MetaboHUB
Infrastructure Nationale
en Métabolomique

Metabolomics using FTMS

François Fenaille

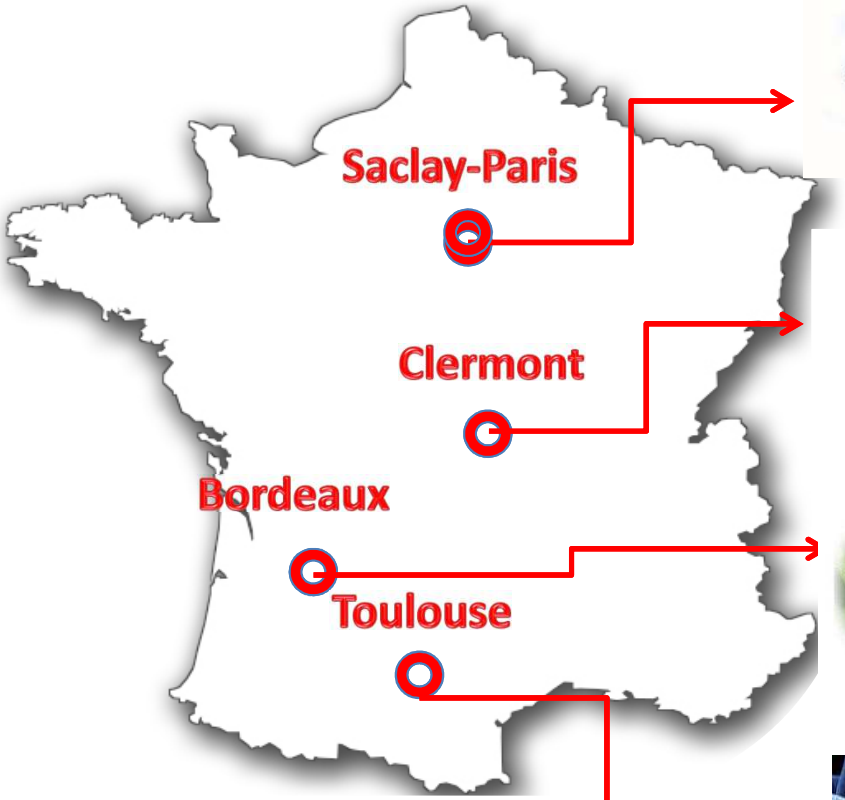
CEA/Laboratoire d'Etude du Métabolisme des Médicaments
CEA-Saclay

francois.fenaille@cea.fr

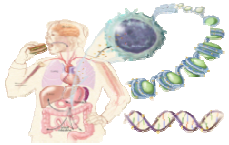
MetaboHUB
Infrastructure Nationale en Métabolomique

MetaboHUB (2013-2020):

French National Infrastructure of Metabolomics and Fluxomics

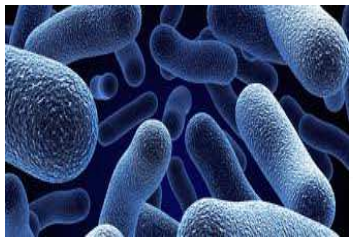


Pharmacology & Clinical diagnostic



Nutrition, Health & Environment

Plant Biology & Biotechnology



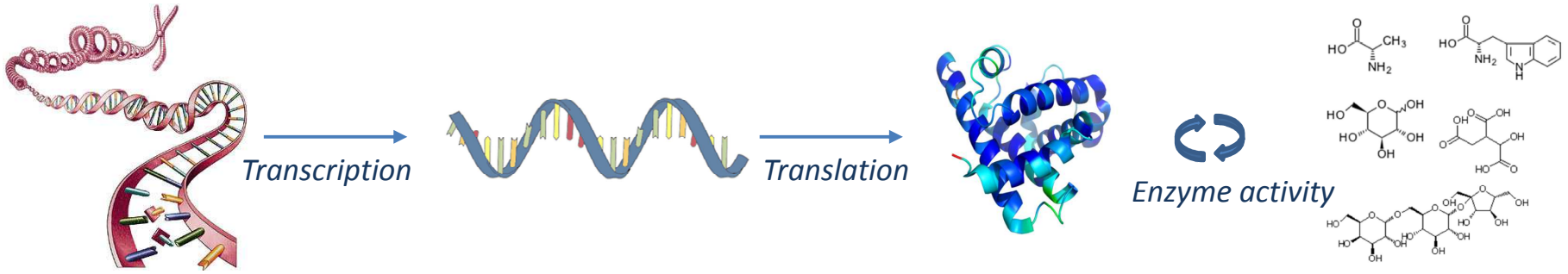
Microbiology, Biotechnology & Toxicology

9 tutelles

https://mama-webapp.metabohub.fr/

MetaboHUB
Infrastructure Nationale en Métabolomique

«Omic»-based approaches



Genes	mRNAs	Proteins	Metabolites
Genome	Transcriptome	Proteome	Metabolome
GENOMICS	**TRANSCRIPTOMICS**	**PROTEOMICS**	**METABOLOMICS**

(Genotype) → (Phenotype)

« What might happen »
« How it is regulated »

« What is happening »

Genotype: part of the genetic makeup of an organism which determines one of its characteristics

Phenotype: observable characteristics or traits of an organism

What is Metabolomics ?

- **The term Metabolomics first used in 1998** (Oliver et al, Trends Biotechnol 1998)

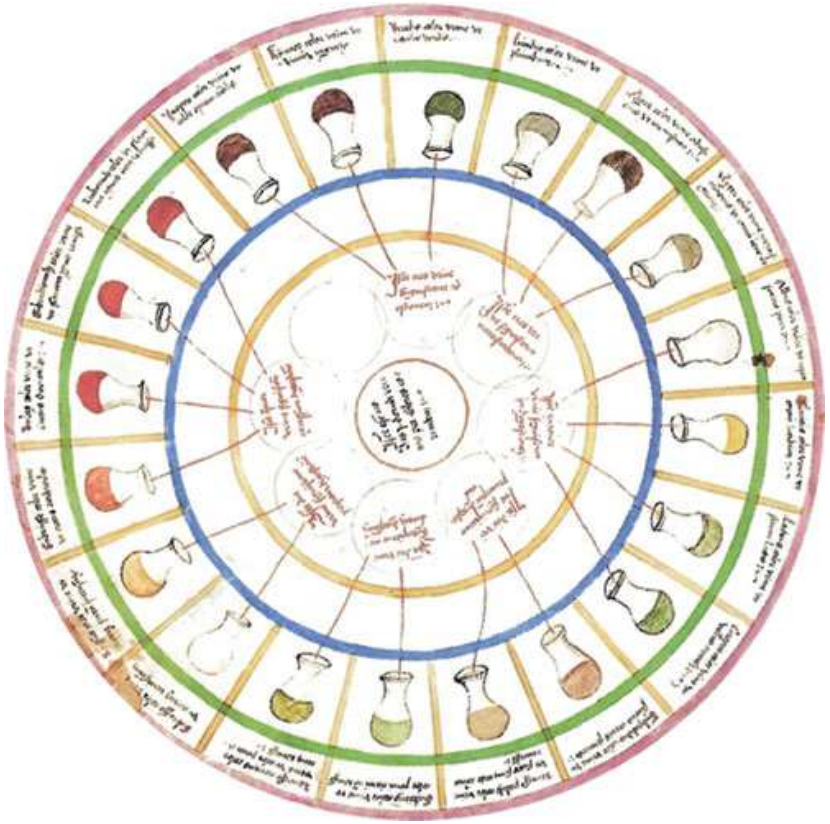
- **Metabolomics is the comprehensive measurement of *all the small molecules* or metabolites* in a given cell, tissue or organism** (i.e., the metabolome)

- **We will consider that Metabolomics is equivalent to Metabonomics**

Metabonomics: "The quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification"
Nicholson et al, Xenobiotica 1999

* no biopolymers (nucleic acids, polypeptides)

Early Metabolomics: What are the origins of the field?



Urine Wheel for diagnosing metabolic diseases
(Ulrich Pinder, 1506, book: *Epiphanie Medicornum*)

- The idea that changes in tissues and biological fluids are indicative of disease dates back to ancient Greece,
- The urine wheel (1506) describes the possible colors, smells and tastes of urine, and uses them to diagnose disease

Nicholson & Lindon, Nature 2008

What is a metabolite?

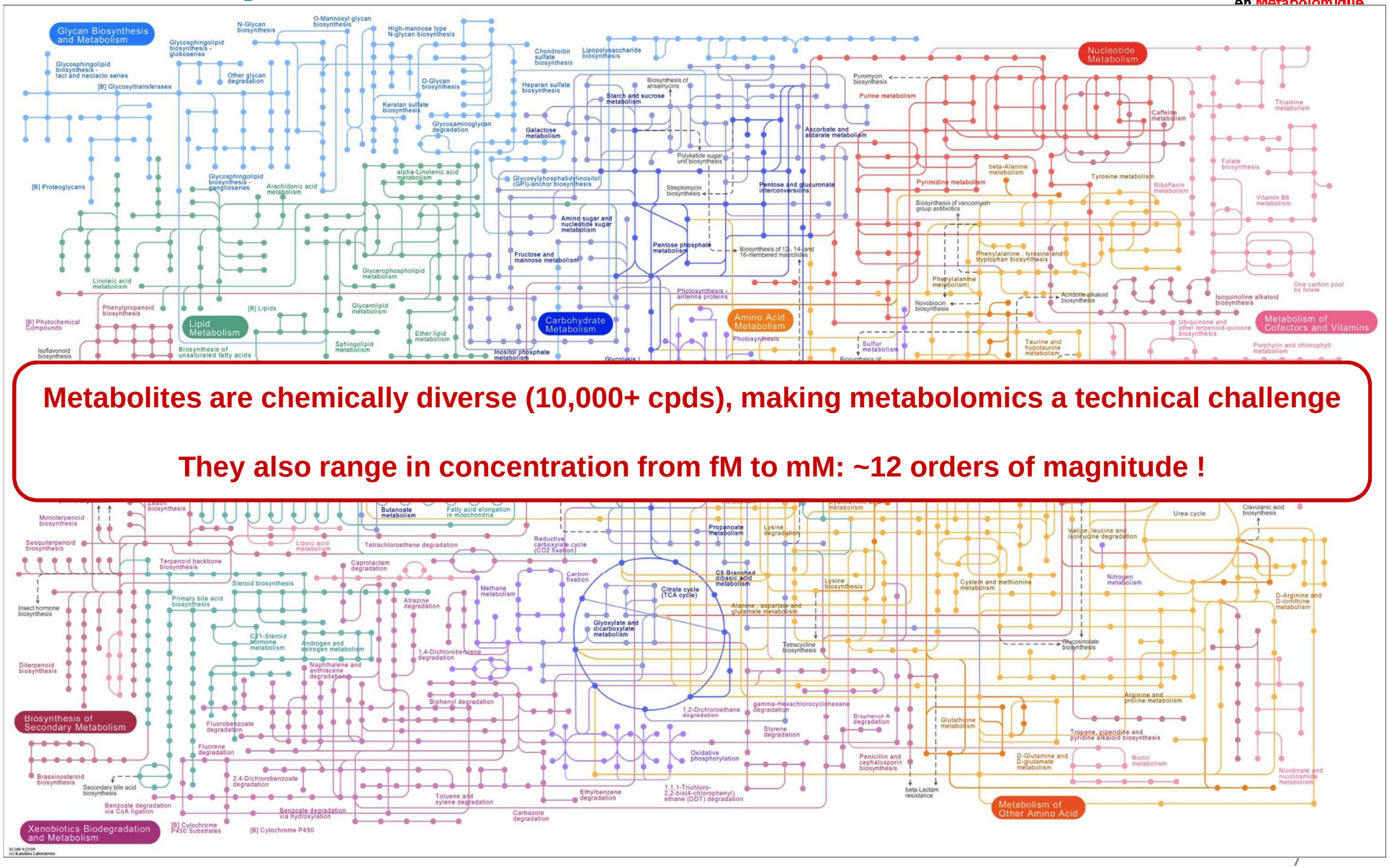
- **Any small organic molecule detectable in the with a molecular weight generally less than 1000 Da (or slightly larger,…)**

- **Includes human and microbial products**

- **High chemical and structural variability: Includes oligonucleotides, sugars, nucleosides, organic acids, amino acids, lipids, steroids, food-derived components, pollutants, drugs and drug metabolites, small peptides, …**

A part of the metabolome…

MetaboHUB
Infrastructure Nationale en Métabolomique



Metabolites are chemically diverse (10,000+ cpds), making metabolomics a technical challenge

They also range in concentration from fM to mM: ~12 orders of magnitude !

Why measuring metabolites?

➢ "Simple" answer

- Studying the metabolome provides system-wide understanding of biological mechanism and pathways
- Infer enzyme activities
- Reflective of any observable phenotype
- Diagnostics, functional genomics

➢ More complex answer

- Not victims, but actors: metabolites have crucial functions (signaling, effects on enzyme activities,…)
- A cause somewhere in the network can have effects elsewhere

The Metabolome is sensitive to

- **Genotype**
- **Changes in mRNA**
- **Changes in proteins**
- **Associated microbes**
- **Environment: food, disease, treatment, exposure to drugs or toxins,…**

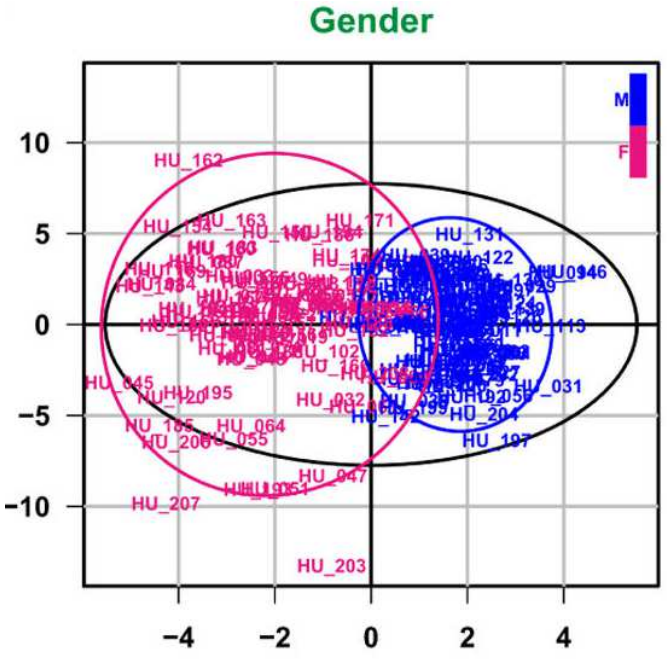
↳ *This sensitivity is both an advantage and disadvantage in Metabolomics*

↳ *Difficult to identify direct and specific associations between cause and effect…*

MetaboHUB
Infrastructure Nationale en Métabolomique

Metabolomics:
Natural and environment variability

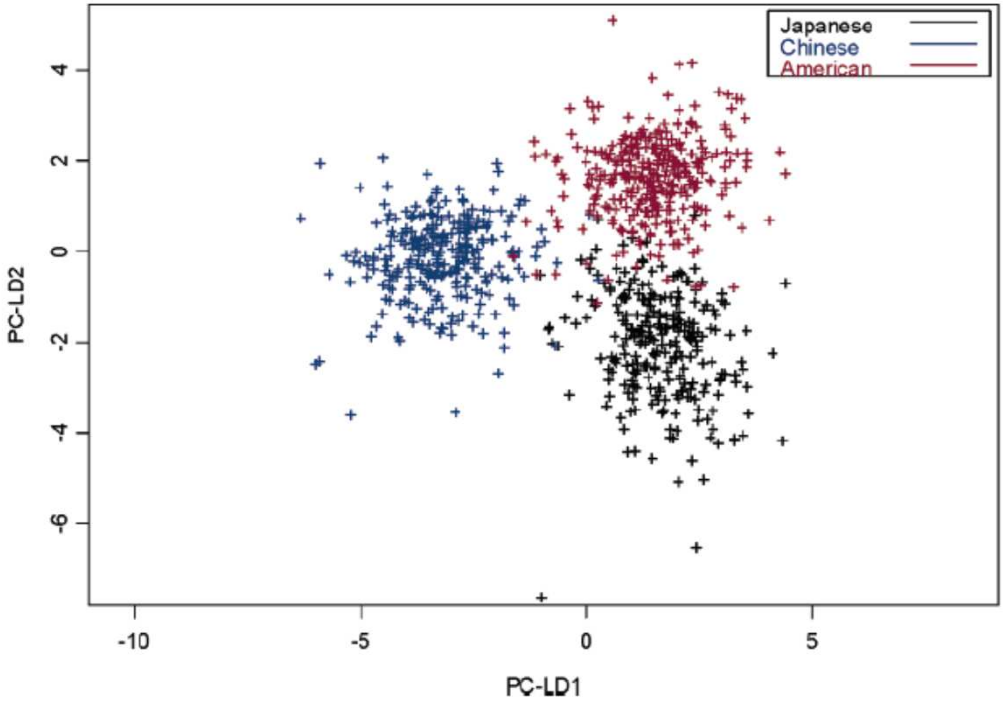
Ex: urine metabolomics



Male/female distinction

LC-MS

Thevenot et al, J Proteome Res 2015



Japanese/Chinese/American distinction

NMR

Dumas ME et al, Anal Chem 2006

MetaboHUB
Infrastructure Nationale
en Métabolomique

Metabolomics and Health

Biomarker Discovery

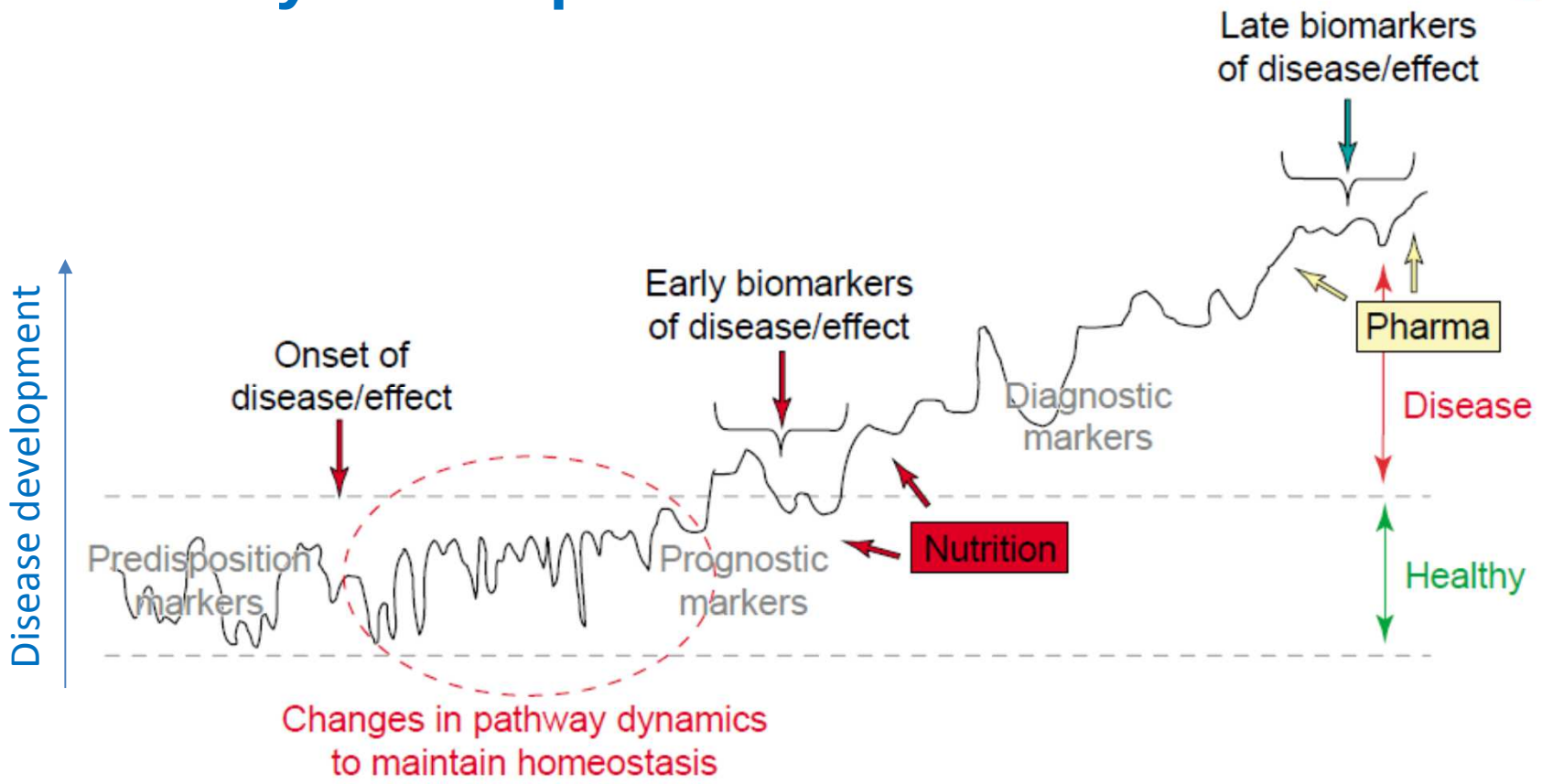
Biomarkers

- *Biomarkers are measurable internal indicators of molecular and/or cellular alterations, that may appear in an organism after or during exposure to a toxicant and possible disease*

