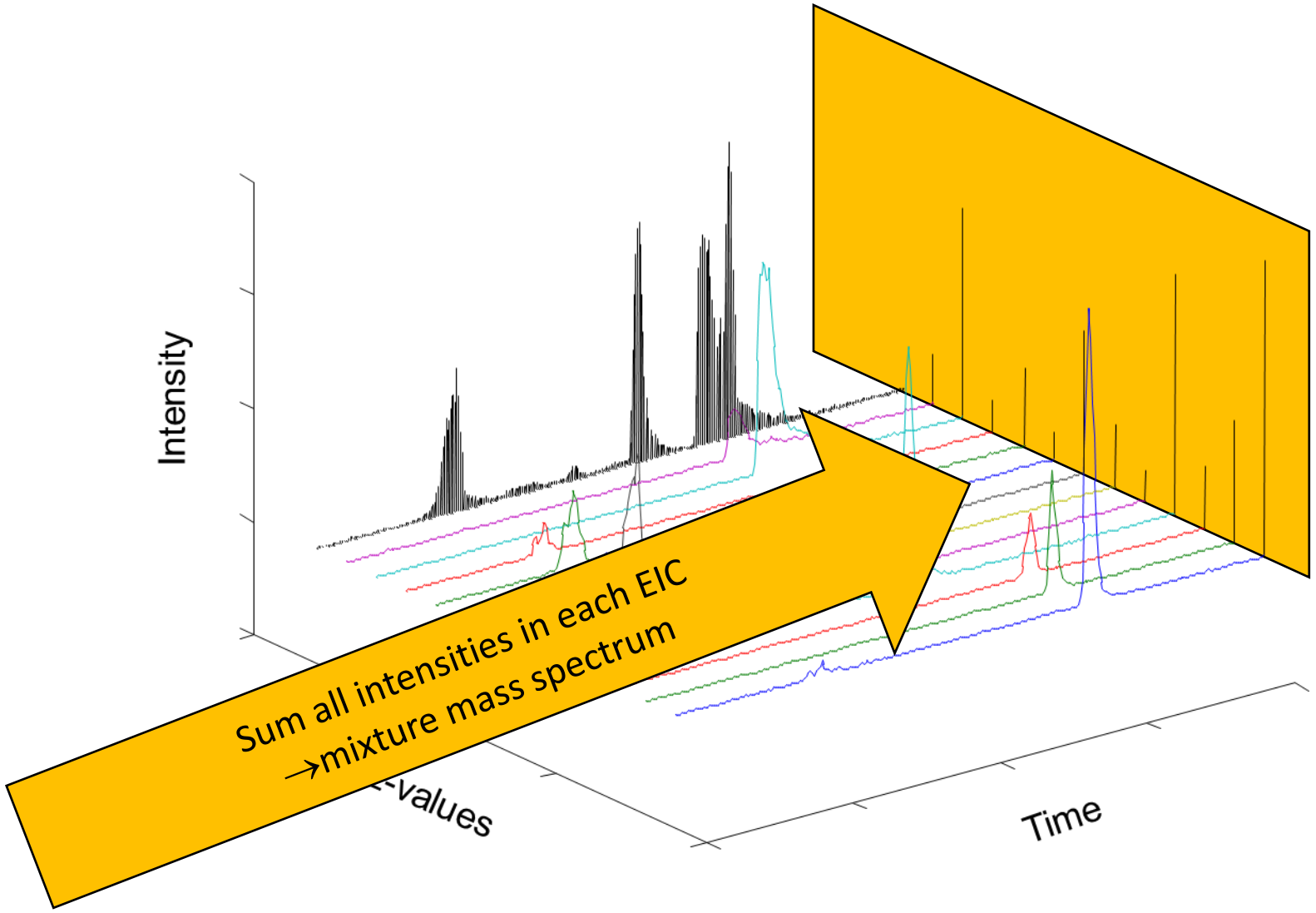
Extracted Ion Chromatogram (EIC)



Reconstruction of chromatogram for specific m/z values

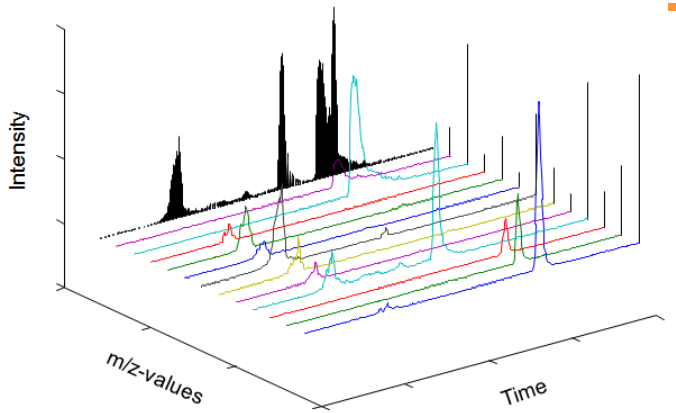
Sum of the above three chromatograms (= intensities of m/z values)

LC-MS data (chromatograms are shown→ EICs)