- **Biomarker of exposure:** detection of the toxic compound
- **Biomarker of effect:** interaction between the toxicant and a biological target (e.g., DNA adduct)
- **Biomarker of susceptibility:** inter-individual differences in response to toxicants
  - *Metabolic profiling*

MetaboHUB
Infrastructure Nationale en Métabolomique

Why is it important?

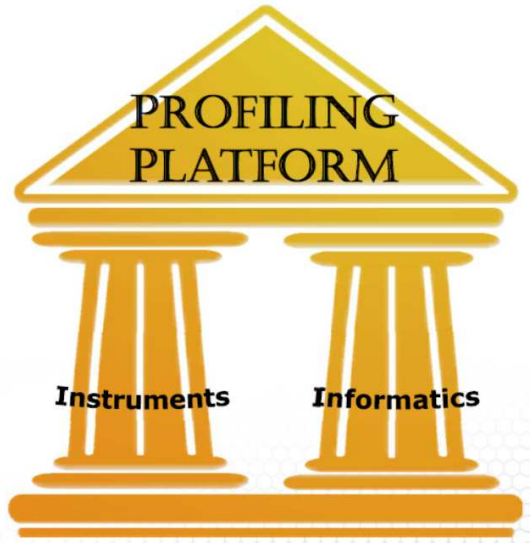


- The system can restore itself via self-regulatory mechanisms, thereby maintaining health
- Disease develops when the system loses this ability
- understanding transitional biomarker profiles in terms of mechanism and validation is crucial

van der Greef et al, Curr Opin Chem Biol 2004

MetaboHUB
Infrastructure Nationale
en Métabolomique

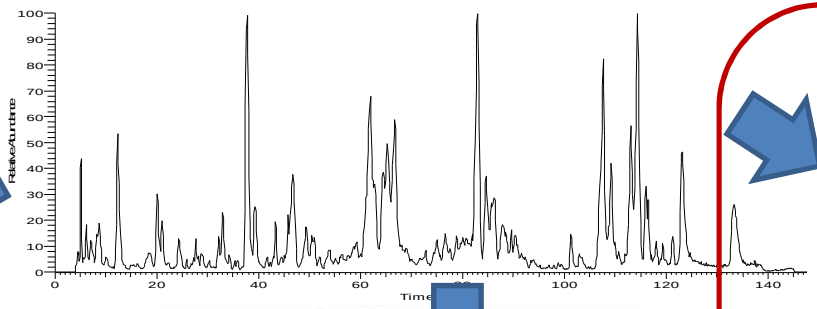
Analytical Tools



How to detect metabolites in biological media?

MetaboHUB
Infrastructure Nationale en Métabolomique

Metabolic fingerprint



NMR

- Simple, non invasive
- Rapid
- Robust: analysis of large series of samples

But:

- Limited sensitivity

GC-EI-MS

- Sensitive
- Reproducible
- Spectral libraries

But:

- Chemical derivatization of non volatile compounds
- Issue of thermolabile compounds

LC-MS

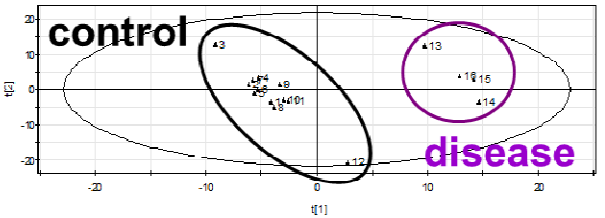
- Molecular mass of intact compounds
- Analysis of thermolabile compounds
- sensitive

But:

Poor inter-platform reproducibility

Metabolomic objectives

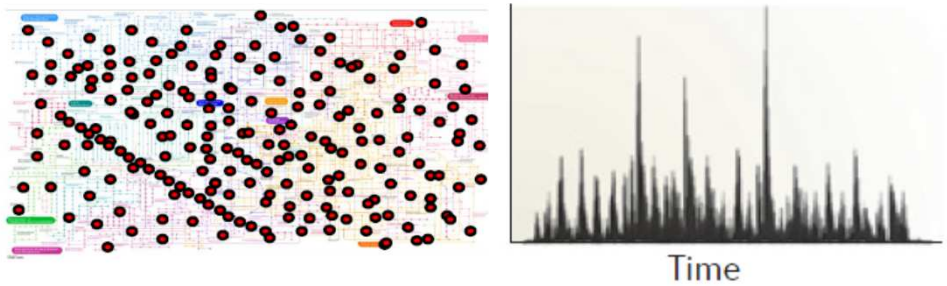
- **Differentiate groups** (e.g., healthy subjects *vs* diseased patients)



- **Quantification:** differences in metabolite concentrations
- **Identification of metabolites that have changed**
- **Systems biology integration:** interactions with genes, proteins

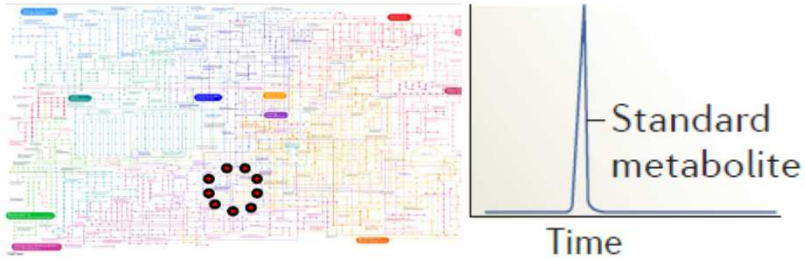
Non-targeted *vs.* targeted metabolomics

Untargeted metabolomics
(Metabolic profiling)



- Semi-quantitative profiling
- LC-HRMS
- Measurement of unexpected changes in known and unknown metabolites (acquisition without any *a priori)*
- Multi- and uni-variate statistical analyses
- > 1,000 metabolites measured
- Identification of (few) relevant metabolites (biomarkers)

Targeted metabolomics
(Quantification)



- Absolute quantification of a limited set of metabolites
- LC-MS/MS
- Metabolites known with corresponding standards available
- Multi- and uni-variate statistical analyses
- Validation of biomarker relevance or hypothesis

Patty GJ et al, Nat Rev Mol Cell Biol 2012; Dunn WB et al, Chem Soc Rev 2011.

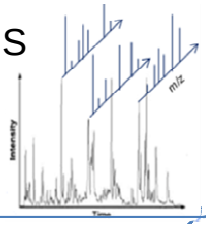
The untargeted metabolomic workflow

Biological question

1 Sample preparation

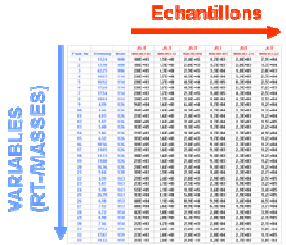
2 Metabolome analysis

LC/MS



3 Data processing

Automatic detection of signals

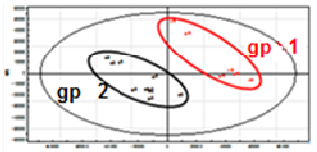


Selection of relevant features

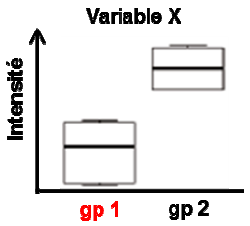
Annotation :
Databases
public / internal

4 Statistical analyses

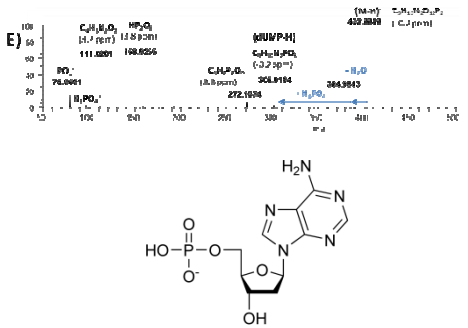
Multivariate



Univariate

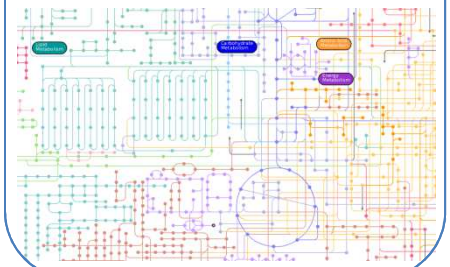


5 Metabolite identification



- To confirm annotations
- To characterize unknowns

6 Data visualization



Main steps of a metabolomic analysis

1. Sample preparation

2. Obtention of metabolic profiles

3. Data treatment and statistical analysis

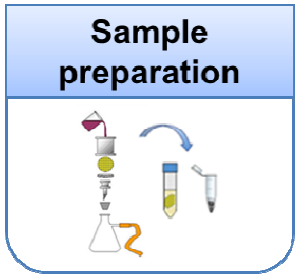
Evidencing biologically-relevant signals

4. Metabolite annotation and identification

MetaboHUB
Infrastructure Nationale
en Métabolomique

Main steps of a metabolomic analysis

1. Sample preparation



2. Obtention of metabolic profiles

3. Data treatment and statistical analysis

Evidencing biologically-relevant signals

4. Metabolite annotation and identification

Sample preparation

➢ Metabolite pre-extraction and extraction

➢ **Quenching** (stopping unwanted biochemical reactions)

Inhibit enzymatic activity

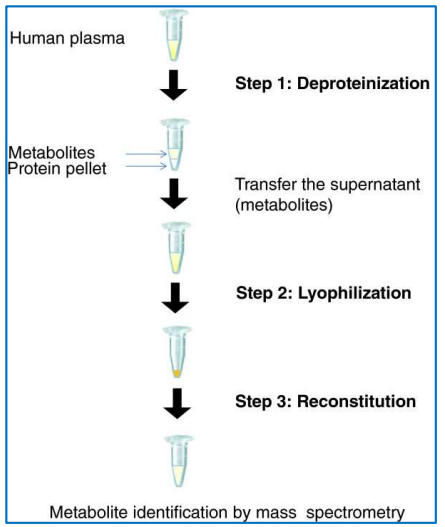
e.g., by sudden temperature shock with ice-cold methanol or liquid nitrogen

➢ **Minimum sample preparation**

Protein precipitation (*e.g.,* plasma deproteinization by methanol)

Lipid extraction

SPE



MetaboHUB
Infrastructure Nationale en Métabolomique

Sample preparation

➢ How many samples for statistically relevant results?

Study Subjects Per Group

	Cell Culture	Small Animals	Human Studies
Optimal	>7	>10	>50
Rigorous	6-7	8-10	40-50
Acceptable	4-5	6-7	25-40

Fewer Required

- Strong phenotype or treatment effect (toxicology study)
- Repeated sampling from the same subject
- Multiple time points
- Multiple doses of a drug/inhibitor

More Required

- Subtle phenotype or treatment effect (dietary supplements, exercise-induced changes
- Mixed populations of subjects (mixed gender, wide-ranging age or BMI)
- Multiple-site collections

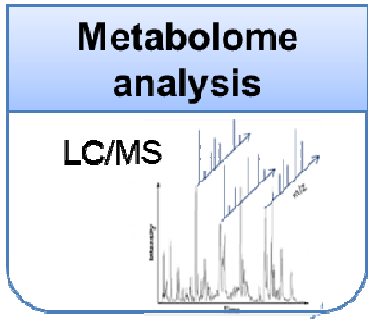
Adapted from Metabolon

MetaboHUB
Infrastructure Nationale en Métabolomique

Main steps of a metabolomic analysis

1. Sample preparation

2. Obtention of metabolic profiles



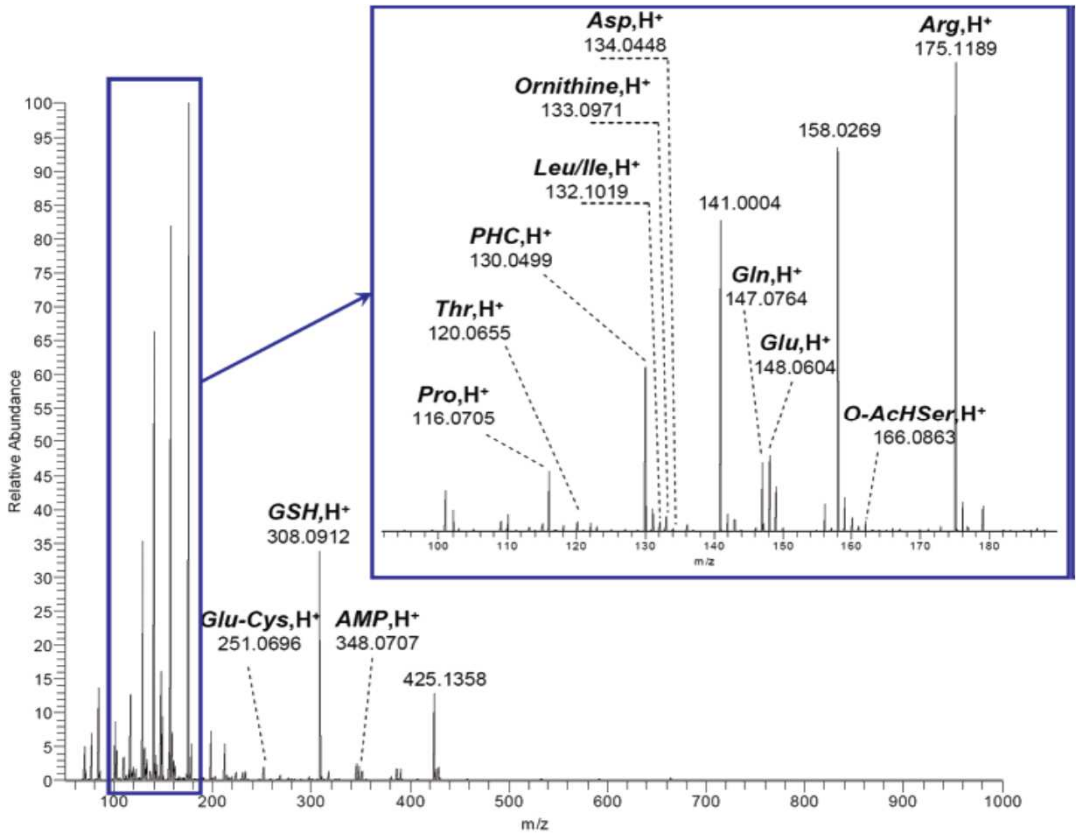
3. Data treatment and statistical analysis

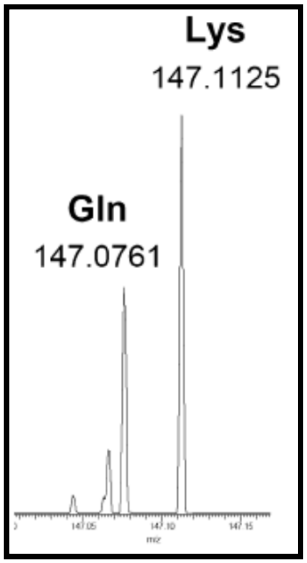
Evidencing biologically-relevant signals

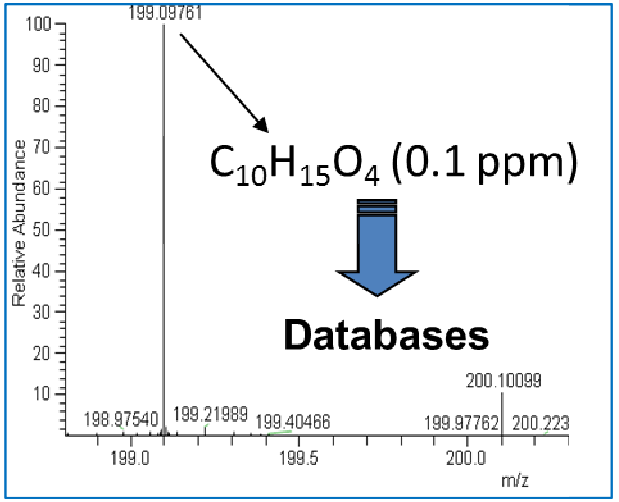
4. Metabolite annotation and identification

MetaboHUB
Infrastructure Nationale en Métabolomique

High resolution mass spectrometry detects more metabolites and improves their identification



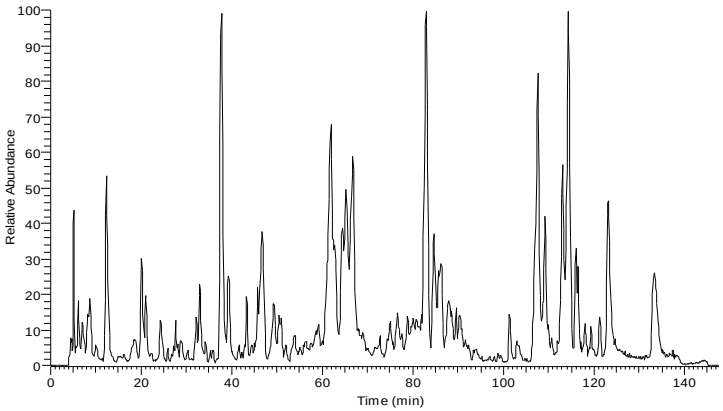




Yeast metabolic extract, LTQ-Orbitrap @100,000 resolution

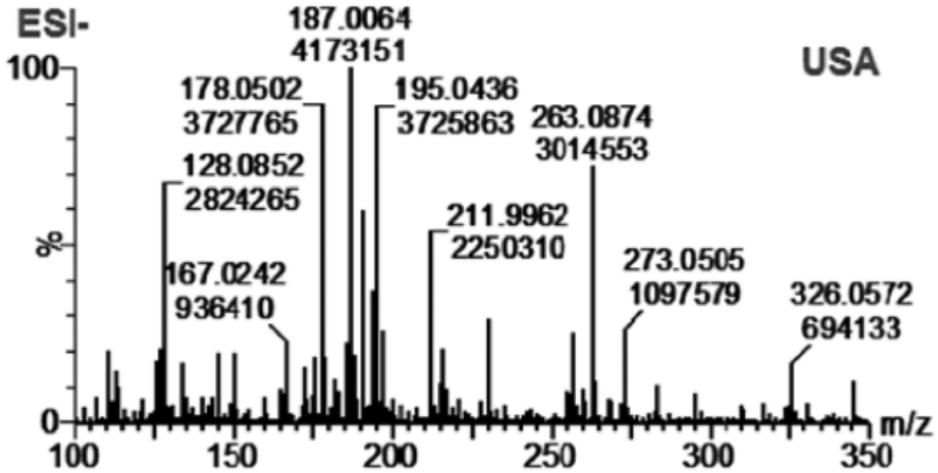
Madalinski G et al, Anal Chem 2008

ESIMS-based metabolomics



LC-MS

- Sensitive
- High metabolome coverage
- Time consuming



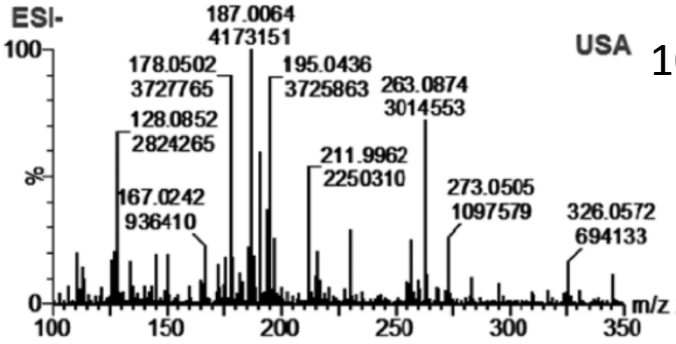
Direct infusion/introduction mass spectrometry (DIMS)

- High throughput
- Lower metabolome coverage

MetaboHUB
Infrastructure Nationale en Métabolomique

High throughput metabolomics by direct-infusion mass spectrometry (DIMS)

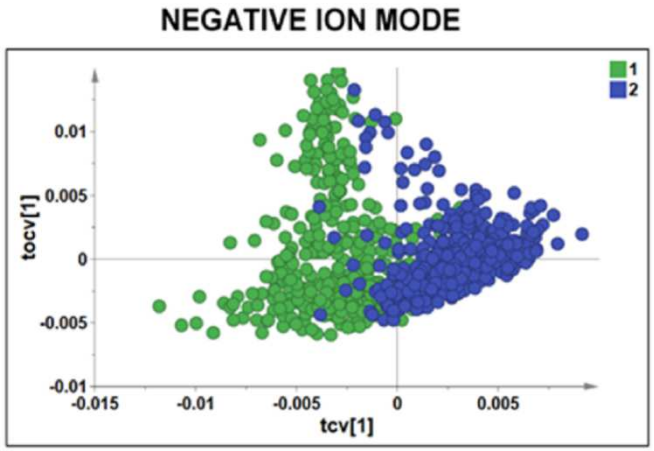
nanoESI/HRMS (Orbitrap, FT-ICR, Q-TOF)



NanoESI/QTOF (-)
Human urine

USA 1000 urine samples

2min/sample
40 metabolites



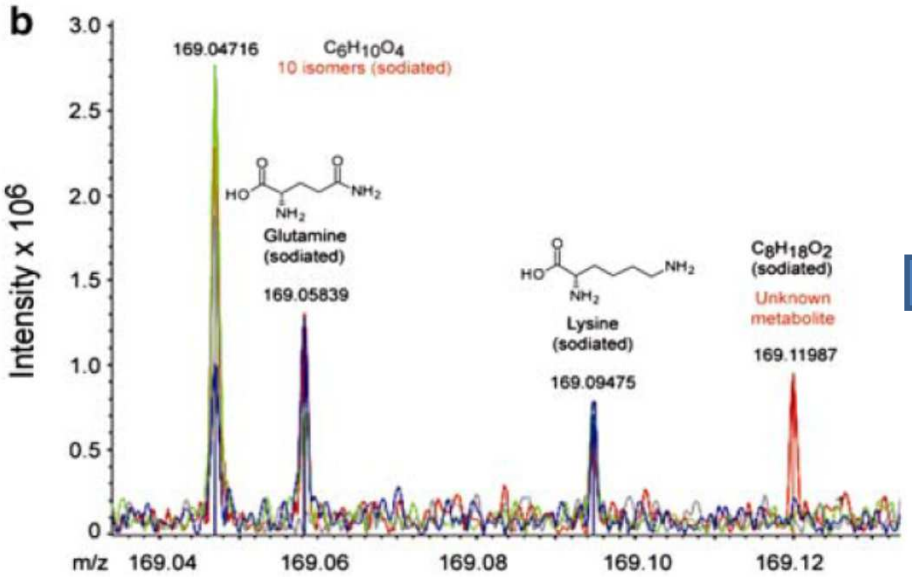
Japan / USA

Chekmeneva et al, J Proteome Res 2017

MetaboHUB
Infrastructure Nationale
en Métabolomique

High throughput metabolomics by direct-infusion mass spectrometry (DIMS)

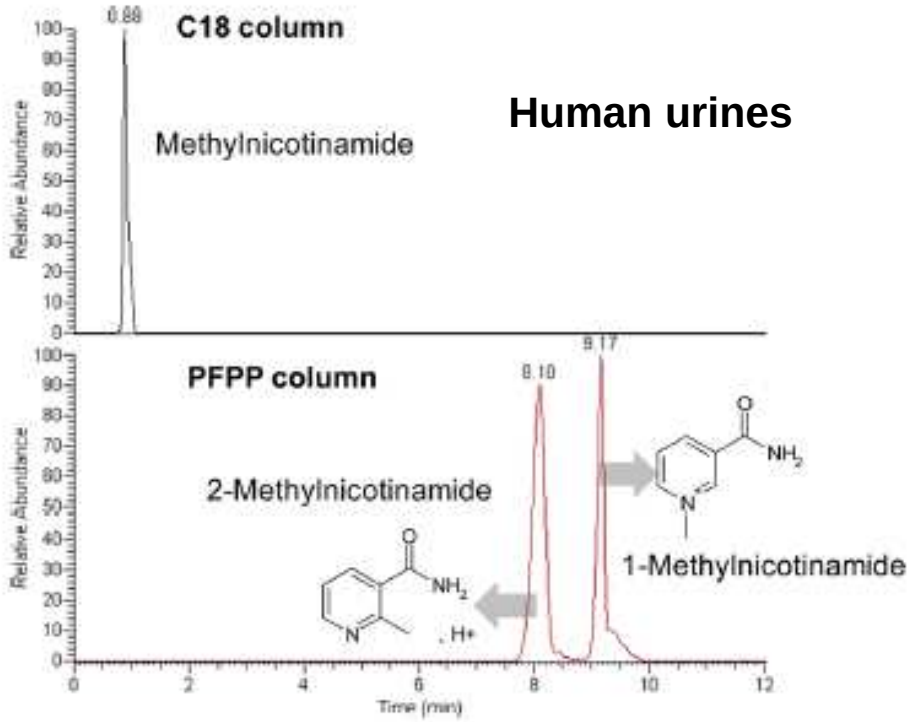
DI/FTICRMS



Annotation of 100 metabolites in human plasma samples

Han et al, Metabolomics 2008

HRMS alone does not distinguish isomers !



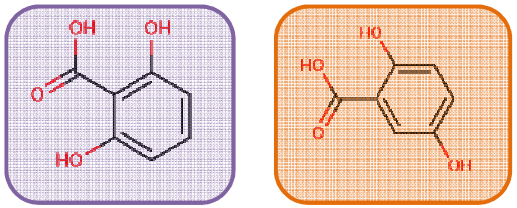
38 metabolites detected in C_{18} conditions actually correspond to 83 métabolites in PFPP conditions.

Roux et al, Anal Chem 2012

MetaboHUB
Infrastructure Nationale en Métabolomique

Value of a multi-LC approach to discriminate between isomers

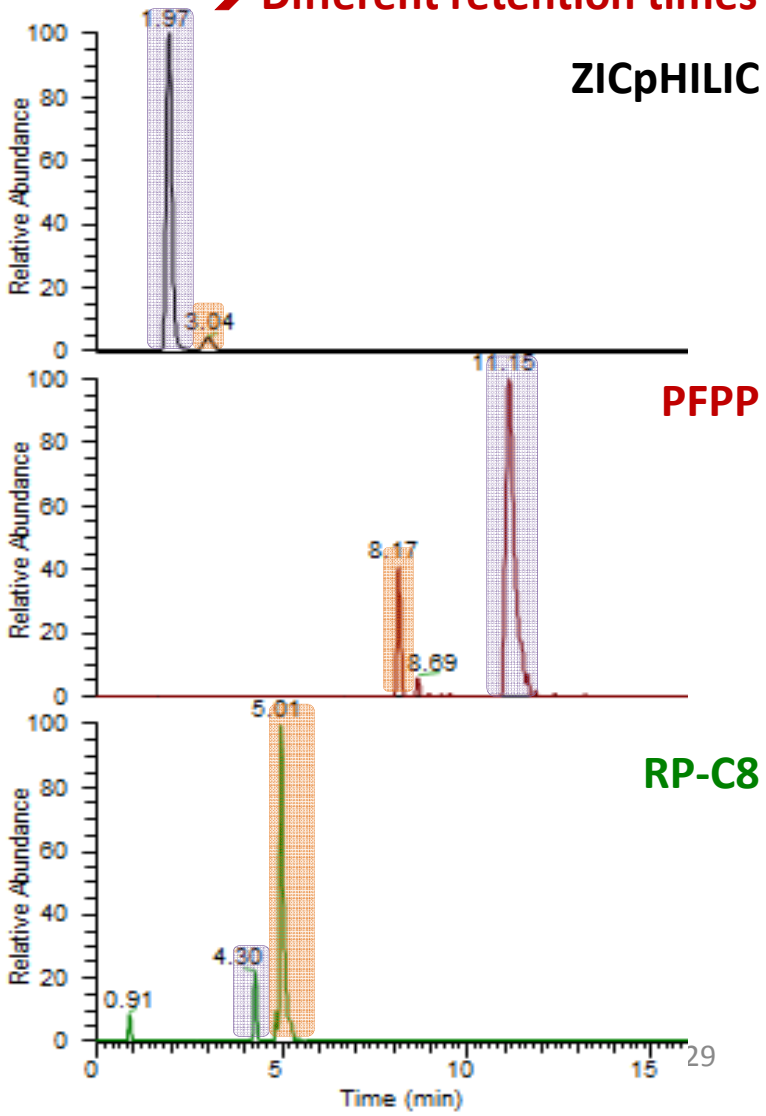
Isomers of dihydroxybenzoic acid



[(M-H)]-
$C_7H_5O_4^-$
153.018

➔ Same MS/MS spectra

➔ Different retention times



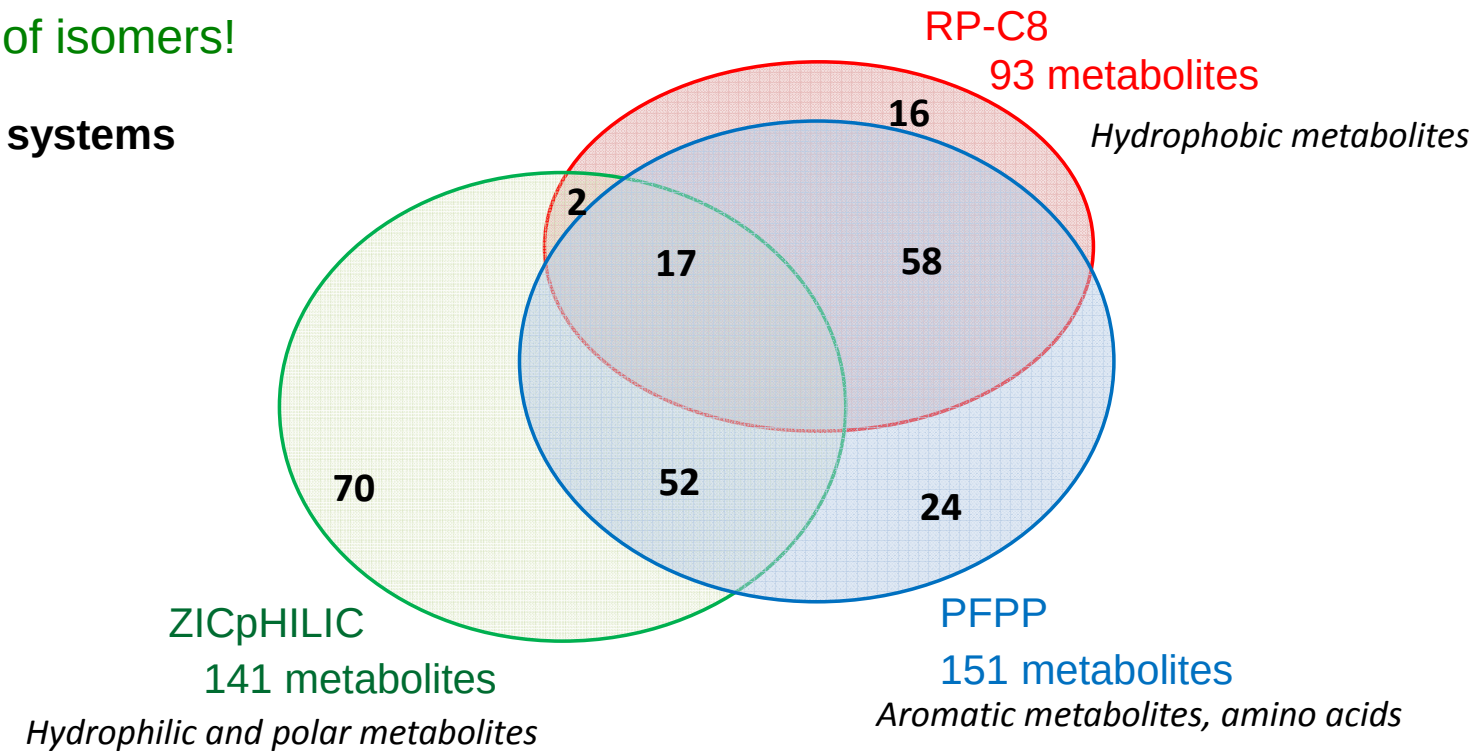
Boudah et al, J Chromatogr B 2014

Several complementary chromatographic methods are required to achieve optimal metabolome coverage

270 metabolites identified in human plasma

Up to 27% of isomers!

3 LC/HRMS systems



Boudah et al, J Chromatogr B 2014

MetaboHUB
Infrastructure Nationale en Métabolomique

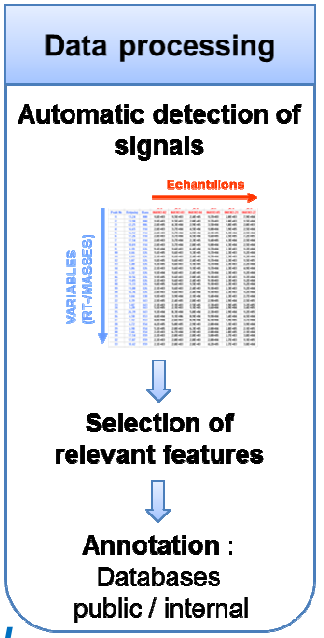
Main steps of a metabolomic analysis

1. Sample preparation

2. Obtention of metabolic profiles

3. Data treatment and statistical analysis

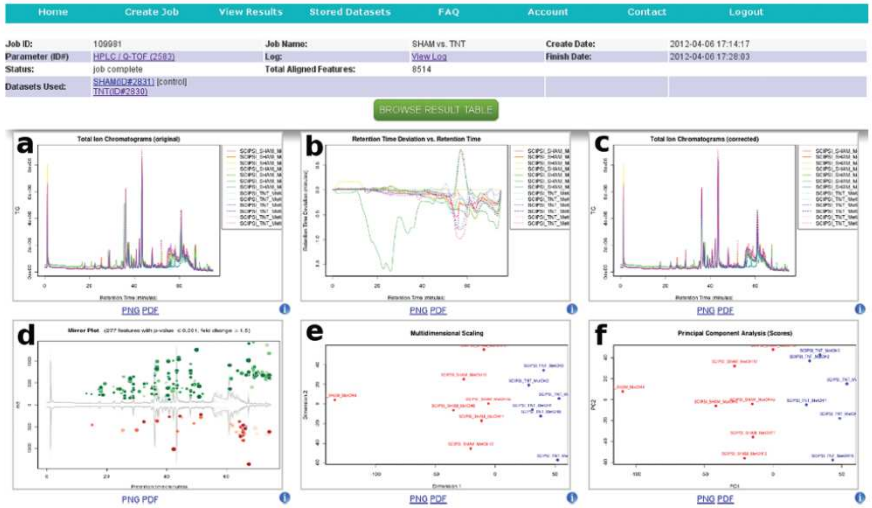
Evidencing biologically-relevant signals