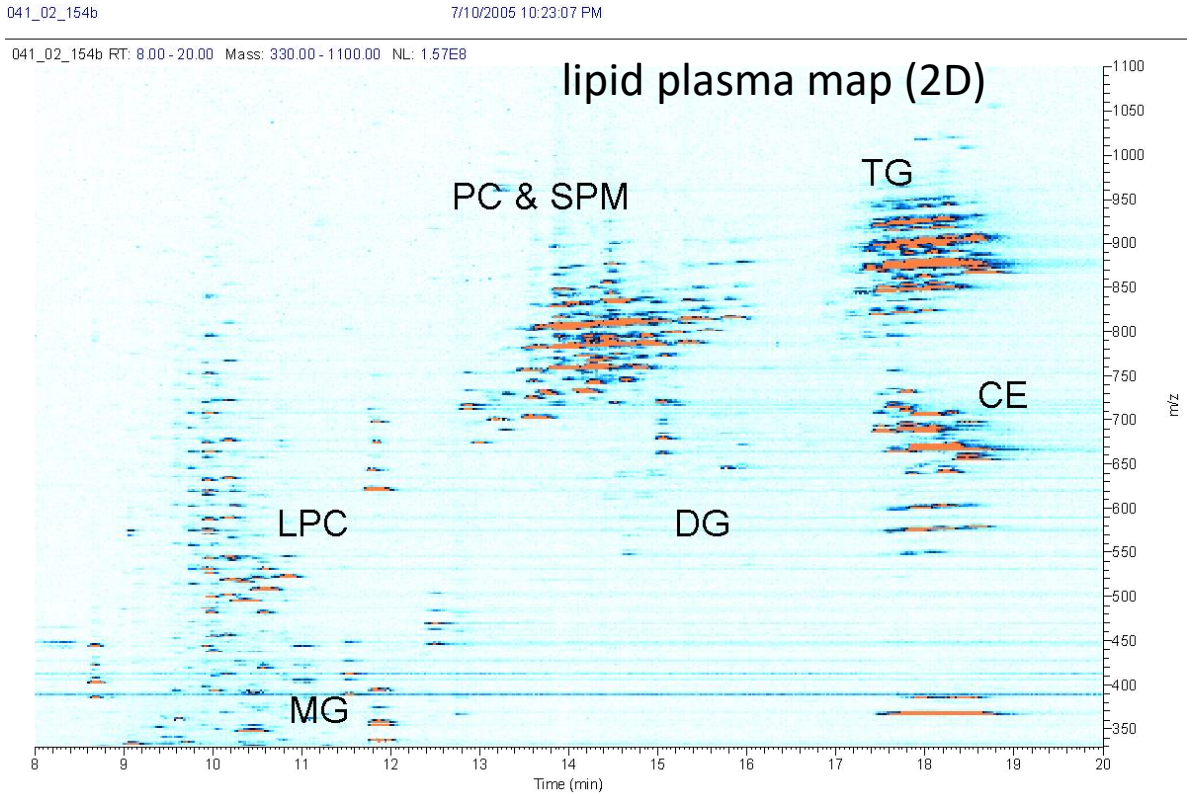
NUGO workshop 2007, TNO

LC-MS data

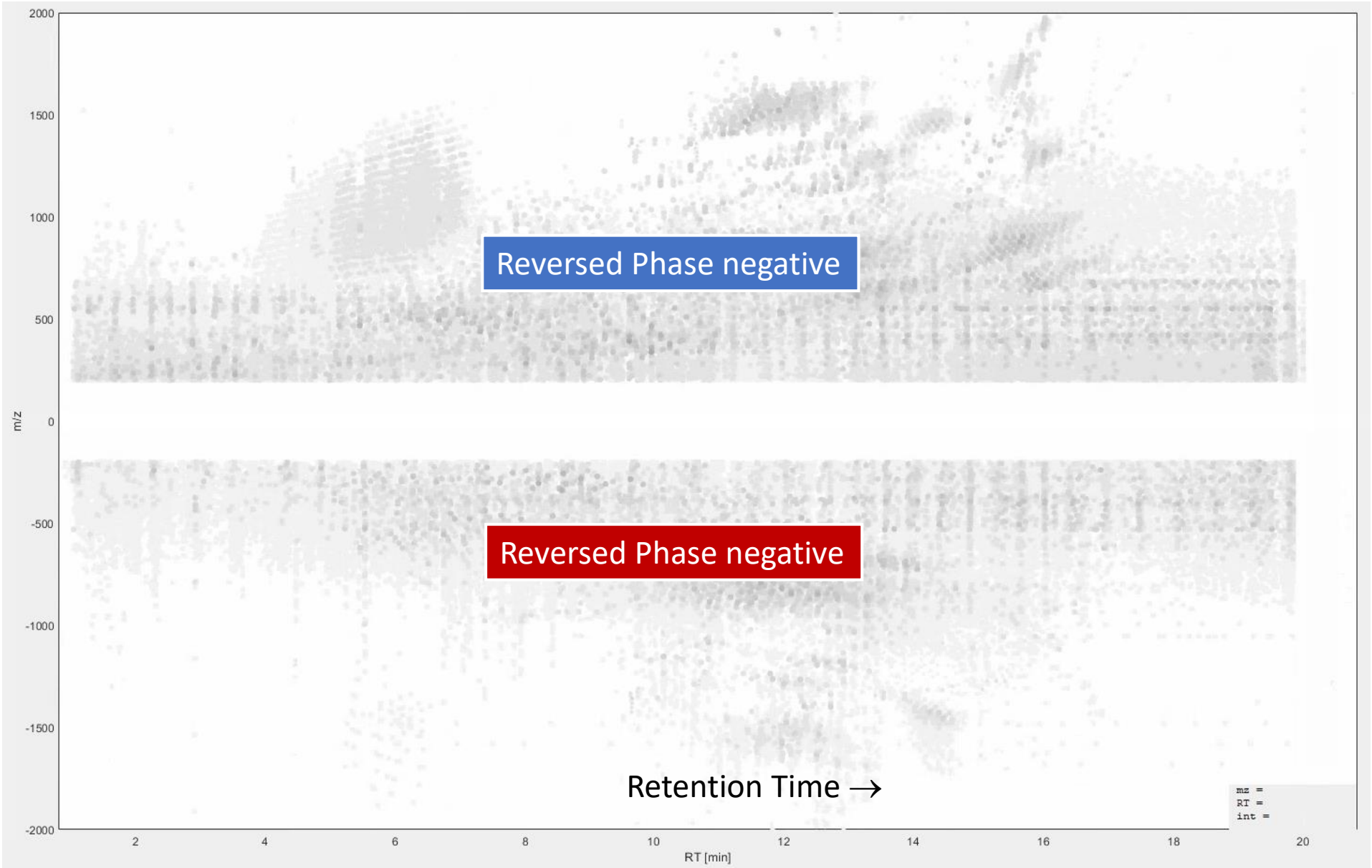


3D visualization