4. Metabolite annotation and identification

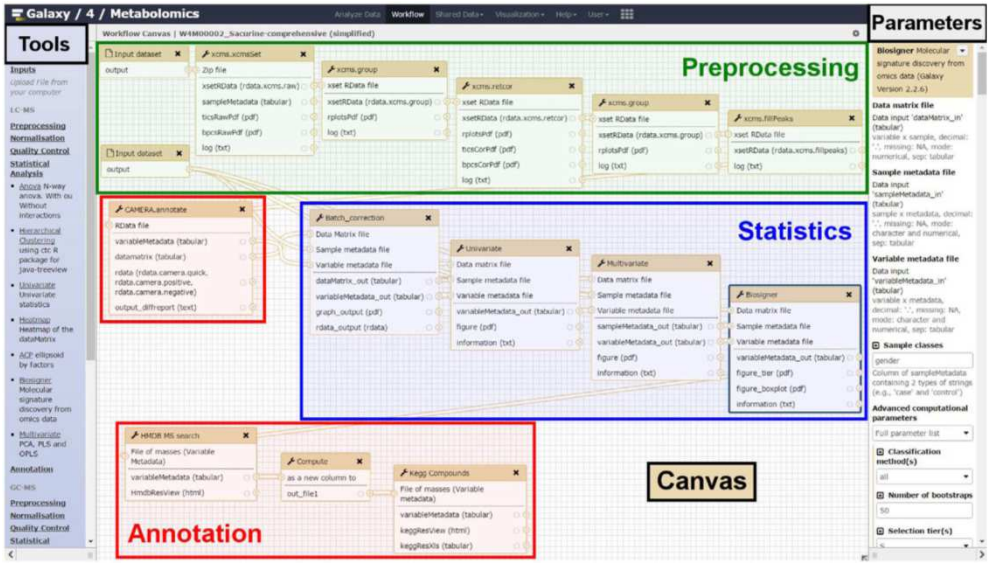
Some online tools

MetaboHUB
Infrastructure Nationale
en Métabolomique



Tautenhan R et al, Anal Chem 2012

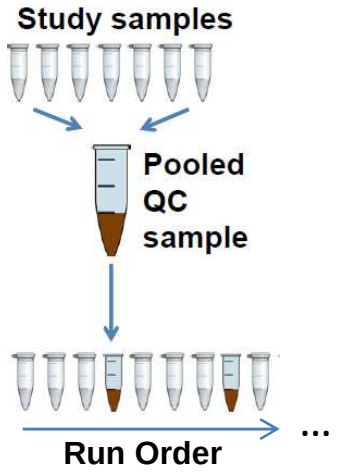
Workflow4metabolomics

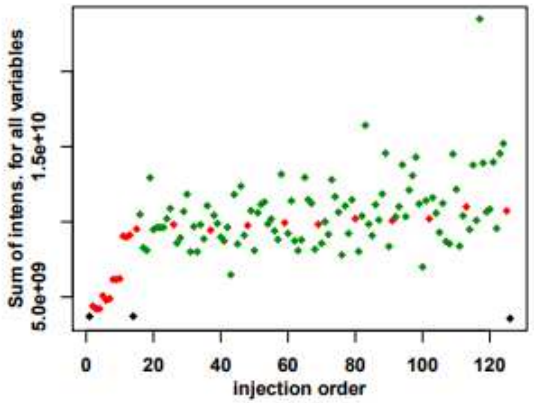


Giacomoni F et al, Bioinformatics 2015
Guitton Y et al, Int J Biochem Cell Biol 2017

How to build a sample batch to "avoid" experimental biases?

- **Add internal standards to all samples**
- **Sample randomization is mandatory**
- **~200 samples/batch**
- **Include blank samples**
- **Include Quality Control (QC) samples:**
  must be representative biological samples
  (e.g., pool of study samples)
- **If needed include interbatches QCs**





MetaboHUB
Infrastructure Nationale en Métabolomique

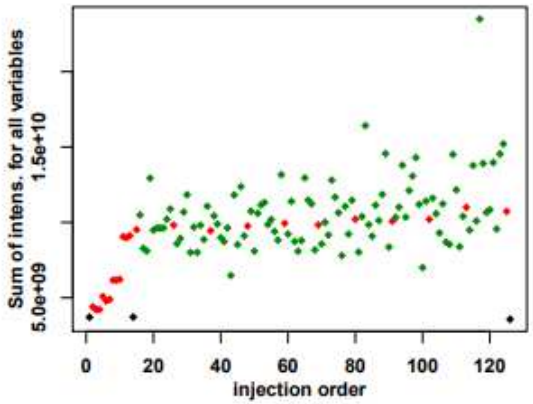
What does a typical sequence look like?

Injection Order	Sample
1	Blank
2	Blank
3	QC
4	QC
5	QC
6	QC
7	QC
8	Blank
9	8x dil. QC
10	4x dil. QC
11	2x dil. QC
12	QC
13	Blank
14	QC
15	Sample 1
16	Sample 2
17	Sample 3
…	…
24	Sample 10
25	Blank
26	QC
27	Sample 11
…	…
36	Sample 20
37	Blank
38	QC
…	…

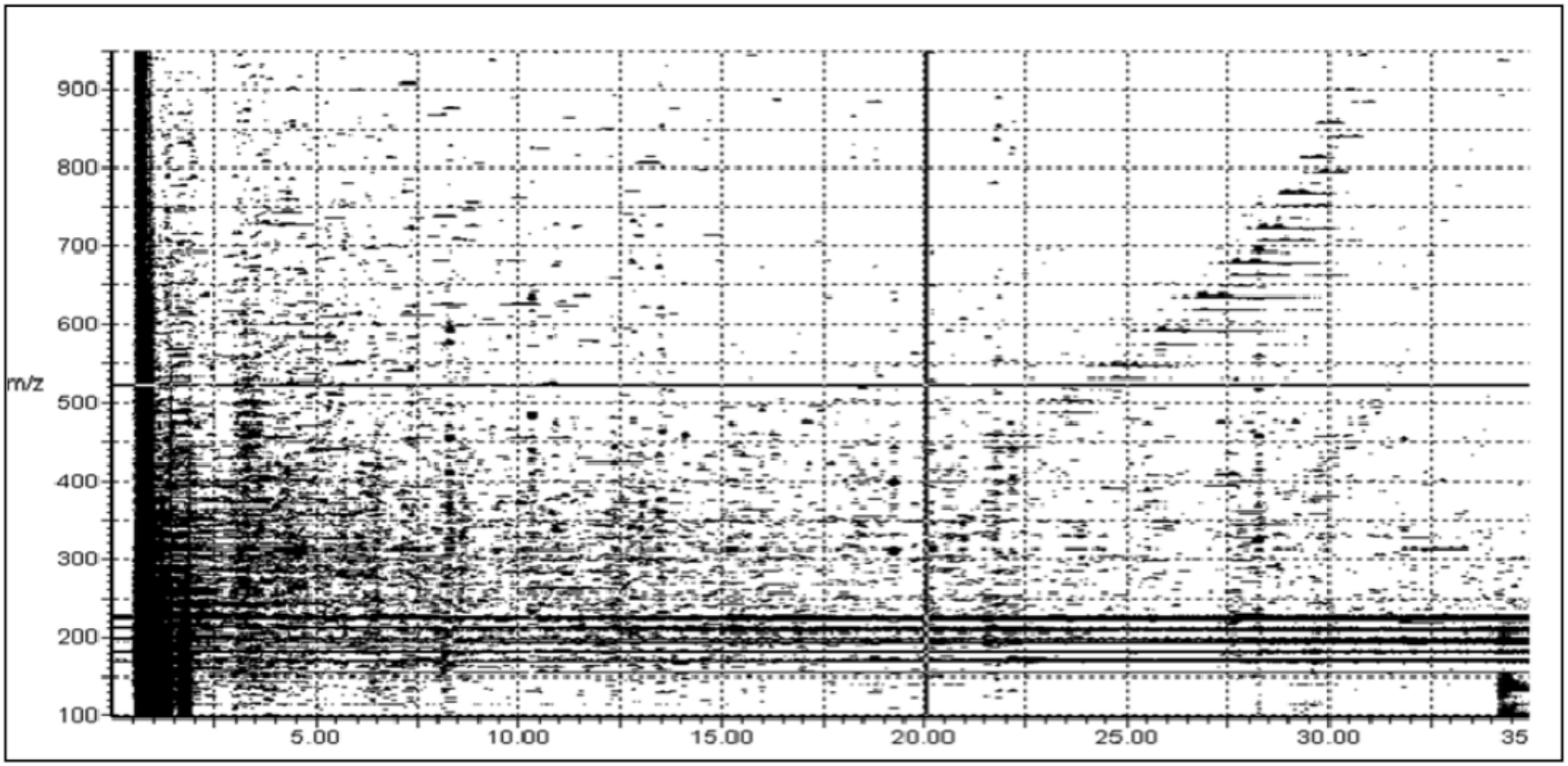
QCs for equilibration

Diluted QCs for data treatment (n=3 each)

10 biological samples

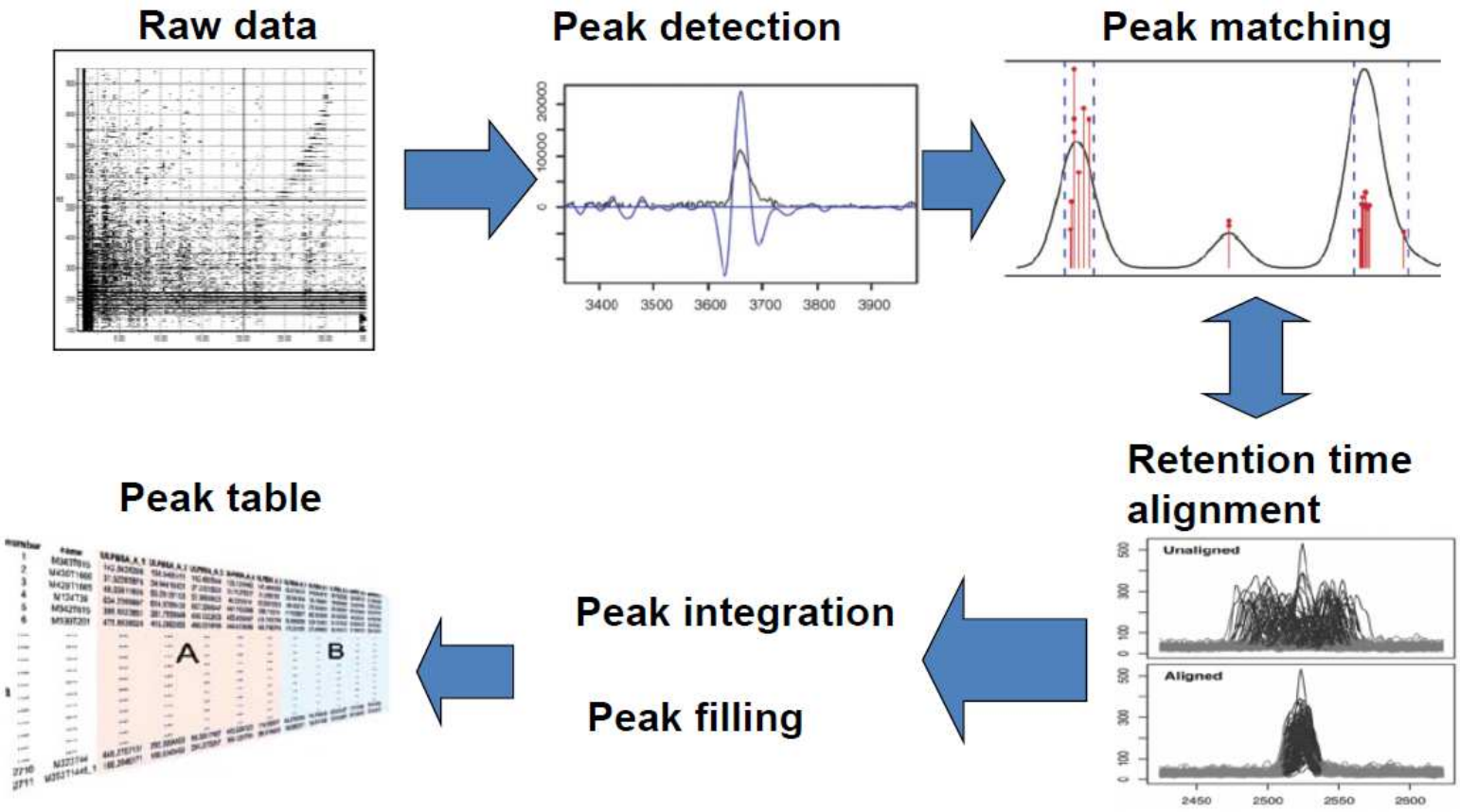


LC-MS metabolic profiles



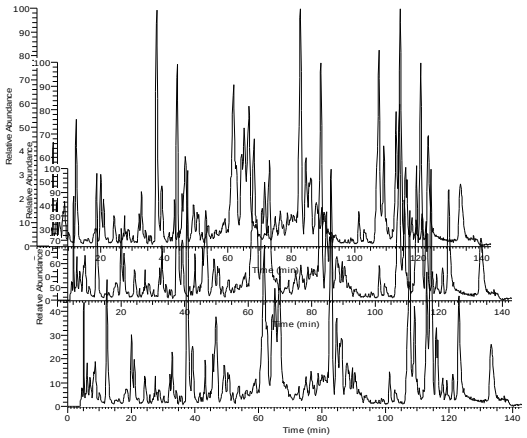
10,000+ signals, 100-1000s metabolites ?

LC-MS data preprocessing (XCMS)



Smith CA et al, Anal Chem 2006

Obtention of peak lists



Samples

Variables (Rt-mass)

Peak Nr	Ret(min)	Mass	J5-T 060303-02	J5-T 060303-03	J5-T 060303-04	J5-T 060303-05	J5-T 060303-21	J5-T 060303-22	
1	13.24	100	1.8E+03	1.5E+03	2.4E+05	1.7E+03	2.0E+03	7.9E+04	?
2	13.98	100	1.8E+03	1.5E+03	2.0E+03	1.7E+03	2.0E+03	1.5E+04	?
3	42.25	106	2.0E+05	4.3E+04	2.9E+05	3.5E+04	1.8E+05	2.0E+03	?
4	16.65	114	2.0E+03	3.7E+04	4.5E+04	1.0E+04	3.9E+05	2.5E+04	?
5	16.92	114	2.0E+03	3.7E+04	4.5E+04	2.1E+05	8.3E+04	2.5E+04	?
6	17.26	114	2.0E+03	3.7E+04	4.5E+04	1.6E+05	2.5E+05	7.2E+05	?
7	17.54	114	2.0E+03	3.7E+04	2.3E+05	1.4E+05	4.3E+04	2.5E+04	?
8	18.01	114	2.0E+03	3.7E+04	2.8E+05	1.0E+04	4.3E+04	2.5E+04	?
9	4.19	126	9.4E+04	1.6E+03	6.4E+04	1.7E+04	2.3E+03	1.2E+04	?
10	4.66	126	1.4E+05	1.6E+03	1.3E+05	1.7E+04	2.3E+03	1.2E+04	?
11	4.93	126	2.1E+03	1.6E+03	2.4E+03	1.7E+04	2.3E+03	1.2E+04	?
12	5.07	126	1.8E+05	1.6E+03	2.4E+03	1.7E+04	2.3E+03	1.5E+05	?
13	5.40	126	1.3E+05	1.6E+03	1.1E+05	1.7E+04	2.3E+03	2.2E+04	?
14	5.86	126	2.1E+03	1.6E+03	1.1E+05	1.7E+04	2.3E+03	4.9E+04	?
15	6.32	126	1.5E+04	1.6E+03	2.4E+03	1.7E+04	2.3E+03	1.2E+04	?
16	10.56	126	1.9E+05	1.6E+03	2.0E+05	9.1E+03	2.3E+03	1.2E+04	?
17	11.05	126	2.1E+03	1.6E+03	2.4E+03	9.1E+03	2.3E+03	1.2E+04	?
18	11.33	126	1.0E+05	1.6E+03	1.5E+05	9.1E+03	2.3E+03	1.2E+04	?
19	11.80	126	2.1E+03	1.6E+03	2.4E+03	9.1E+03	2.3E+03	1.2E+04	?
20	16.36	126	2.0E+03	1.6E+03	2.4E+04	9.1E+03	2.3E+03	1.2E+04	?
21	9.04	138	1.9E+03	3.9E+04	2.1E+04	1.4E+04	2.3E+03	2.7E+04	?
22	4.39	143	2.9E+05	2.4E+05	2.8E+03	2.9E+05	2.9E+04	2.5E+05	?
23	5.07	143	2.3E+03	2.1E+03	2.5E+05	2.3E+03	1.9E+05	3.2E+05	
24	5.20	143	5.5E+05	2.1E+03	1.1E+05	3.6E+04	2.9E+04	3.4E+05	
25	26.39	143	1.3E+04	8.3E+04	5.8E+04	2.3E+03	2.9E+04	1.2E+05	
26	6.58	153	4.0E+04	1.1E+06	8.9E+04	9.9E+04	3.4E+04	4.5E+04	
27	7.12	153	4.0E+04	2.6E+03	8.9E+04	6.1E+04	1.9E+05	3.7E+04	
28	6.72	154	4.2E+05	5.0E+05	2.9E+03	2.0E+04	3.1E+03	3.9E+04	
29	6.98	154	5.3E+05	2.0E+03	6.3E+05	2.0E+04	2.8E+05	2.1E+05	
30	7.66	154	2.2E+03	6.7E+04	2.9E+03	2.0E+04	2.8E+05	2.1E+04	
31	17.54	159	2.3E+03	2.0E+03	2.8E+03	3.0E+05	2.7E+03	3.8E+04	
32	17.87	159	2.3E+03	2.0E+03	2.8E+03	2.8E+04	2.7E+03	1.1E+05	
33	18.42	159	2.3E+03	2.0E+03	2.8E+03	4.2E+05	2.7E+03	3.8E+04	

Few thousands of variables…
(per chromatographic condition!)

Obtention of peak lists: Quality control and data filtering

MetaboHUB Infrastructure Nationale en Métabolomique

Samples

Peak Nr	Ret(min)	Mass	J5-T 060303-02	J5-T 060303-03	J5-T 060303-04	J5-T 060303-05	J5-T 060303-21	J5-T 060303-22
1	13.24	100	1.8E+03	1.5E+03	2.4E+05	1.7E+03	2.0E+03	7.9E+04
2	13.98	100	1.8E+03	1.5E+03	2.0E+03	1.7E+03	2.0E+03	1.5E+04
3	42.25	106	2.0E+05	4.3E+04	2.9E+05	3.5E+04	1.8E+05	2.0E+03
4	16.65	114	2.0E+03	3.7E+04	4.5E+04	1.0E+04	3.9E+05	2.5E+04
5	16.92	114	2.0E+03	3.7E+04	4.5E+04	2.1E+05	8.3E+04	2.5E+04
6	17.26	114	2.0E+03	3.7E+04	4.5E+04	1.6E+05	2.5E+05	7.2E+05
7	17.54	114	2.0E+03	3.7E+04	2.3E+05	1.4E+05	4.3E+04	2.5E+04
8	18.01	114	2.0E+03	3.7E+04	2.8E+05	1.0E+04	4.3E+04	2.5E+04
9	4.19	126	9.4E+04	1.6E+03	6.4E+04	1.7E+04	2.3E+03	1.2E+04
10	4.66	126	1.4E+05	1.6E+03	1.3E+05	1.7E+04	2.3E+03	1.2E+04
11	4.93	126	2.1E+03	1.6E+03	2.4E+03	1.7E+04	2.3E+03	1.2E+04
12	5.07	126	1.8E+05	1.6E+03	2.4E+03	1.7E+04	2.3E+03	1.5E+05
13	5.40	126	1.3E+05	1.6E+03	1.1E+05	1.7E+04	2.3E+03	2.2E+04
14	5.86	126	2.1E+03	1.6E+03	1.1E+05	1.7E+04	2.3E+03	4.9E+04

1. Repeatability filter

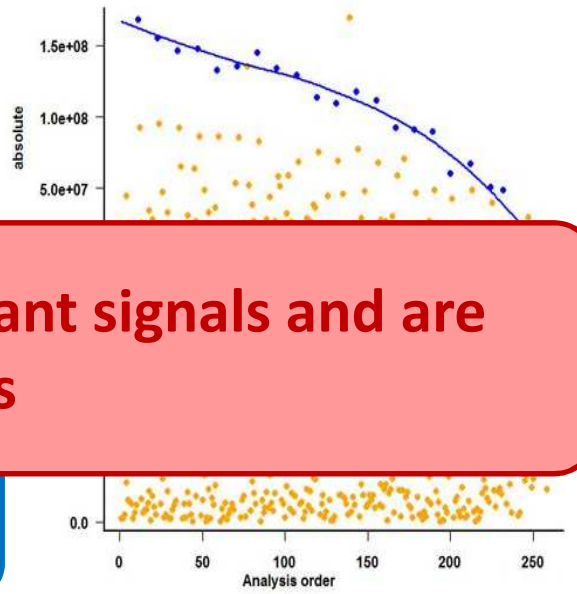
- CV of Internal Standards < 30%
- CV of QC Samples < 30%

2. Intensity and linearity filters

Peak lists now include only analytically relevant signals and are ready for statistical analysis

Few thousands of variables…
(per chromatographic condition)

- Correct feature specific drift within a batch



LOESS Algorithm
(Low Order non linear locally Estimated Smoothing Function)

4. Batch correction

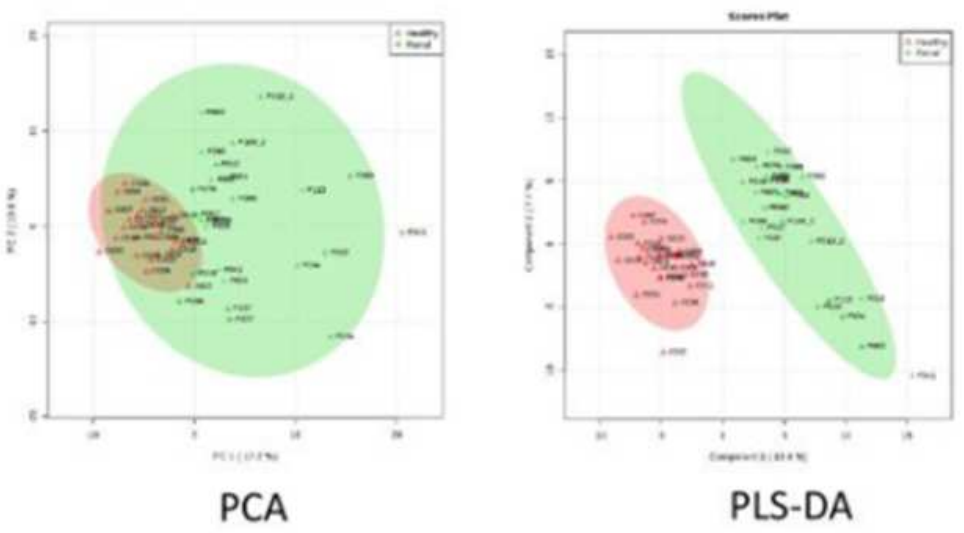
- Correct drift across batches

Cleveland, J. Am. Stat. Assoc. 1979
Dunn WB, Nat. Protoc. 2011

MetaboHUB
Infrastructure Nationale en Métabolomique

Facilitate data comparison by lowering the number of dimensions

			J5-T	J5-T	J5-T	J5-T	J5-T	J5-T
Peak Nr	Ret(min)	Mass	060303-02	060303-03	060303-04	060303-05	060303-21	060303-22
1	13.24	100	1.8E+03	1.5E+03	2.4E+05	1.7E+03	2.0E+03	7.9E+04
2	13.98	100	1.8E+03	1.5E+03	2.0E+03	1.7E+03	2.0E+03	1.5E+04
3	42.25	106	2.0E+05	4.3E+04	2.9E+05	3.5E+04	1.8E+05	2.0E+03
4	16.65	114	2.0E+03	3.7E+04	4.5E+04	1.0E+04	3.9E+05	2.5E+04
5	16.92	114	2.0E+03	3.7E+04	4.5E+04	2.1E+05	8.3E+04	2.5E+04
6	17.26	114	2.0E+03	3.7E+04	4.5E+04	1.6E+05	2.5E+05	7.2E+05
7	17.54	114	2.0E+03	3.7E+04	2.3E+05	1.4E+05	4.3E+04	2.5E+04
8	18.01	114	2.0E+03	3.7E+04	2.8E+05	1.0E+04	4.3E+04	2.5E+04
9	4.19	126	9.4E+04	1.6E+03	6.4E+04	1.7E+04	2.3E+03	1.2E+04
10	4.66	126	1.4E+05	1.6E+03	1.3E+05	1.7E+04	2.3E+03	1.2E+04
11	4.93	126	2.1E+03	1.6E+03	2.4E+03	1.7E+04	2.3E+03	1.2E+04
12	5.07	126	1.8E+05	1.6E+03	2.4E+03	1.7E+04	2.3E+03	1.5E+05
13	5.40	126	1.3E+05	1.6E+03	1.1E+05	1.7E+04	2.3E+03	2.2E+04
14	5.86	126	2.1E+03	1.6E+03	1.1E+05	1.7E+04	2.3E+03	4.9E+04
15	6.32	126	1.5E+04	1.6E+03	2.4E+03	1.7E+04	2.3E+03	1.2E+04
16	10.56	126	1.9E+05	1.6E+03	2.0E+05	9.1E+03	2.3E+03	1.2E+04
17	11.05	126	2.1E+03	1.6E+03	2.4E+03	9.1E+03	2.3E+03	1.2E+04
18	11.33	126	1.0E+05	1.6E+03	1.5E+05	9.1E+03	2.3E+03	1.2E+04
19	11.80	126	2.1E+03	1.6E+03	2.4E+03	9.1E+03	2.3E+03	1.2E+04
20	16.36	126	2.0E+03	1.6E+03	2.4E+04	9.1E+03	2.3E+03	1.2E+04
21	9.04	138	1.9E+03	3.9E+04	2.1E+04	1.4E+04	2.3E+03	2.7E+04
22	4.39	143	2.9E+05	2.4E+05	2.8E+03	2.9E+05	2.9E+04	2.5E+05
23	5.07	143	2.3E+03	2.1E+03	2.5E+05	2.3E+03	1.9E+05	3.2E+05
24	5.20	143	5.5E+05	2.1E+03	1.1E+05	3.6E+04	2.9E+04	3.4E+05
25	26.39	143	1.3E+04	8.3E+04	5.8E+04	2.3E+03	2.9E+04	1.2E+05
26	6.58	153	4.0E+04	1.1E+06	8.9E+04	9.9E+04	3.4E+04	4.5E+04
27	7.12	153	4.0E+04	2.6E+03	8.9E+04	6.1E+04	1.9E+05	3.7E+04
28	6.72	154	4.2E+05	5.0E+05	2.9E+03	2.0E+04	3.1E+03	3.9E+04
29	6.98	154	5.3E+05	2.0E+03	6.3E+05	2.0E+04	2.8E+05	2.1E+05
30	7.66	154	2.2E+03	6.7E+04	2.9E+03	2.0E+04	2.8E+05	2.1E+04
31	17.54	159	2.3E+03	2.0E+03	2.8E+03	3.0E+05	2.7E+03	3.8E+04
32	17.87	159	2.3E+03	2.0E+03	2.8E+03	2.8E+04	2.7E+03	1.1E+05
33	18.42	159	2.3E+03	2.0E+03	2.8E+03	4.2E+05	2.7E+03	3.8E+04



2 types of multivariate analyses

non supervised (PCA): **no information provided regarding sample type**

supervised (PLS...): **Introduction of a factor explaining sample variance to optimize sample distinction (e.g., healthy/disease, gender,…)**

+ univariate analyses

Experimental design	Normal distribution (compare means)	For from normal (compare medians)
Compare two unpaired groups	Unpaired t-test	Mann-Whitney
Compare two paired groups	Paired t-test	Wilcoxon signed-rank
Compare more than two unmatched groups	One-way ANOVA with multiple comparison	Kruskal Wallis
Compare more than two matched groups	Repeated-measures ANOVA	Friedman

MetaboHUB
Infrastructure Nationale
en Métabolomique

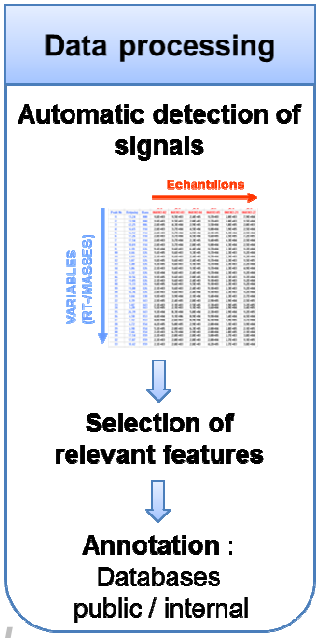
Main steps of a metabolomic analysis

1. Sample preparation

2. Obtention of metabolic profiles

3. Data treatment and statistical analysis

Evidencing biologically-relevant signals



4. Metabolite annotation and identification

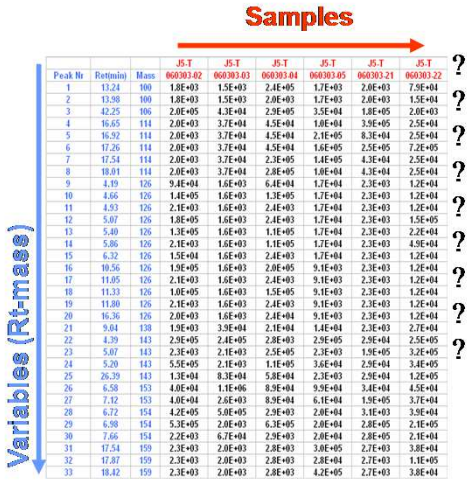
Annotation *vs.* Identification

- **Annotation: One** (or more) **property** (typically mass) **match to databases** (not necessarily acquired under identical analytical conditions)

- **Identification: At least two orthogonal properties** (RT, MS/MS) **compares to authentic chemical standard analyzed under identical conditions**

Sumner et al, Metabolomics 2007

Annotation procedure



1. Annotation using publically available databases

Accurate mass (<1ppm)
Isotopic pattern (^{13}C, ^{34}S, ^{18}O,…)
Molecular Formula

Public databases
(HMDB, Metlin, KEGG)

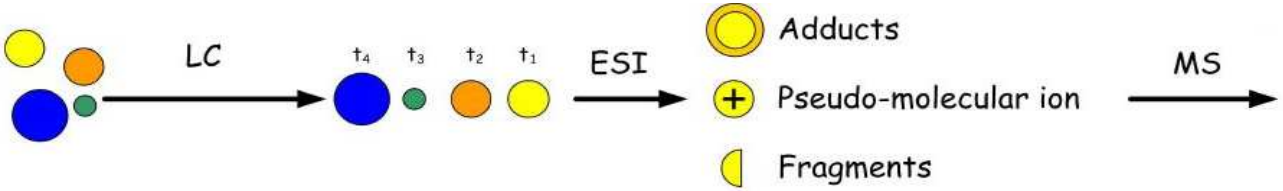
Accurate mass (<1ppm)
Isotopic pattern (^{13}C, ^{34}S, ^{18}O,…)
Molecular Formula
+/- Retention Time

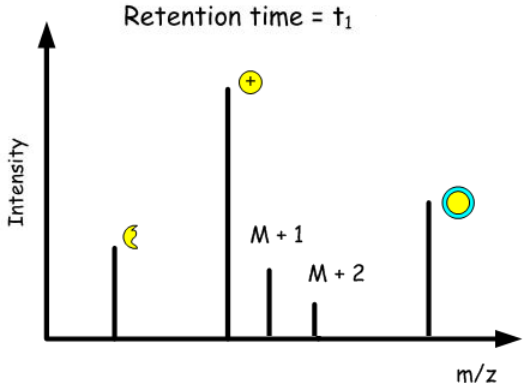
In-house database

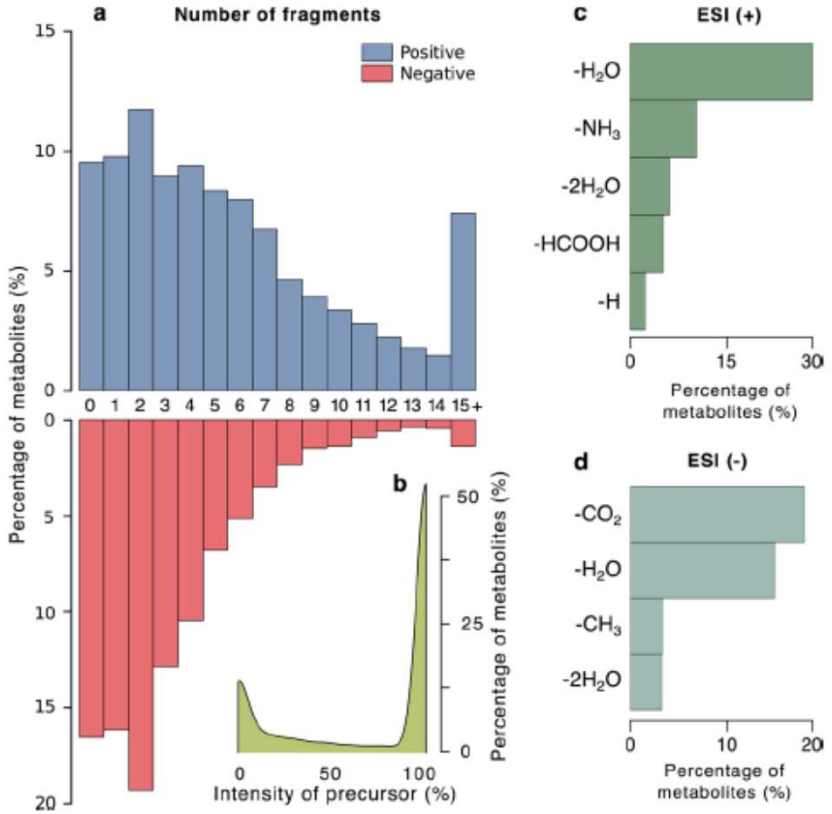
2. Annotation using in-house databases

Relevance of spectral databases

One molecule = several ions





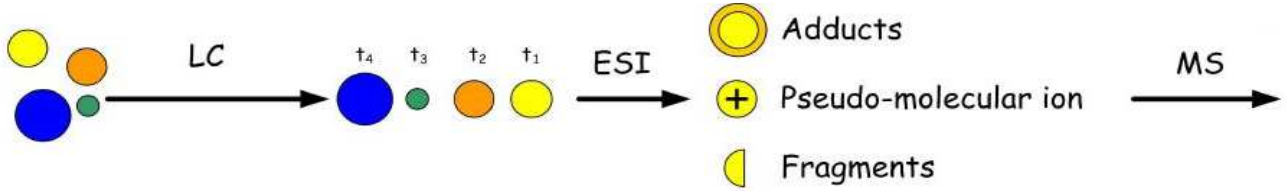


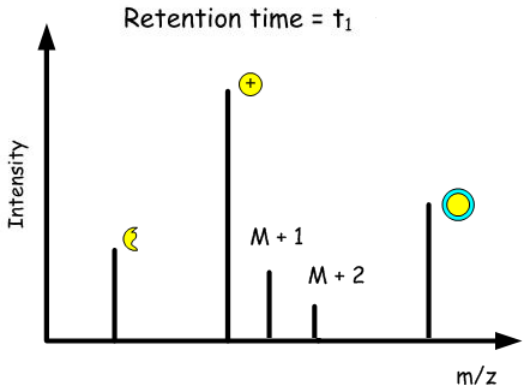
- **METLIN Database** (>10,000 metabolites)
- Up to 15 in-source fragments
- [M+H]+ and [M-H]- as most abundant species in only 50% of the cases

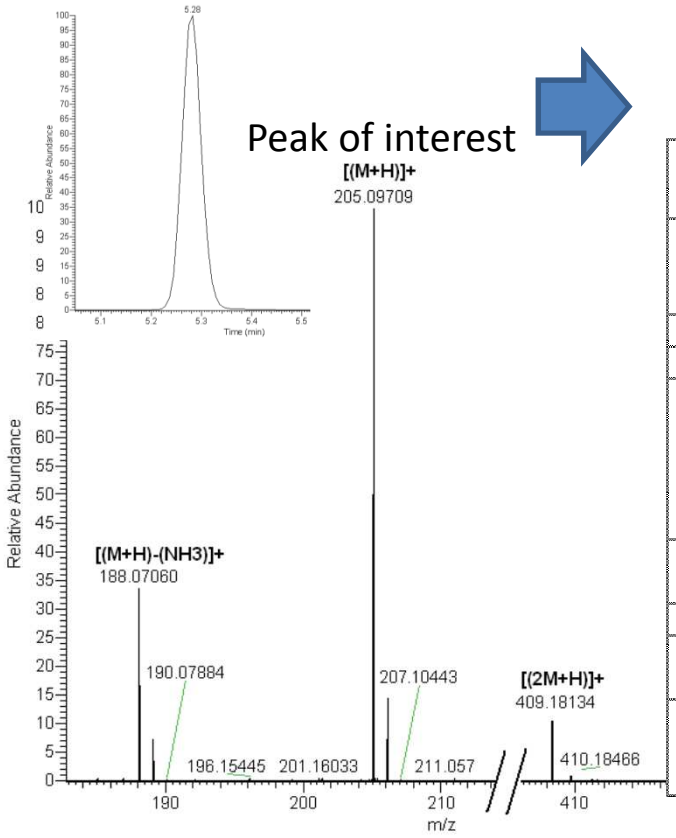
Domingo-Almenara et al, Anal Chem 2018

Relevance of spectral databases

One molecule = several ions





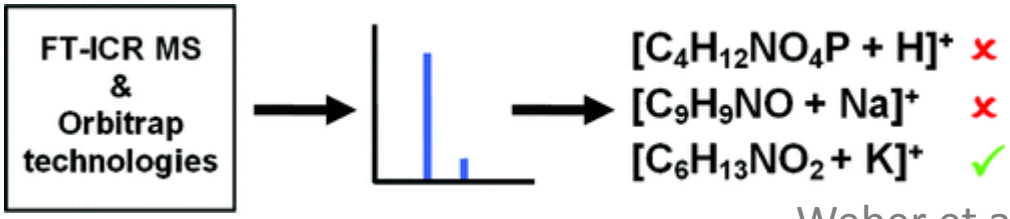


Automated detection of ions, list of annotated features

M/Z	RT	Formula	Compound	Attribution	Annotations (HMDB, KEGG, METLIN)
188.0709	5.28	C11H10NO2	Tryptophan	[(M+H)-(NH3)]+	Deethylatrazine
					3-amino-2-naphthoic acid
					Indoleacrylic acid
189.0757	5.28	C10[13C]H10NO2	Tryptophan	[(M+H)-(NH3)]+ (13C)	Ethyl Oxalacetate
190.0787	5.28	C9[13C]2H10NO2	Tryptophan	[(M+H)-(NH3)]+ (13C2)	
205.0975	5.28	C11H13N2O2	Tryptophan	[(M+H)]+	Tryptophan
					ethotoin
					Vasicinol
					Idazoxan
					Nirvanol
206.1010	5.28	C10[13C]H13N2O2	Tryptophan	[(M+H)]+ (13C)	N-Acetyl-D-fucosamine
					N-Acetyl-D-quinovosamine
207.1051	5.28	C9[13C]2H13N2O2	Tryptophan	[(M+H)]+ (13C2)	
409.1902	5.28	C22H25N4O4	Tryptophan	[(2M+H)]+	Gly Trp Phe (and isomers)
					Lys Met Met (and isomers)
410.1938	5.28	C21[13C]H25N4O4	Tryptophan	[(2M+H)]+ (13C)	Tyr Leu Asp (and isomers)
					Ile Tyr Asp (and isomers)
					Val Tyr Glu (and isomers)