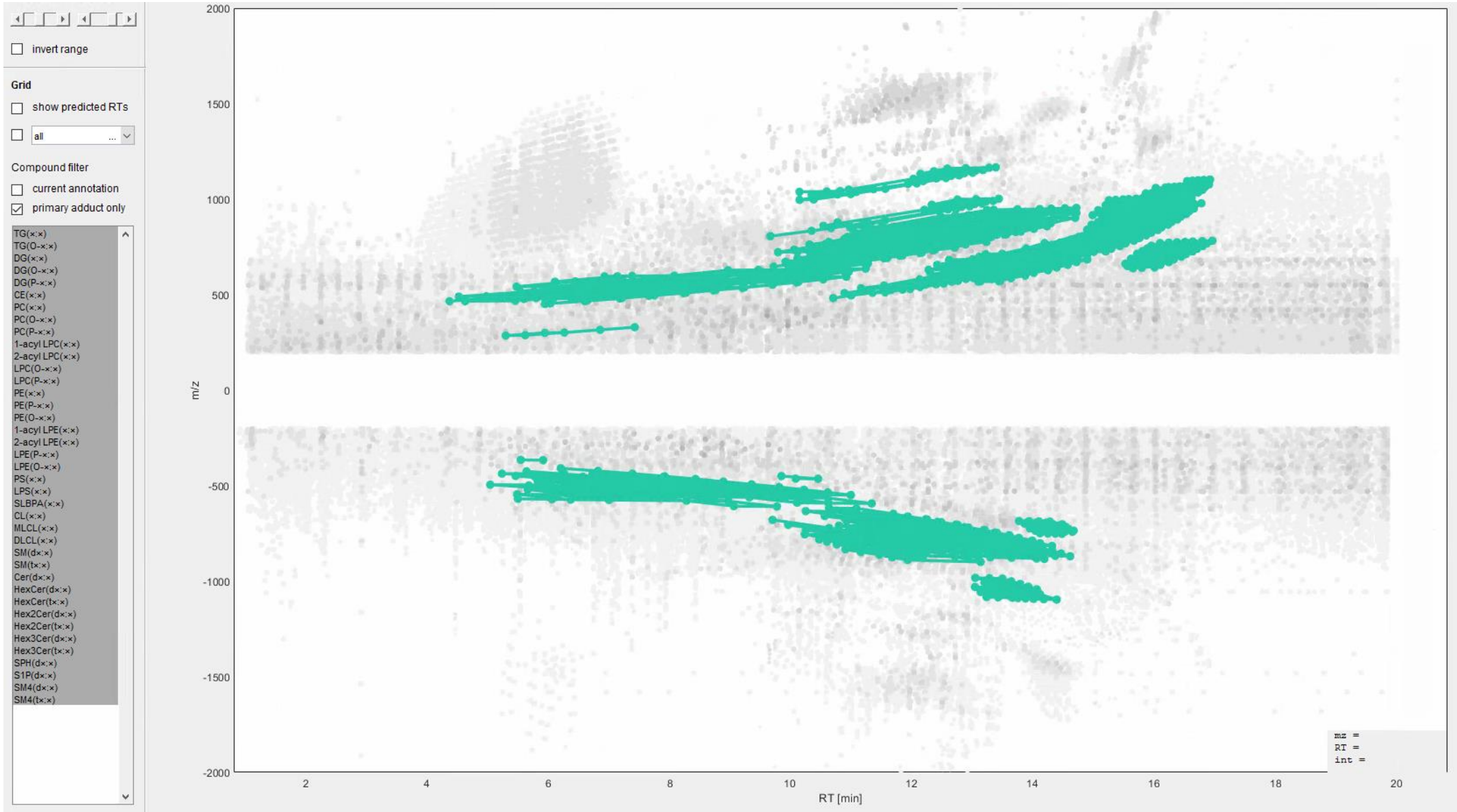
Heatmap



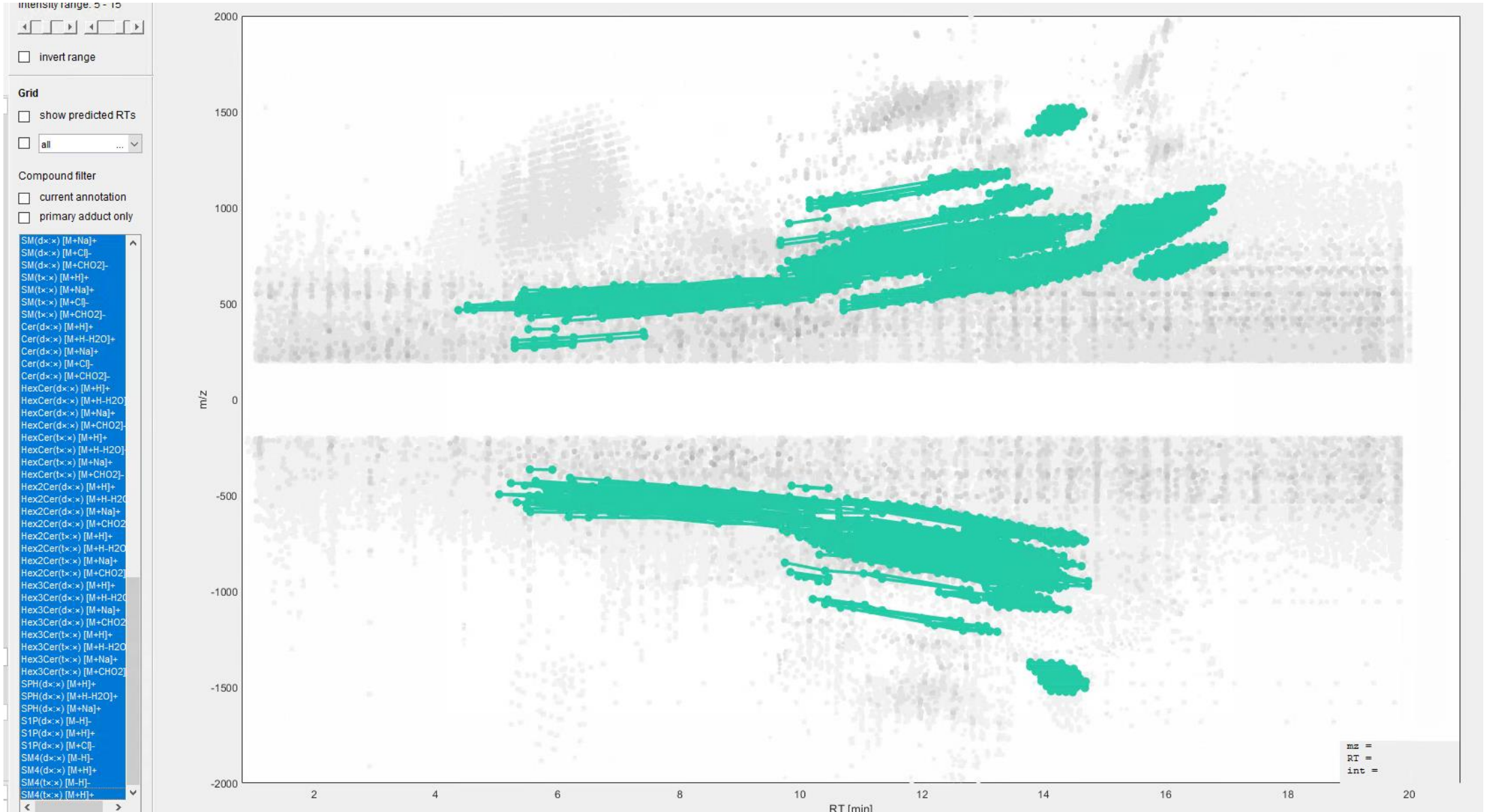
Lipidomics Gray levels represent intensity



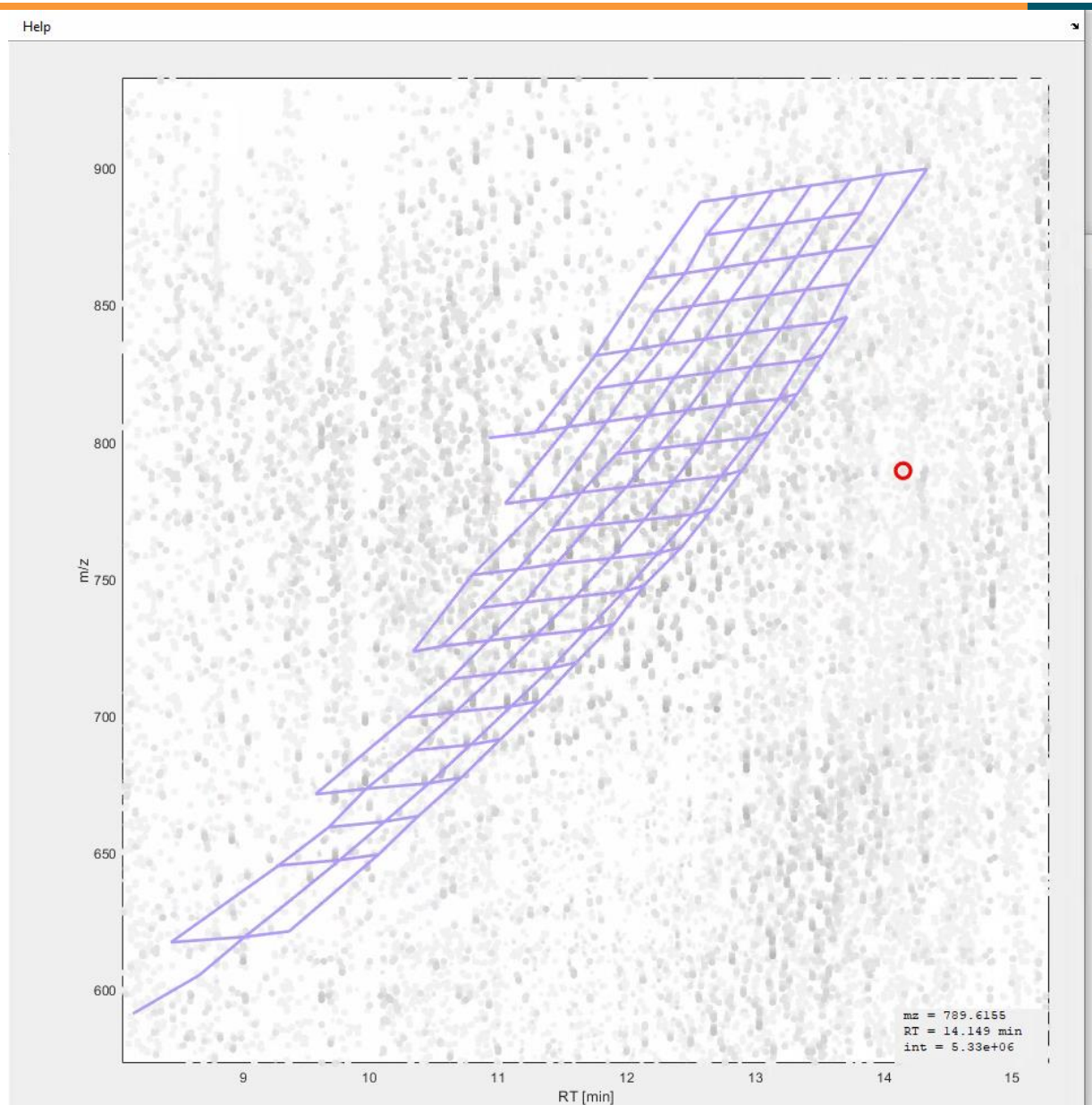
Targeted Lipidomics

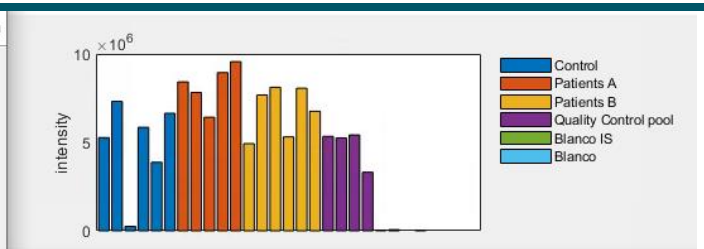


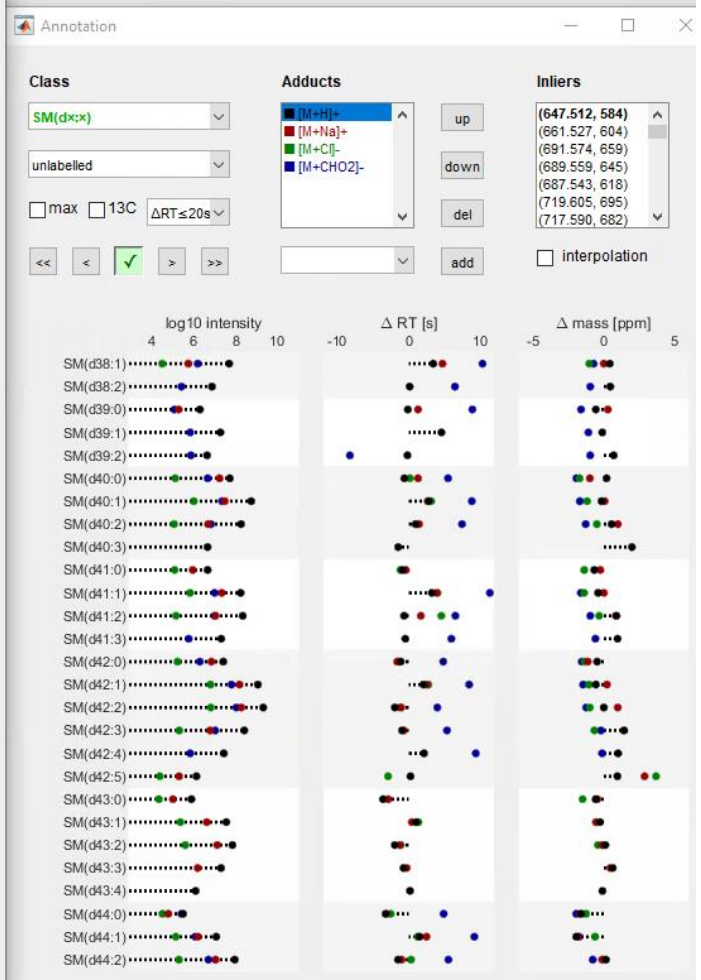
Targeted Lipidomics (more ions)



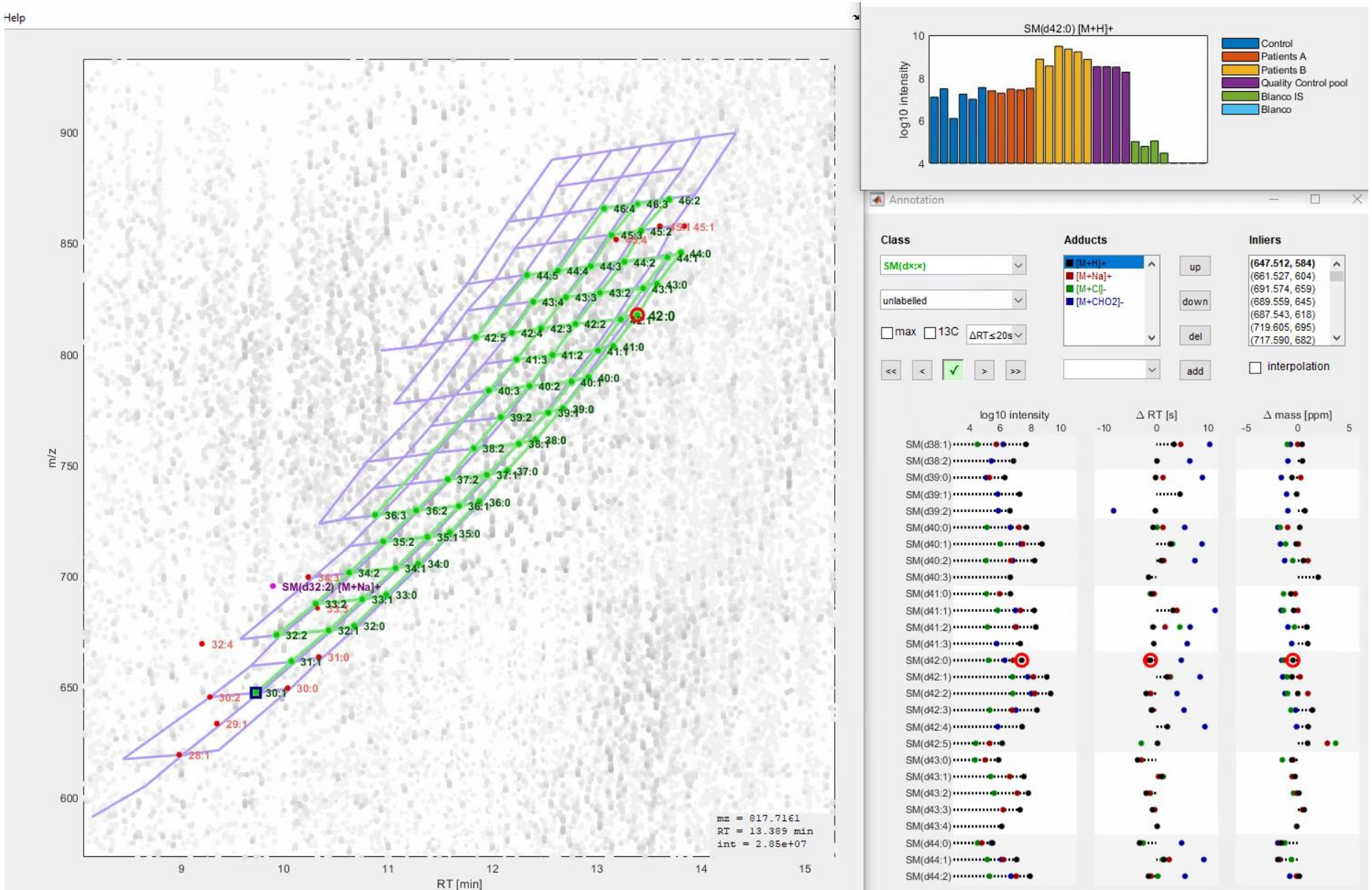
Targeted Lipidomics (Targets)