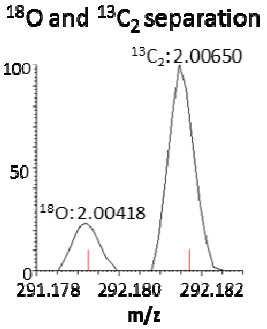
Roux et al., Anal. Chem. 2012

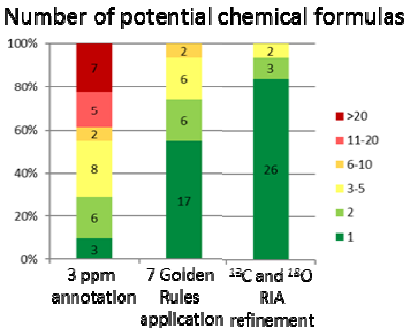
Usefulness of Relative Isotopic Abundances for metabolite annotation



Weber et al, Anal Chem 2011

^{13}C but also ^{18}O, ^{32}S,…

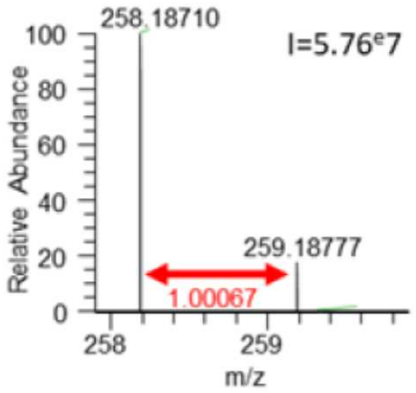




Barbier Saint Hilaire et al., Anal. Chem. 2018

… But beware of space charge effects

(when too many ions are trapped)



Orbitrap Fusion
@240,000
AGC 2^{e}5

1.0007 Da measured *vs* 1.0033 Da expected !!

MetaboHUB
Infrastructure Nationale en Métabolomique

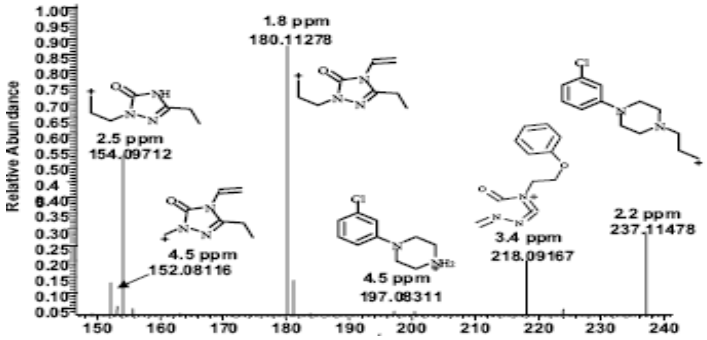
Metabolite Identification

1. Mass Spectrum
- **Identify ion type,** e.g. $[M+H]^+$ or $[M-H]^-$
- **Molecular formula** ($C_xH_yO_z$)
- **Isotopes** (^{13}C, ^{18}O, ^{34}S…)

→ **Annotation**

2. Database

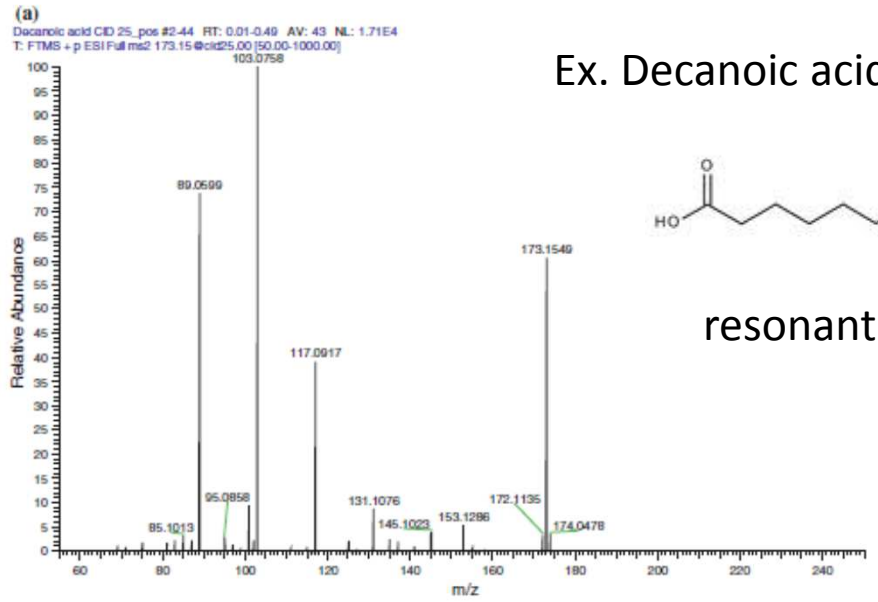
3. MS/MS or MS^n fragmentation (CID, …)



4. Complementary experiments (H/D exchange, NMR)

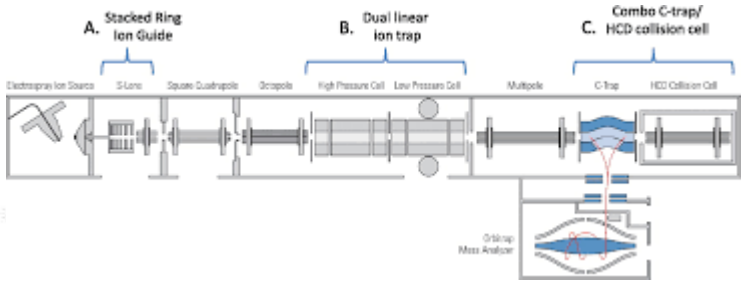
5. Structure confirmation - **Chemical synthesis of standard**
- **Comparative LC-MS analysis**

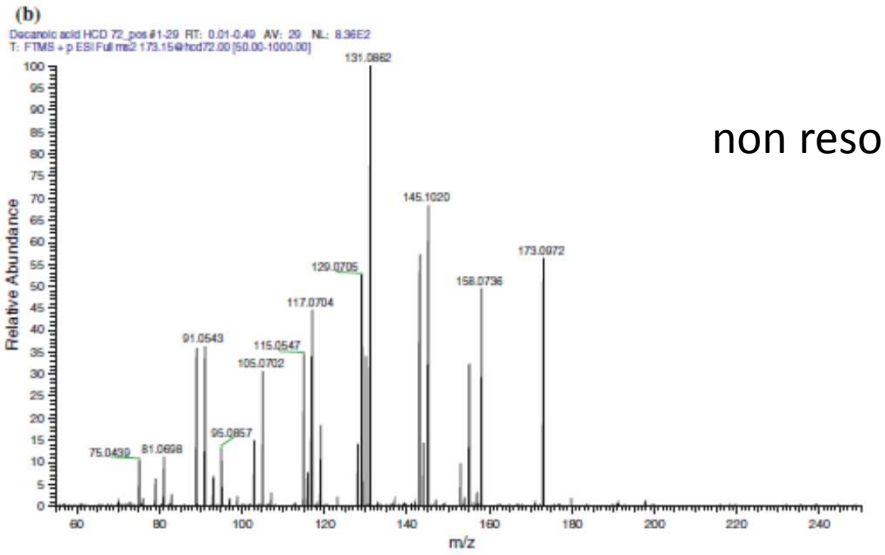
Comparison of MS/MS spectra



Ex. Decanoic acid

resonant





non resonant

Two types of spectra bring different structural information

Dunn et al., Metabolomics 2013

MetaboHUB
Infrastructure Nationale en Métabolomique

Metabolite Identification

4 Levels of confidence

Metabolomics Standard Initiative criteria
(Sumner et al, Metabolomics 2007)

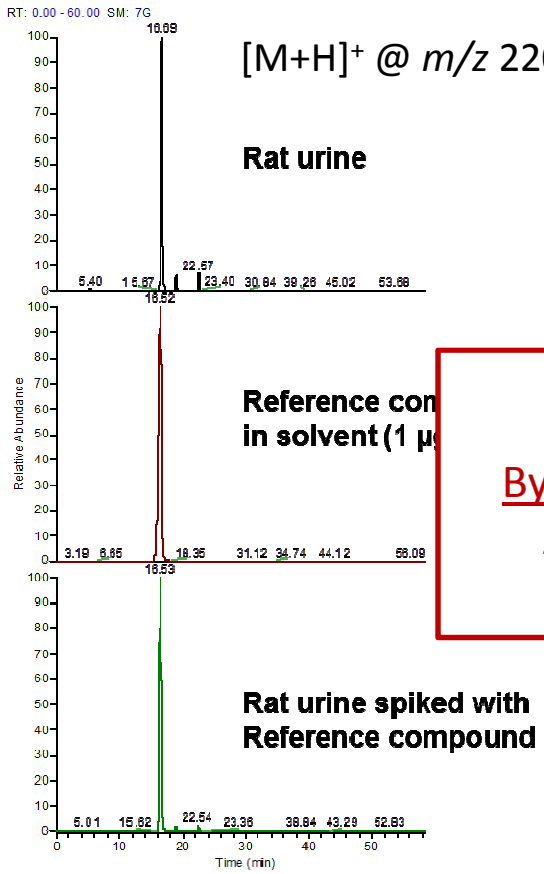
Level	Confidence of Identity	Level of Evidence
1	Identified compounds	Comparison of two or more independent and orthogonal data with **an authentic chemical** standard analyzed under identical experimental conditions
2	Putatively annotated compounds	based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries, **without chemical reference standards**
3	Putatively characterized compounds	Based upon characteristic physicochemical properties of a **chemical class** of compounds, or by spectral similarity to known compounds of a chemical class
4	Unknown compounds	Although unidentified or unclassified these metabolites can still be differentiated and quantified based upon spectral data

MetaboHUB
Infrastructure Nationale en Métabolomique

Formal identification

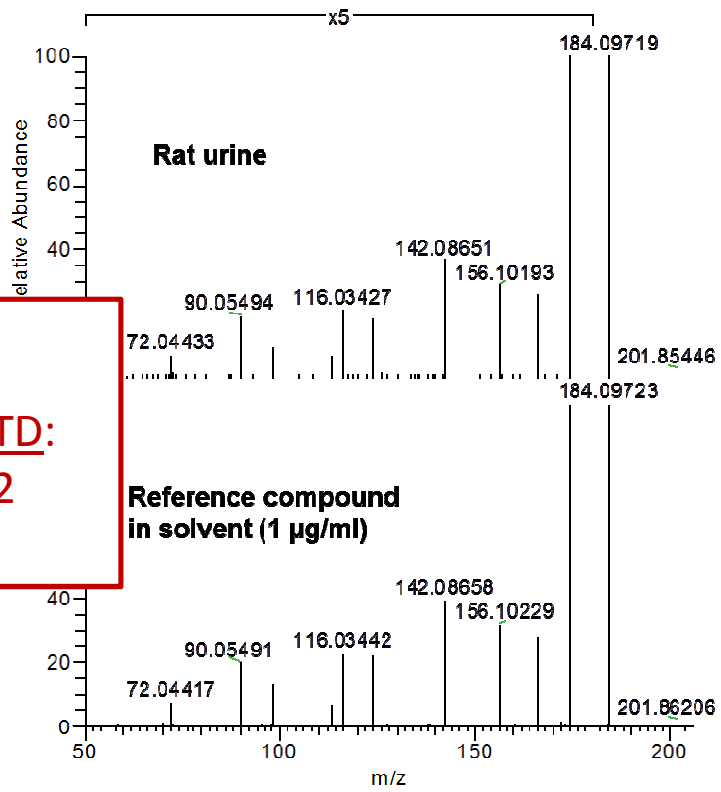
Example of pantothenic acid in rat urine

RP-LCMS, LTQ-Orbitrap



$[M+H]^+$ @ *m/z* 220.1180 (+0.5ppm)

MS/MS spectra (@7,500 reso.)



LEVEL 1

By comparison to pure STD:
Accurate mass, RT, MS2
(2 stereoisomers on HMDB)

Werner E et al, J Chromatogr B 2008

MetaboHUB
Infrastructure Nationale en Métabolomique

Putative annotation

Human urine, RP-HRMS (LTQ-Orbitrap)

$[M-H]^-$ @ *m/z* 595.3463

XCMS output			CAMERA output			Inter-sample correlation	Public database annotation
Variable number	m/z	Retention time (min)	isotopes	adduct	pcgroup		
1806	303.1443	9.33	**	**	531	NA	**
4683	593.3312	9.34	[681][M]+	**	512	NA	L-Urobilin
4688	594.3388	9.34	[681][M+1]+	**	512	NA	**
4679	595.3463	9.40	[650][M]+	[M-H]-	394	1.00	C-Curarine / L-Urobilinogen
4682	596.3514	9.40	[650][M+1				
4878	631.3256	9.40	**				
3797	481.2789	9.48	**				
2763	381.1910	9.53	**				
3834	485.1792	9.61	**				
1255	253.1440	9.67	**				

LEVEL 2

No pure STD available

2 hits in public databases (accurate mass)

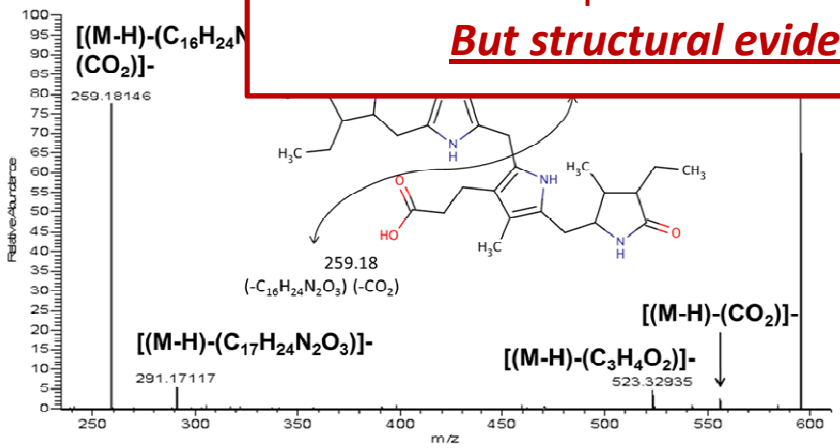
No MS2 spectra in databases

But structural evidence

Urobilinogen or Stercobilinogen ???

Urobilinogen

MS/MS



Roux A et al, Anal Chem 2012

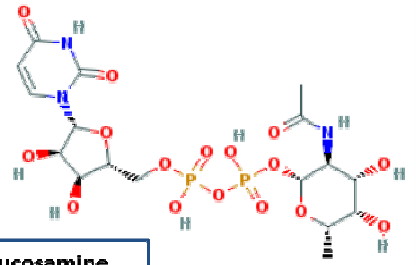
MetaboHUB
Infrastructure Nationale en Métabolomique

Putative characterization

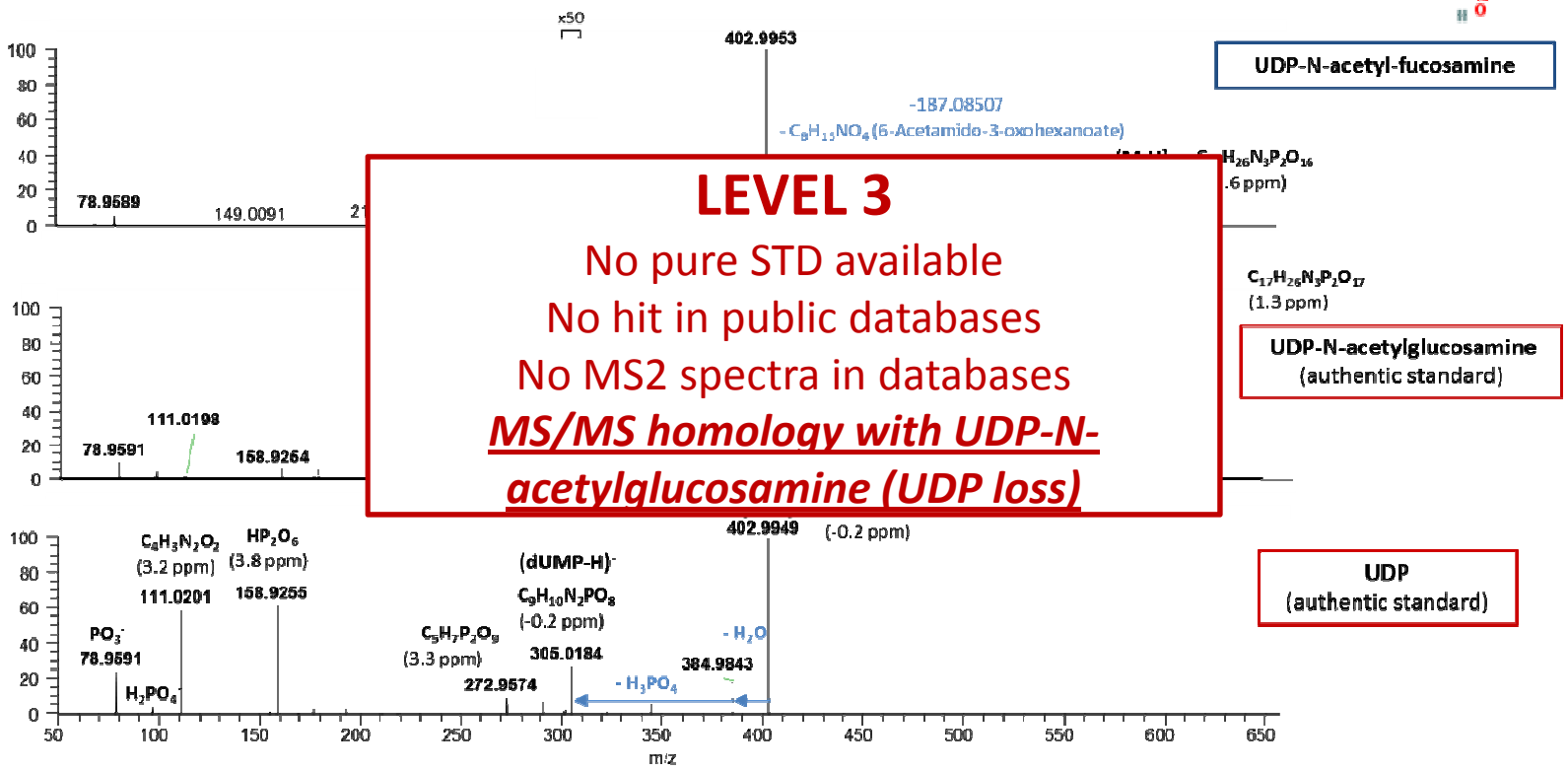
S. aureus metabolic extract, HILIC-HRMS (Q-Orbitrap)

$[M-H]^-$ @ *m/z* 590.0794

MS/MS spectra



UDP-N-acetyl-fucosamine



LEVEL 3

No pure STD available

No hit in public databases

No MS2 spectra in databases

MS/MS homology with UDP-N-acetylglucosamine (UDP loss)

$C_{17}H_{26}N_3P_2O_{17}$ (1.3 ppm)

UDP-N-acetylglucosamine (authentic standard)

UDP (authentic standard)

Aros-Calt S et al, J Proteome Res 2015

metabolites

MDPI

Communication

The Time Is Right to Focus on Model Organism Metabolomes

Arthur S. Edison [1], Robert D. Hall [2], Christophe Junot [3], Peter D. Karp [4], Irwin J. Kurland [5], Robert Mistrik [6], Laura K. Reed [7], Kazuki Saito [8], Reza M. Salek [9], Christoph Steinbeck [9], Lloyd W. Sumner [10] and Mark R. Viant [11,*]

Table 1. Prioritized list of model organisms that the new MOM task group recommend for deep investigations of their metabolomes.