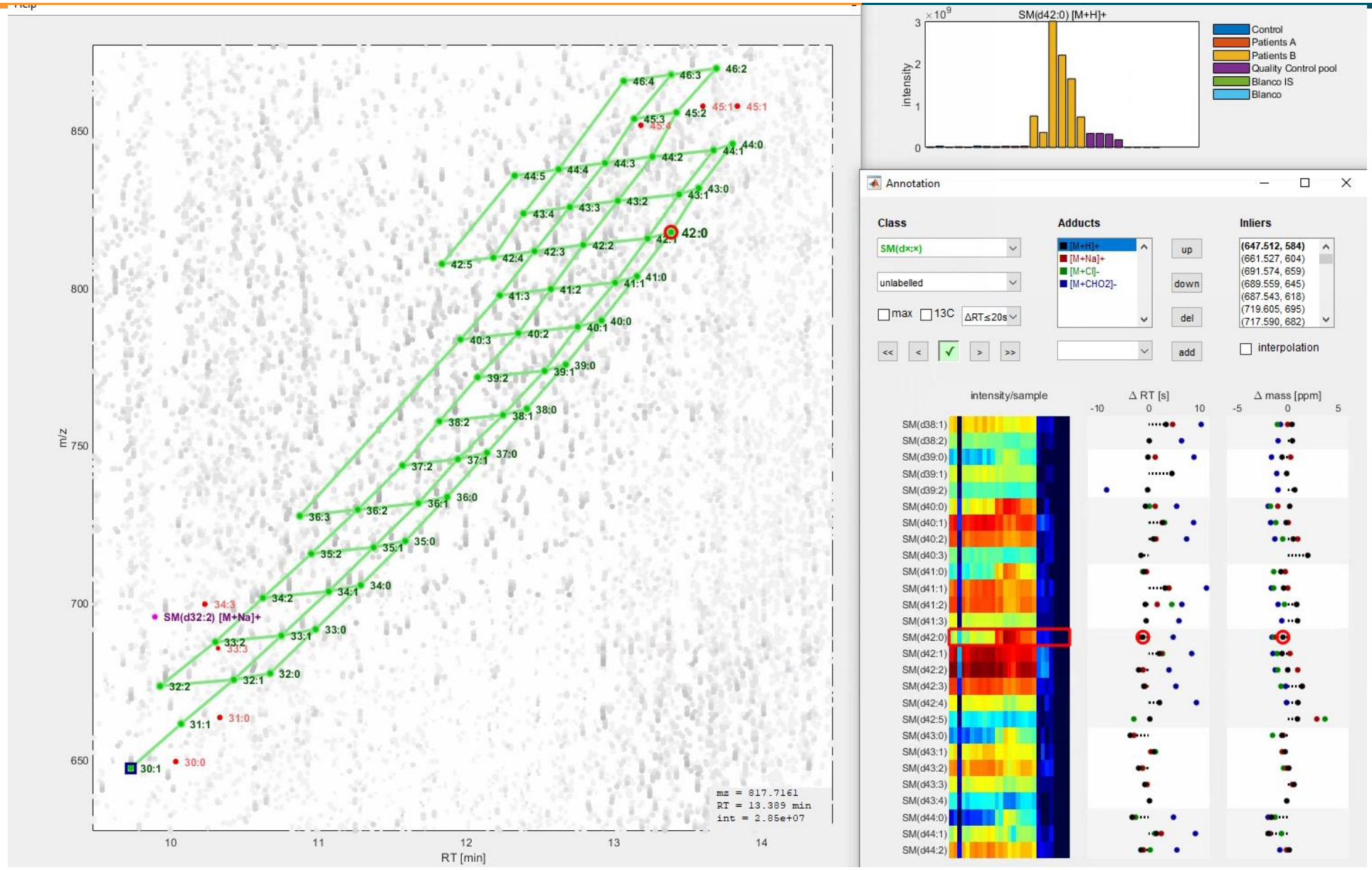




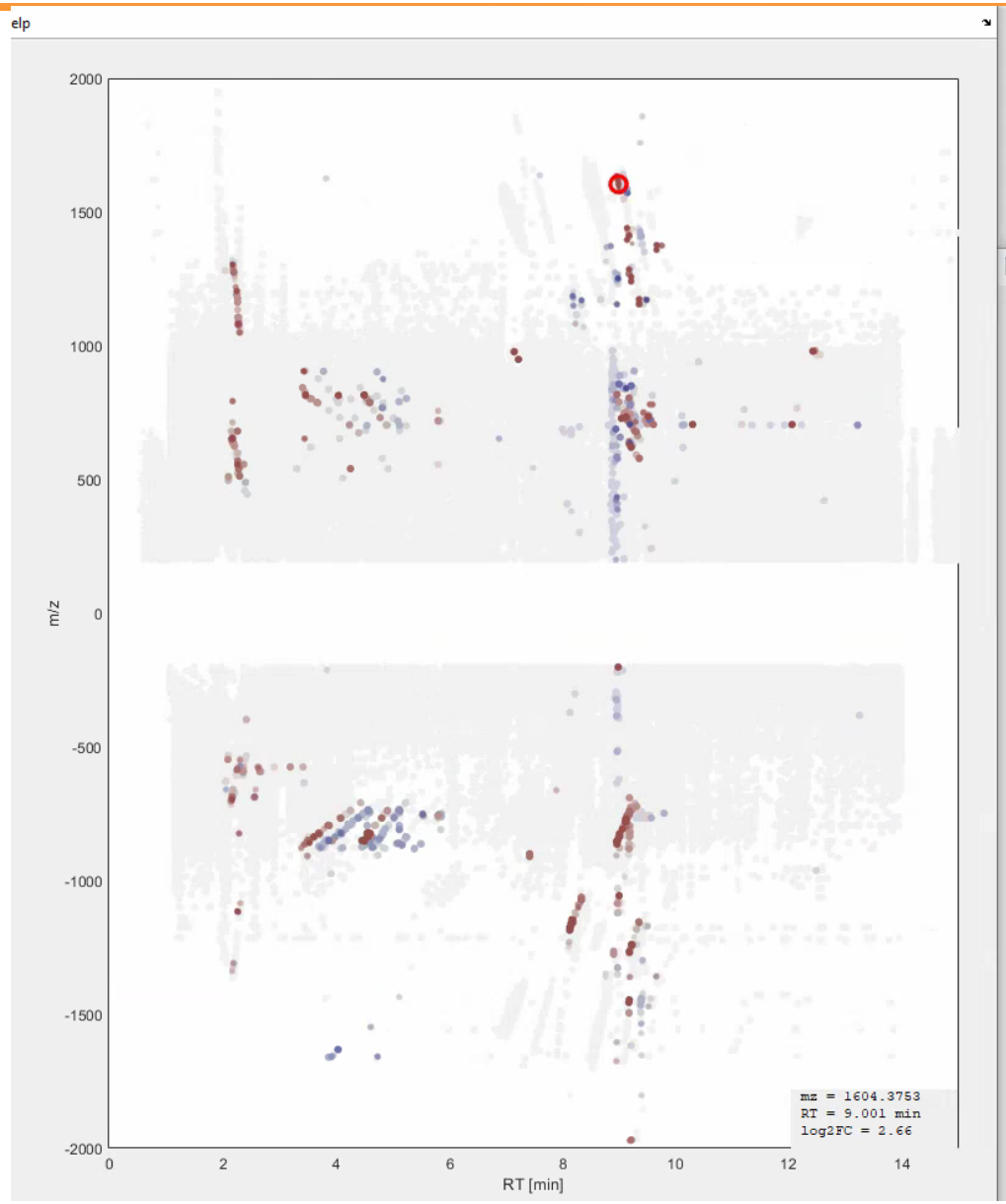
Targeted Lipidomics (Targets found)