Kingdom	Latin Name	Common Name
Bacteria	*Escherichia coli*	-
Fungi	*Saccharomyces cerevisiae*	yeast
Animal (invertebrate)	*Caenorhabditis elegans* *Daphnia magna* *Drosophila melanogaster* *	nematode water flea fruit fly
Animal (vertebrate)	*Danio rerio* *Mus musculus*	zebrafish mouse
Plant	*Arabidopsis thaliana* ** *Medicago truncatula* *Oryza sativa* *Solanum lycopersicum*	thale cress barrel medic, model legume rice tomato

* International Drosophila Metabolomics Curation Consortium [30]; ** Metabolomics subcommittee (chaired by Kazuki Saito) within the Multinational Arabidopsis Steering Committee [31].

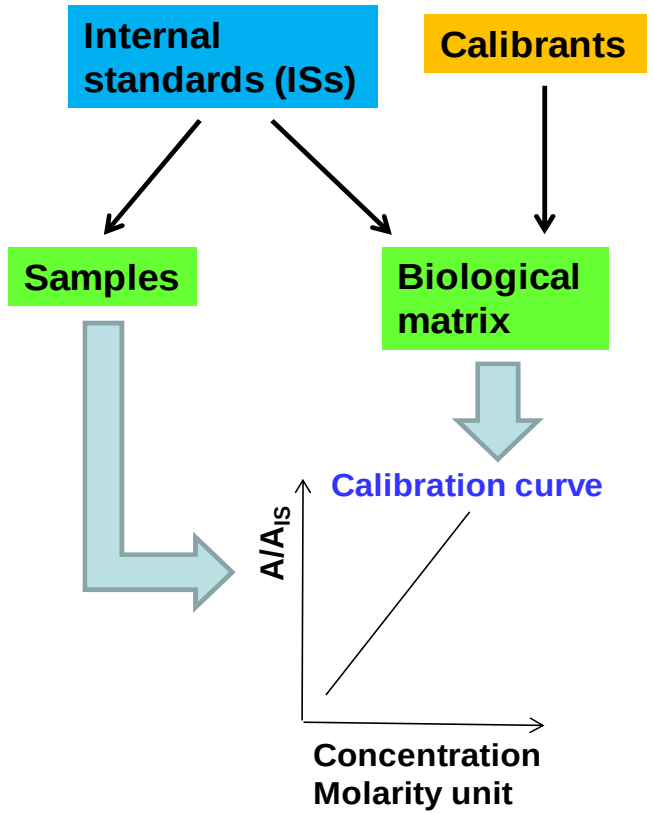
Summary and prospects

- **FTMS of great value to the field:** mass accuracy, high resolution (separation of isobaric species), structural elucidation
- **Metabolomics tools** (data acquisition and treatment) **are constantly improving**
- **Still need to standardize** (e.g., MS and MS/MS data acquisition) **and share informatics tools and databases**
- **Imaging mass spectrometry**
- **Ion mobility**
- **Integration of multi-omics data**
- **Toward large-scale quantitative high throughput metabolomics?**
- **How can we expedite metabolite identification?**

Toward large-scale quantitative Metabolomics

MetaboHUB Infrastructure Nationale en Métabolomique

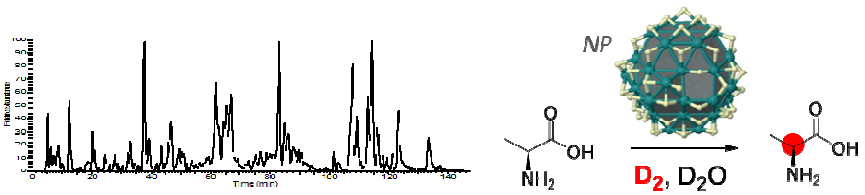
Absolute quantification of metabolites



- ^{13}C- and/or ^{15}N based Metabolic labeling
- Chemical labeling
- Derivatization: dansyl chloride…

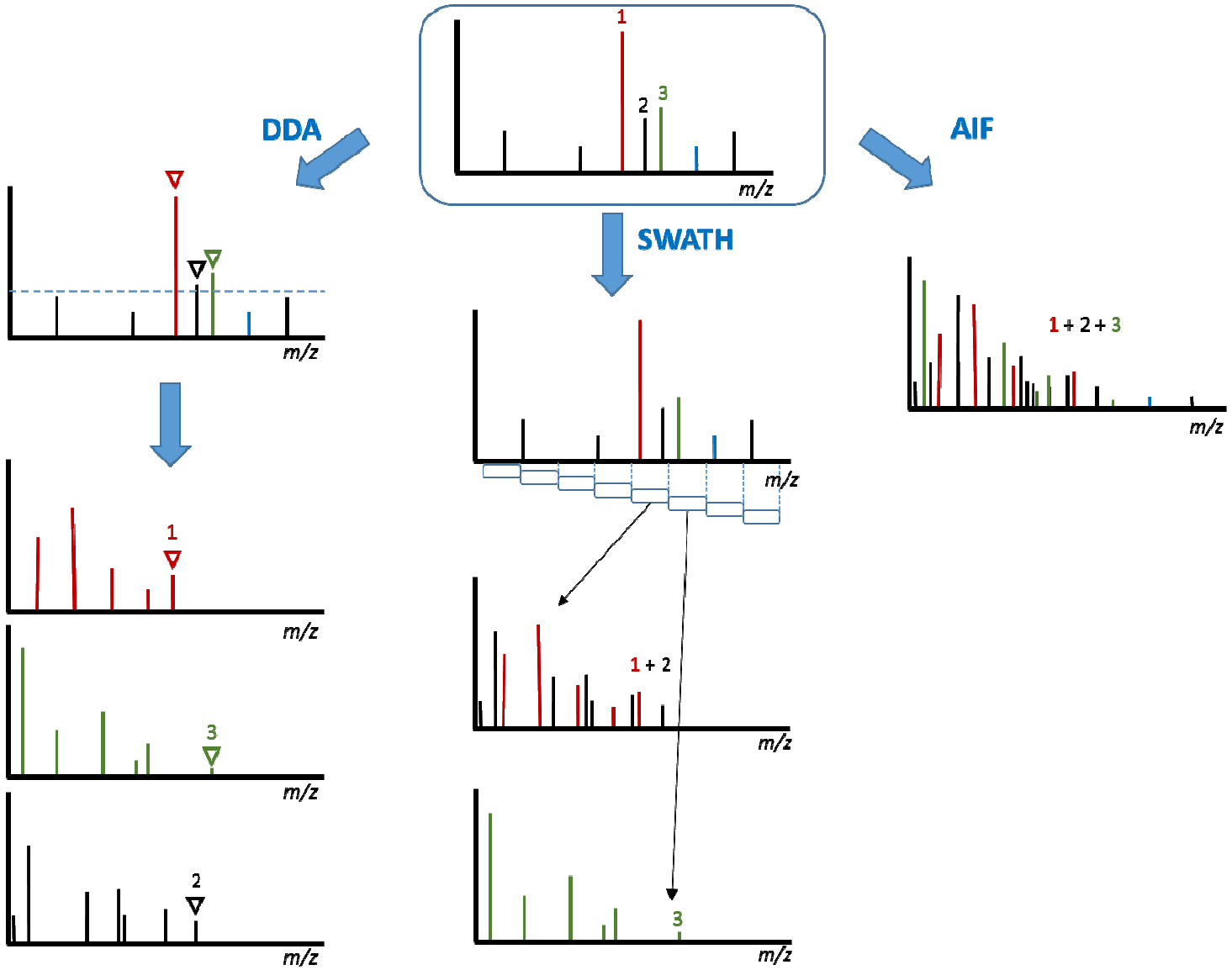
Guo K et al, Anal Chem 2009

- High throughput synthesis of deuterated internal standards



Taglang C et al, Angew Chem Int Ed 2015

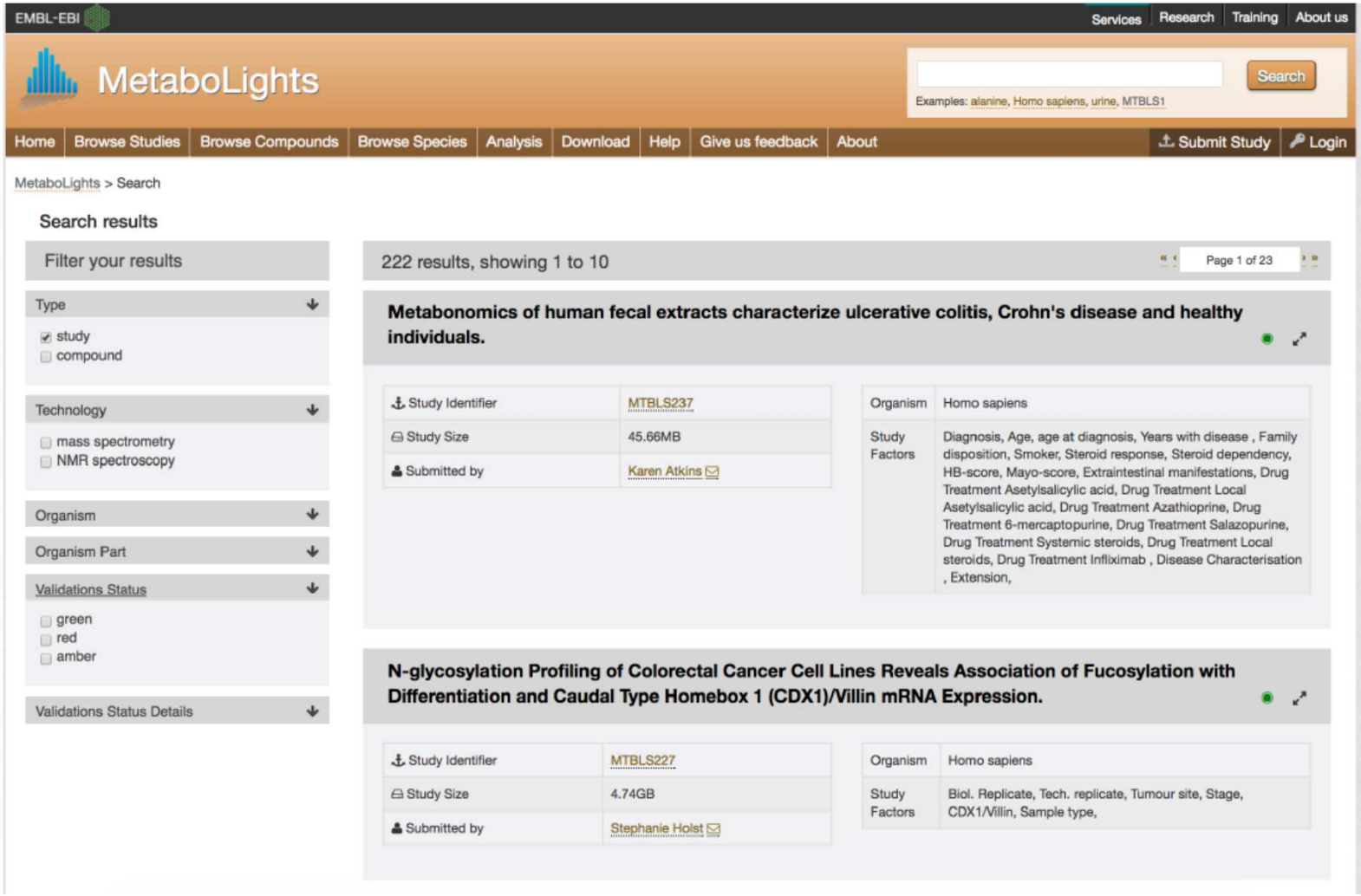
Toward new data acquisition workflows



Mullard G et al, Metabolomics 2015

Bonner R et al, Bioanalysis 2015

MetaboLights: Metabolomics data sharing

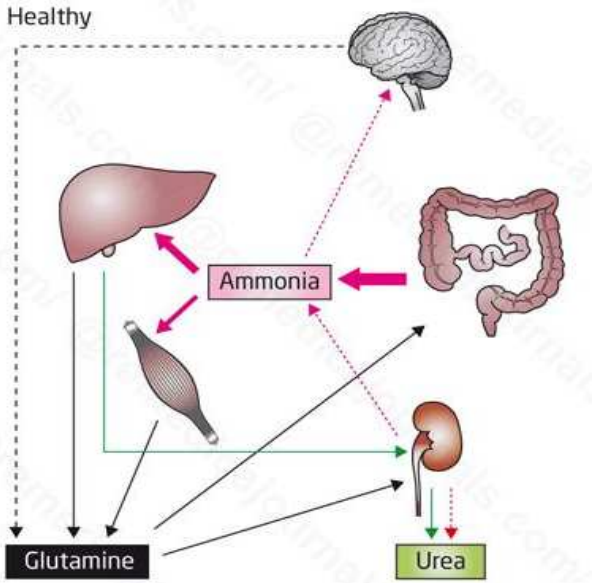


Steinbeck C et al, Metabolomics 2012

A few biological examples

Metabolomics for the study of liver diseases

- **Hepatic Encephalopathy (HE) is a neurological complication observed in patients with liver diseases (e.g., ACLF)**
- **The proportion of cirrhotic patients developing overt HE is about 40-60%**



However, the pathophysiological mechanism of HE remains poorly understood:

- Hyperammonemia
- Inflammation
- Altered permeability of blood-brain barrier

AIM OF THE STUDY. To highlight altered metabolic pathways in HE patients by using CSF LC/MS-based metabolomics

- **patient stratification**
- **pharmacological targets**

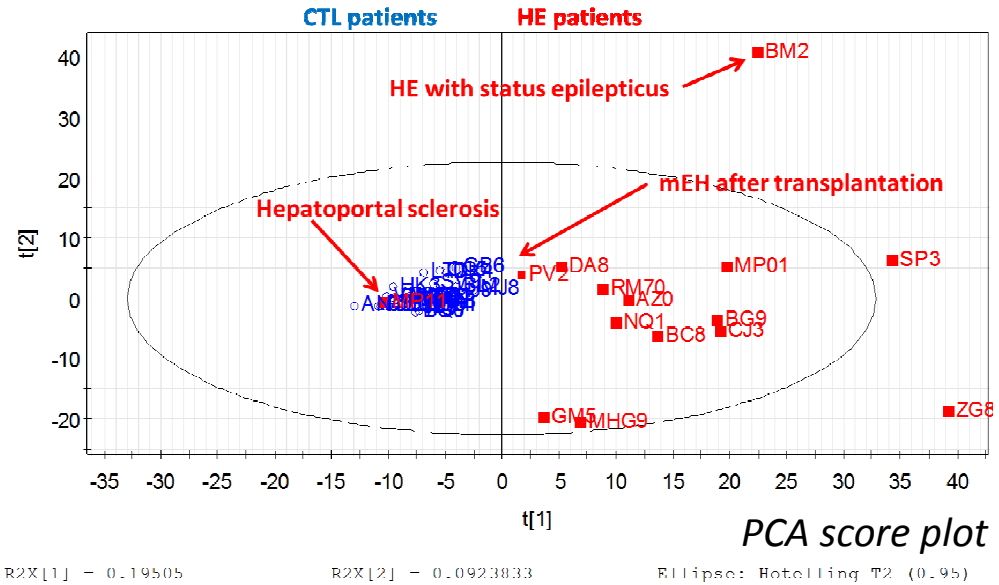
Weiss et al, J Hepatol 2016

Metabolomics for the study of liver diseases

MetaboHUB Infrastructure Nationale en Métabolomique

27 control subjects and 14 HE patients

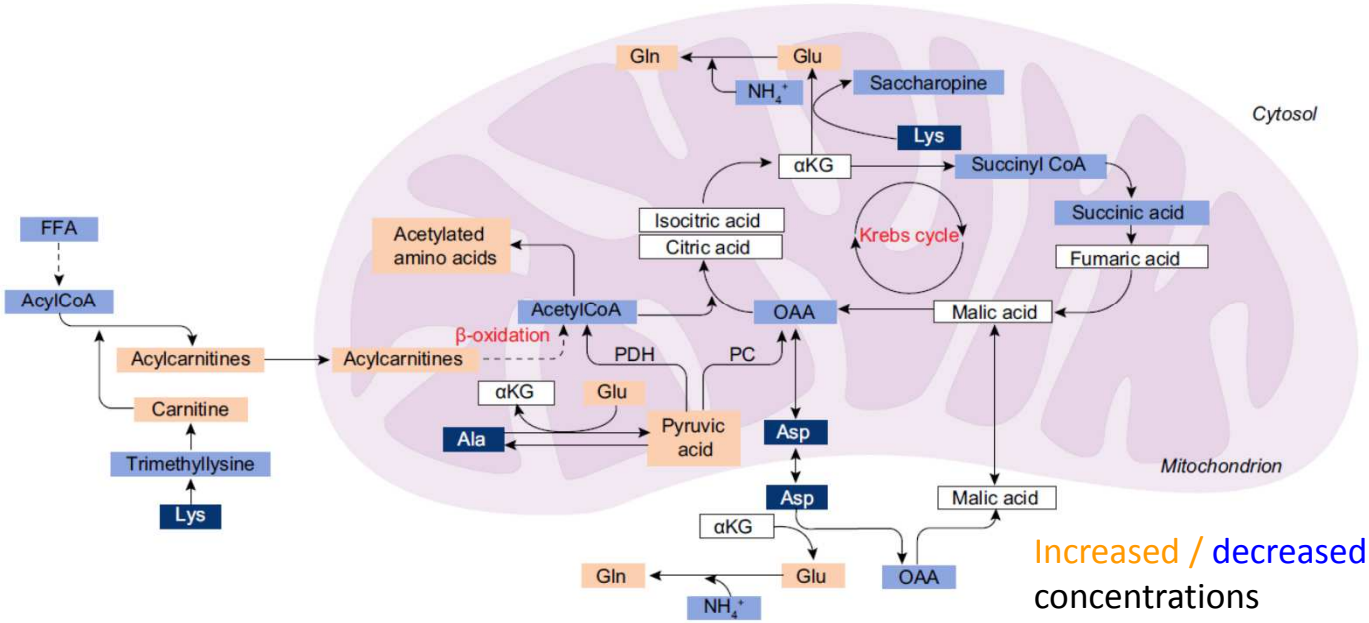
HE and control patients have different CSF-metabotypes



↳ ***72 metabolites** (over 122 monitored) **with altered concentrations in HE***

Weiss et al, J Hepatol 2016

Metabolomics for the study of liver diseases



Increased / decreased metabolite concentrations

CSF metabolomics highlights alterations of metabolic pathways linked to energy metabolism that are not observed in plasma samples.

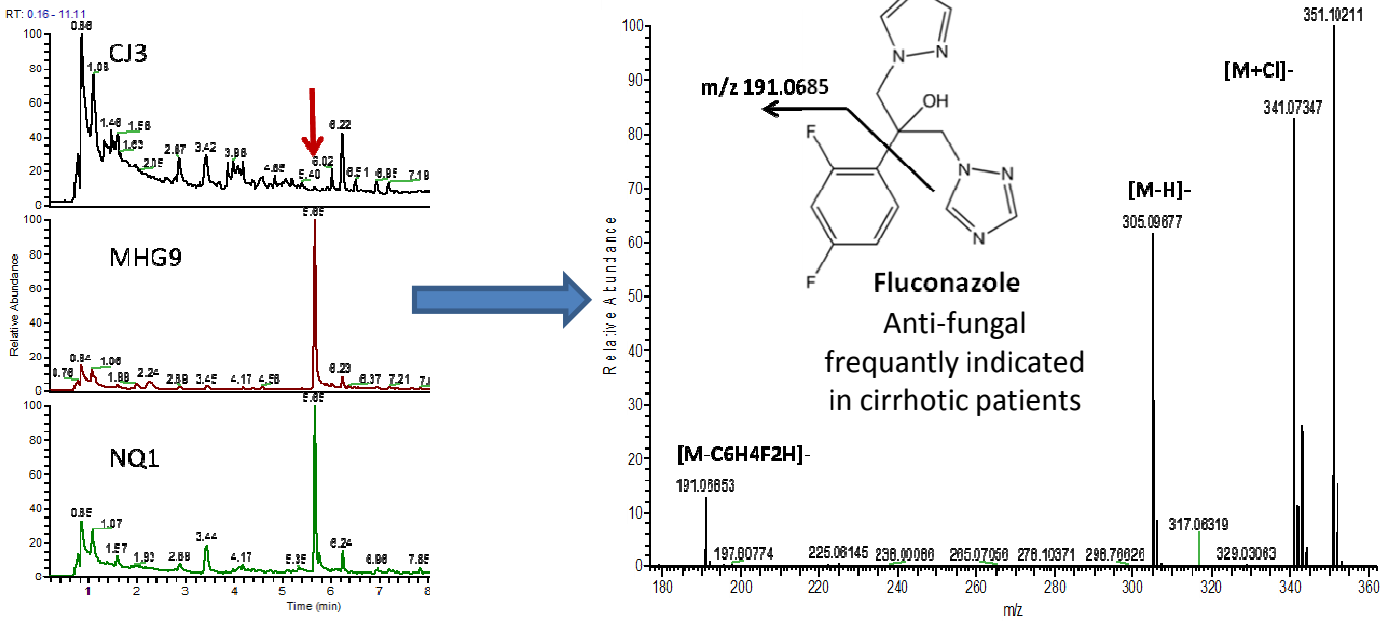
Energy metabolism as a pharmacological target for HE??

Weiss et al, J Hepatol 2016

Metabolomics for the study of liver diseases

MetaboHUB Infrastructure Nationale en Métabolomique

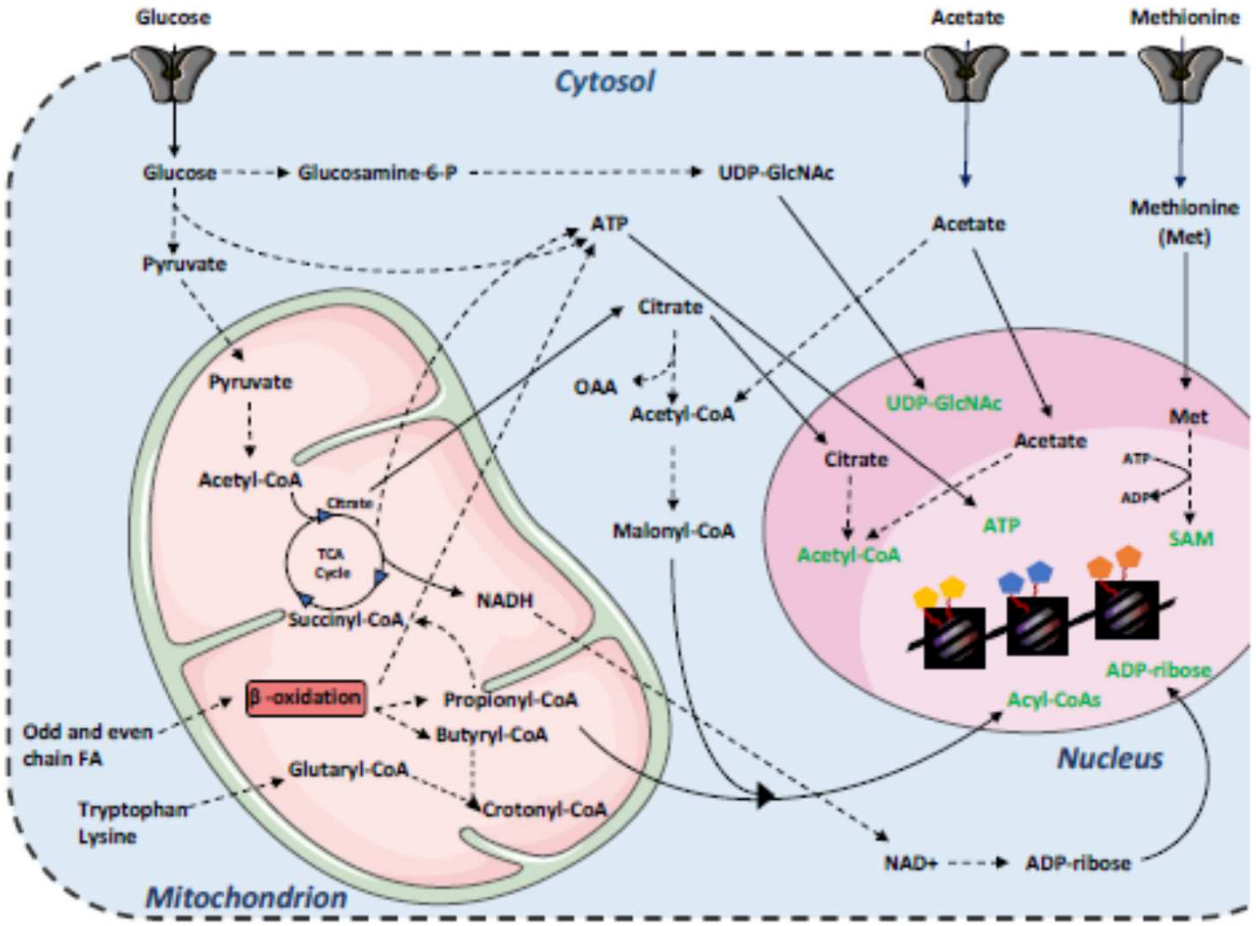
Presence of drugs highlighted by untargeted Metabolomics



Putative annotation	Samples
Levetiracetam	BM2
Metronidazole	BG9
Fluconazole	MGH9, NQ1, ZG8, AZ0
Diazepam	BC8, MHG9, MP01, NQ1, ZG8
N-Desmethyldiazepam	BC8, MHG9, MP01, NQ1, ZG8
Tazobactam	BG9, NQ1, ZG8
Piperacillin	BG9, NQ1, ZG8
Ciprofloxacin	AZ0, BG9, NQ1, ZG8, SP3
Norfloxacin	CJ3, BC8

Weiss et al, J Hepatol 2016

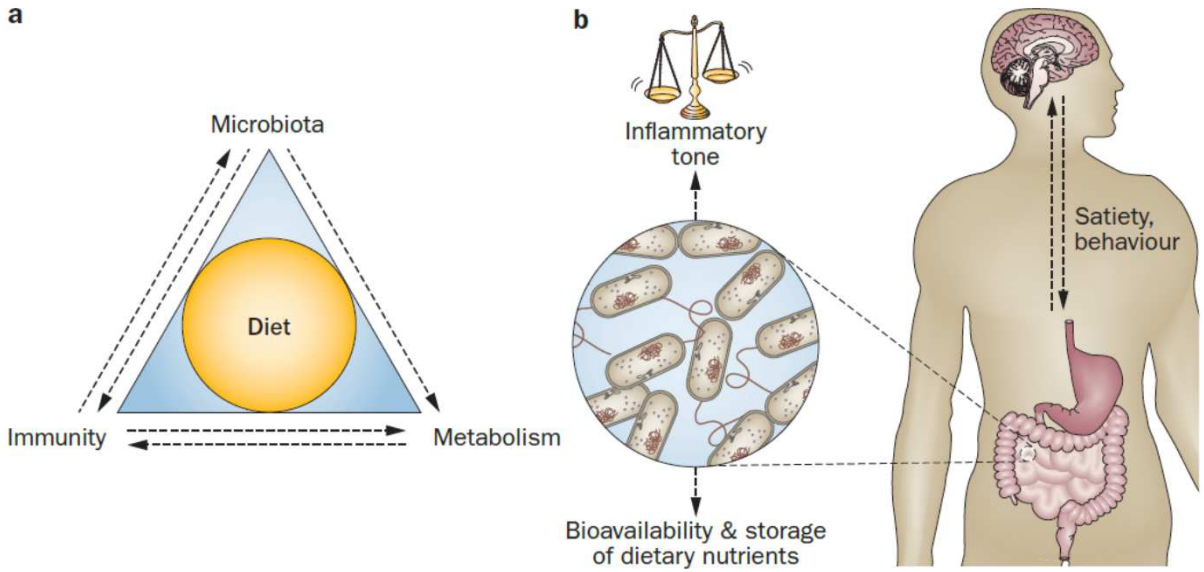
Integrating proteomics and metabolomics to understand changes in histone modifications



Simithy et al, Nat Commun 2017; Simithy et al, Proteomics 2018

MetaboHUB
Infrastructure Nationale
en Métabolomique

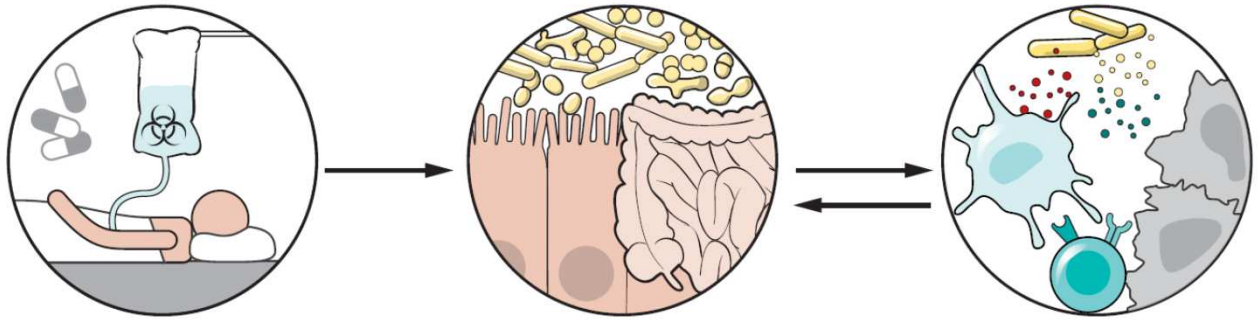
What's next ? Links between Metabolism/Microbiota/Immunity



Cancer therapies
Anticancer treatment modalities and co-medications (such as antibiotics) affect the integrity of the epithelial barrier.

Microbiome
Gut-resident commensals interacting with epithelial, stromal, endocrine, neural, immune intestinal cells to regulate barrier functions and whole-body metabolism.

Immune responses
The gut microbiota has systemic effects throughout the meta-organism via secretion of anti-inflammatory cytokine/chemokines, metabolites, antimicrobial and neuropeptides.

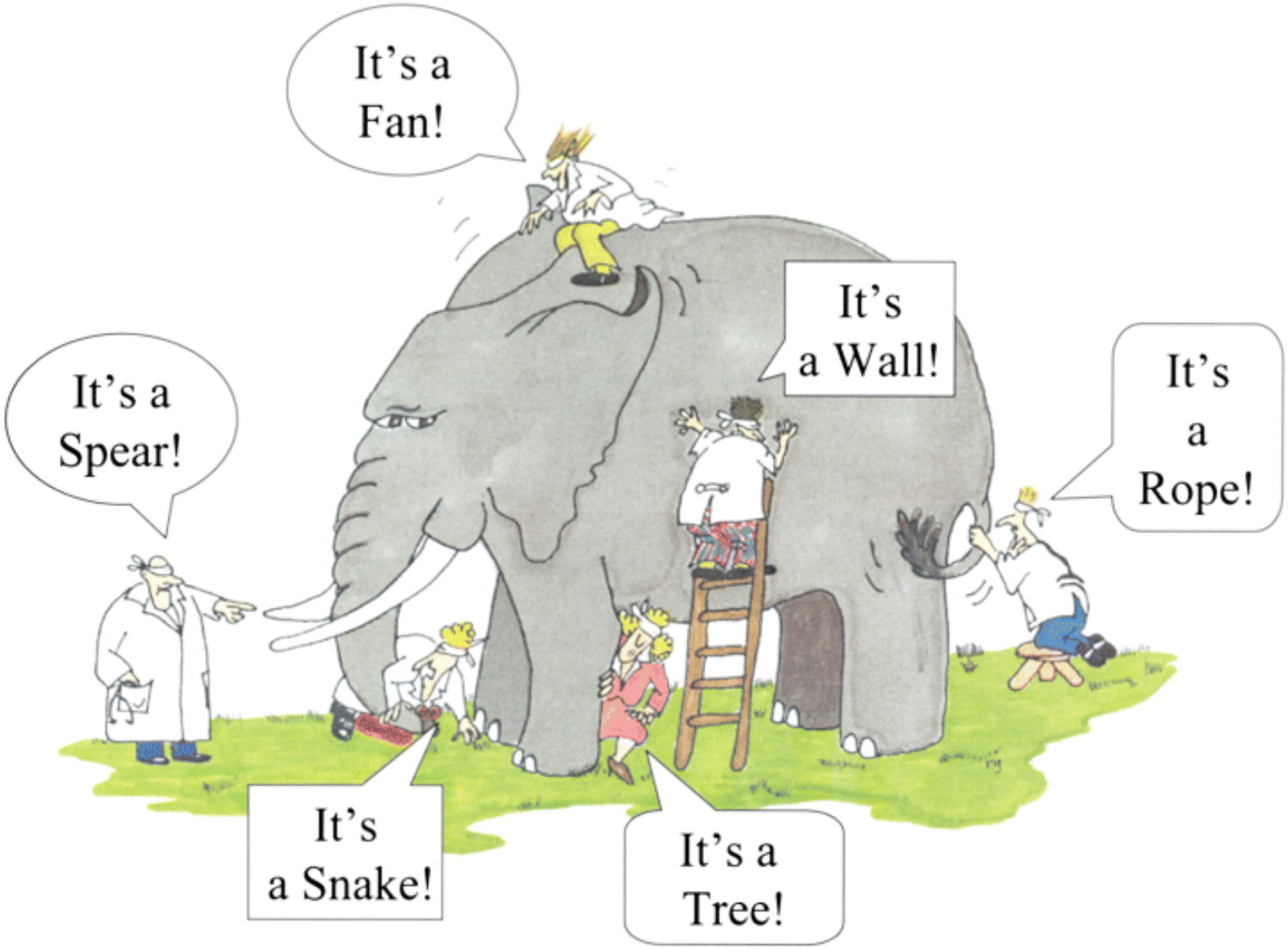


Shanahan et al, Nat Rev Gastroenterol 2012; Zitvogel et al, Science 2018.

MetaboHUB
Infrastructure Nationale
en Métabolomique

System biology:

Six blind biologists examining an elephant



Acknowledgements

CEA/SPI/LEMM

Florence Castelli
Benoit Colsch
Annelaure Damont
Patricia Lamourette
Marie-Françoise Olivier
Emeline Chu-Van
Pierre Barbier Saint Hilaire
Kathleen Rousseau
Hélène Cazier
Thais Hautbergue

Jean-Claude Tabet
Anna Warnet

CEA/SPI

Christophe Junot
Sandrine Leblois
Laurie Ménez
Emmanuel Favry
Florence Vizet

CEA/LIST

Etienne Thévenot
Natacha Lenuzza
Alexis Delabrière

MedDay

Sandrine Aros
Alexandre Seyer
Stéphanie Oursel
Fanny Leroux
Simon Broudin
Jennife Hamoniaux

Sandra Alves
Estelle Paris
Denis Lesage

And thank you for your attention!!!