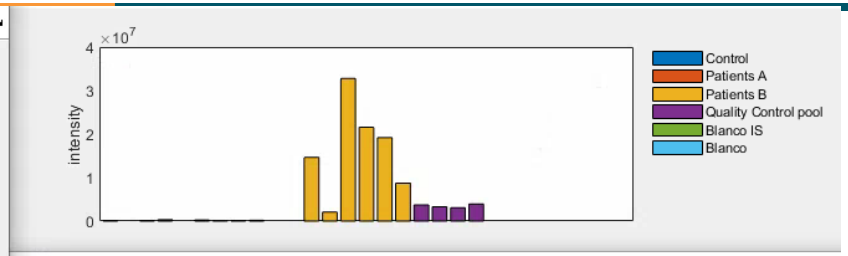


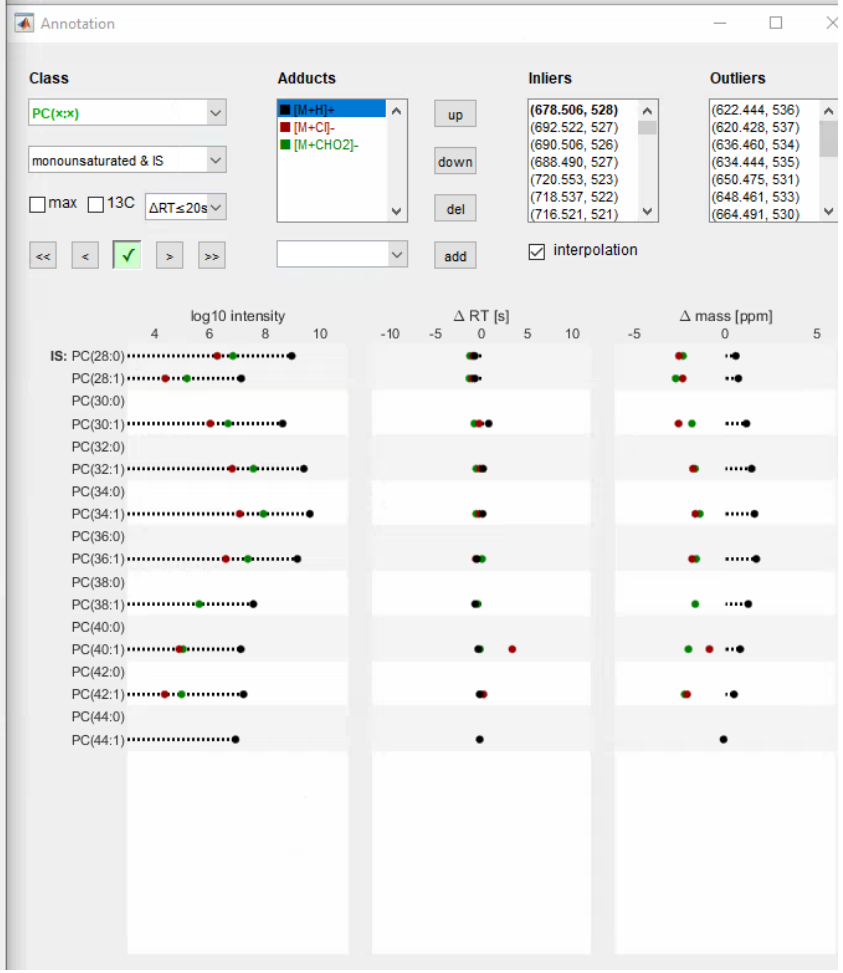
Targeted Lipidomics (Annotated targets and contrasts)



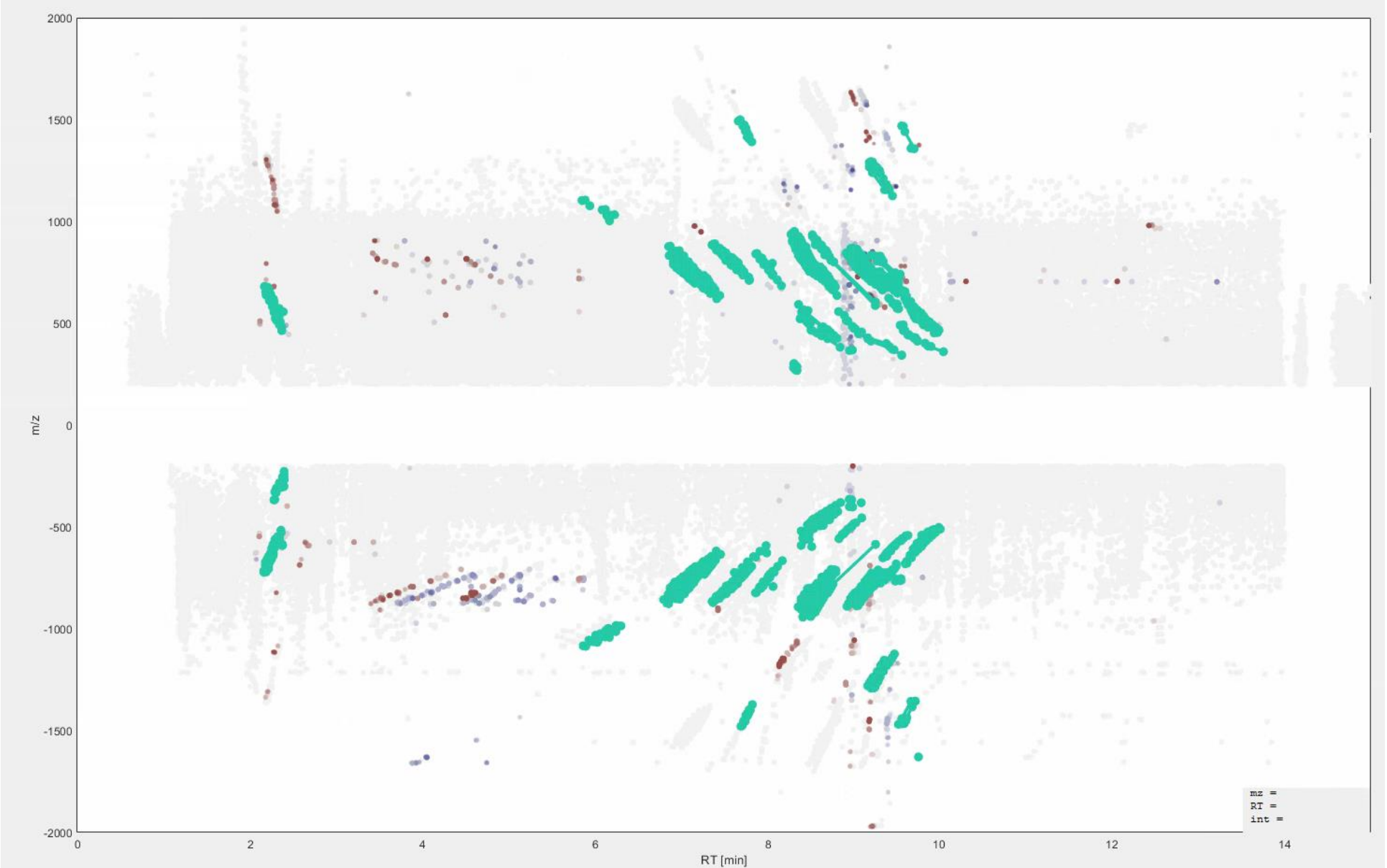
Untargeted Lipidomics





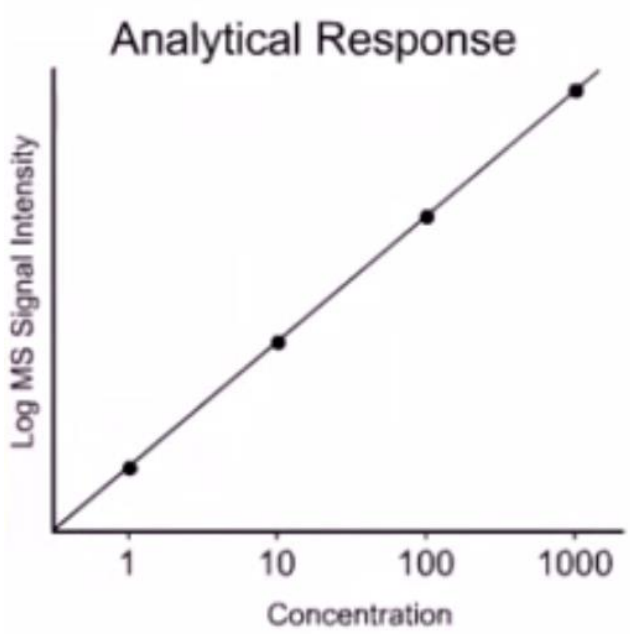


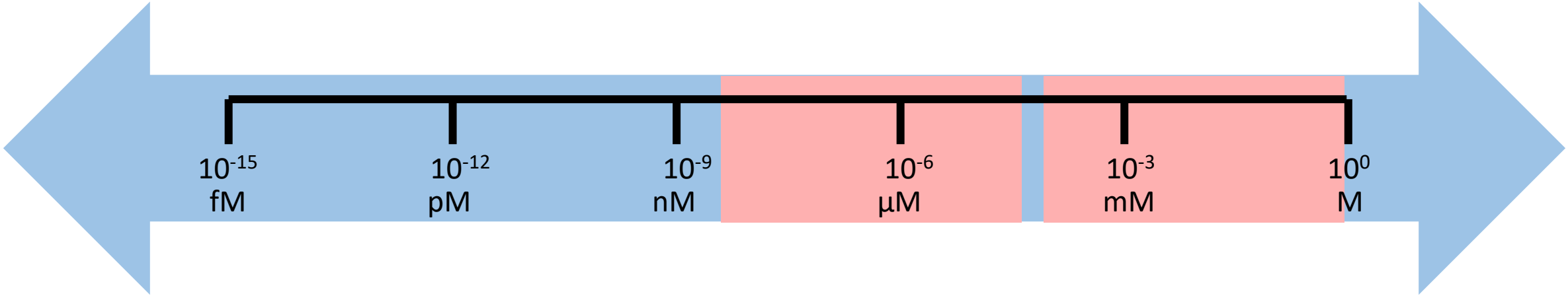
Untargeted (targeted shown)



LC-MS based metabolomics

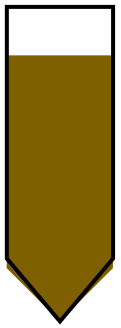
- LC-MS Measures Hundreds of metabolites over 4 orders of magnitude in concentration in a single run
- A single analysis can't provide full coverage
- Lower concentrations require more sample work effort and biomass





LC-MS based metabolomics

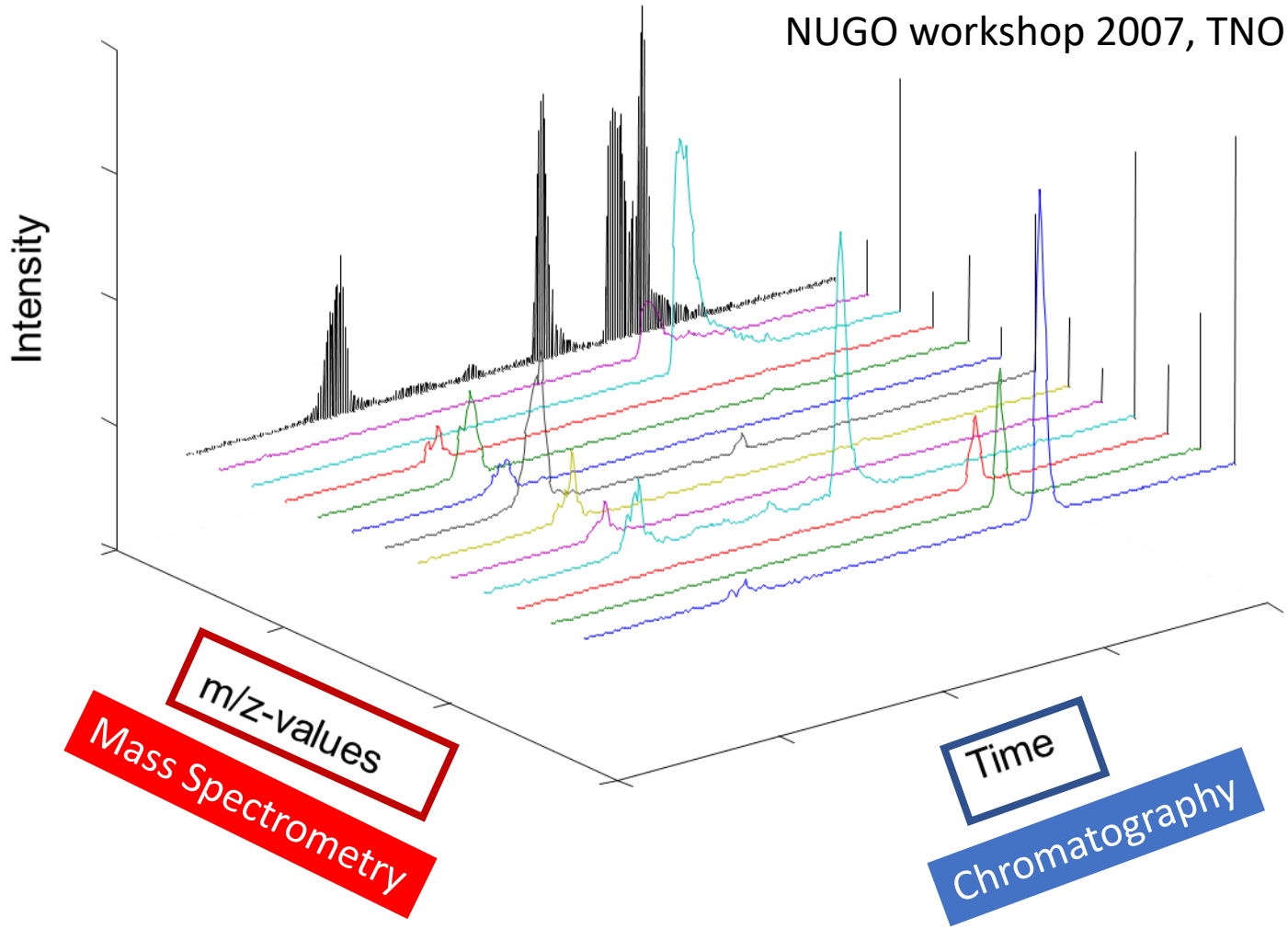
Biological sample



LC-MS separates complex biological samples on:

- Polarity (LC)
- Mass (MS)

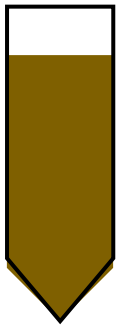
NUGO workshop 2007, TNO



LC-MS based metabolomics

Coverage can be extended by performing LC-MS runs under different experimental conditions

Biological sample

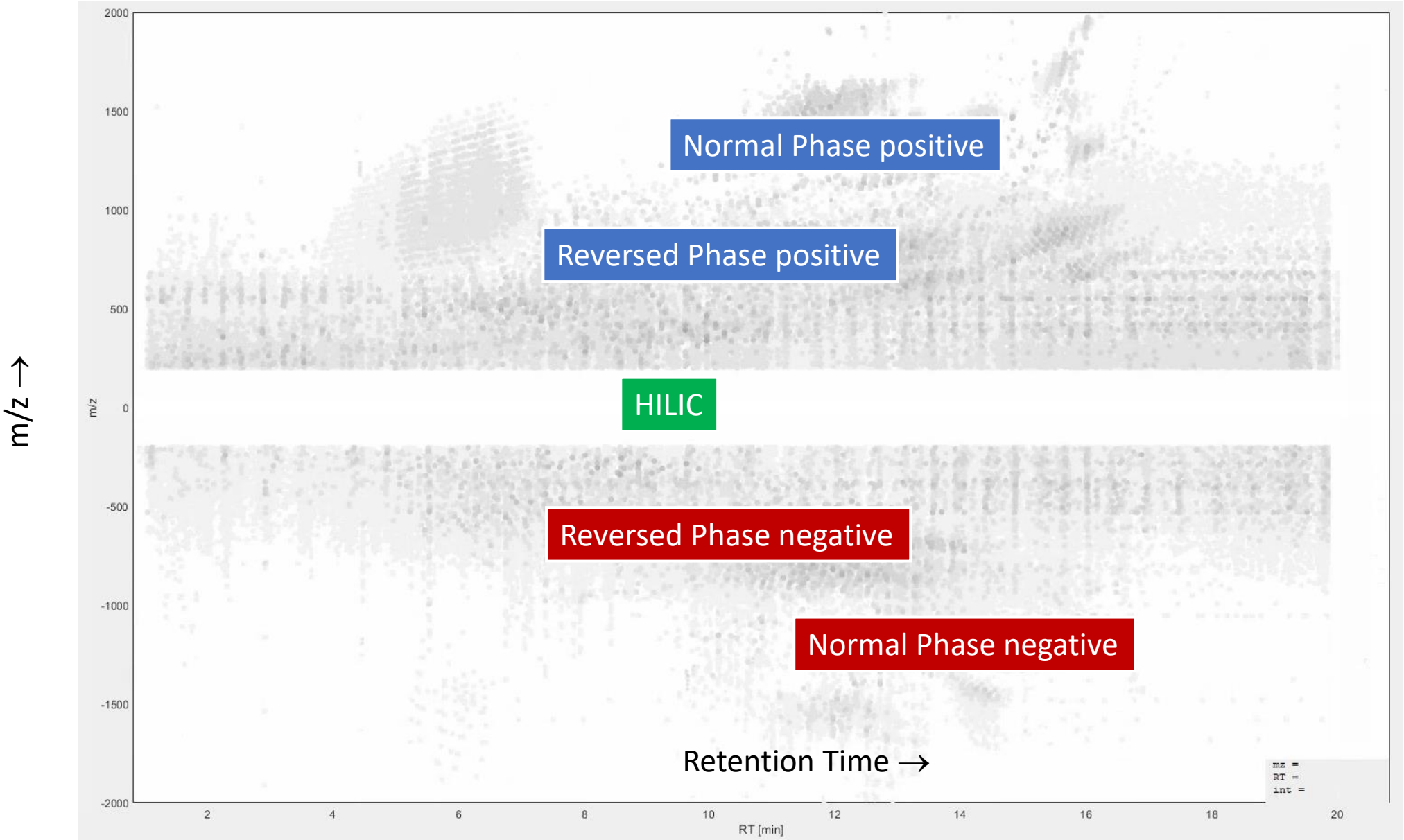


- Lipidomics
  - Normal Phase
    - Negative
    - Positive
  - Reversed Phase
    - Negative
    - Positive
- Polar metabolomics
  - = Small molecule metabolomics
  - HILIC
    - Negative
    - Positive

Sample Preparation | Chromatography | MS Ion mode

Core Facility Metabolomics (CFM)
cfmetabolomics.nl

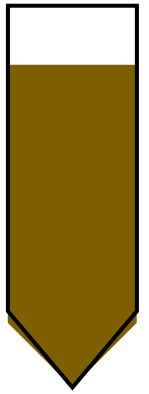
Core Facility Metabolomics



Metabolomics Challenges

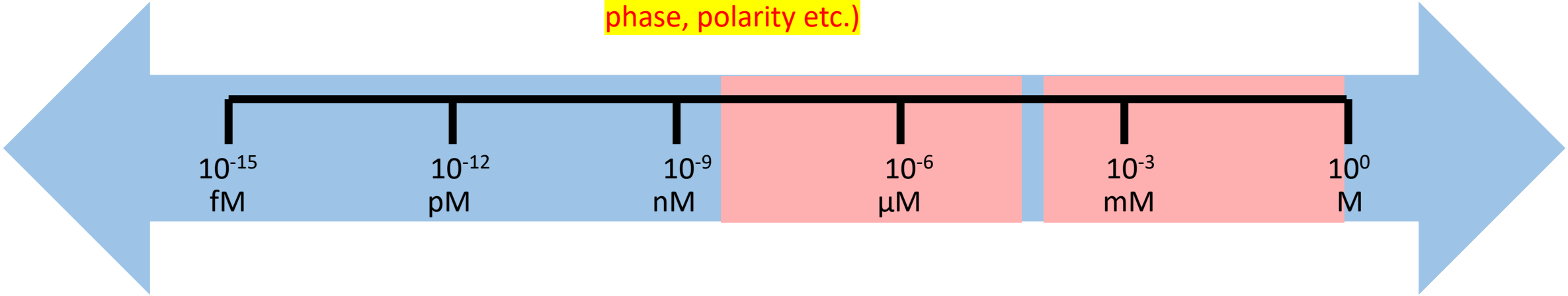
- Sample Complexity
  - Body fluids / tissues
  - Hundreds or thousands metabolites per sample
- Chemical properties
  - Polarity
  - Size / Mass
- Concentrations

Biological sample



LC-MS can separate these complex mixtures by making use of the chemical properties and measures hundreds of metabolites over orders of magnitude in a single run

Coverage can be extended by changing measurement conditions (sample preparation, chromatographic phase, polarity etc.)



LC-MS drawbacks

- Identification is challenging
- Absolute quantification is only possible with good analytical standard materials (isotope labeled)
- Sensitivity is different for each metabolite
- Destructive once a biological sample is measured it can not be measured again
- Ion suppression
- Column degradation