MetaboHUB
Infrastructure Nationale
en Métabolomique

# Bases de données en métabolomique

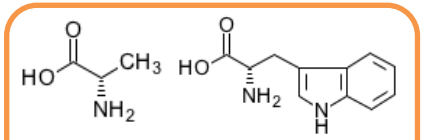
**Christophe Junot**

**CEA/Laboratoire d'Etude du Métabolisme des Médicaments**
**CEA-Saclay (iBiTec-S)**
**christophe.junot@cea.fr**

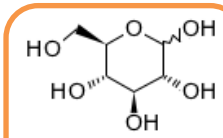
MetaboHUB
Infrastructure Nationale
en Métabolomique

# Metabolites and metabolome

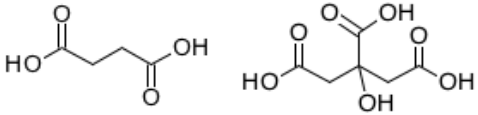
## Primary metabolites



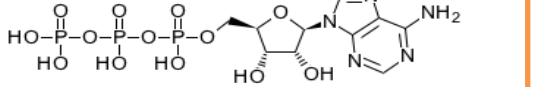
Aminoacids



Sugars



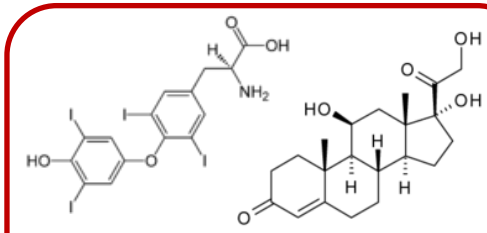
Organic acids



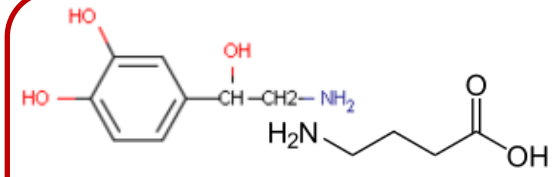
Nucléotides

(…)

## Secondary metabolites



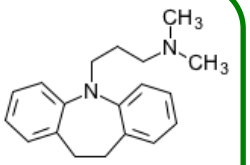
Hormones



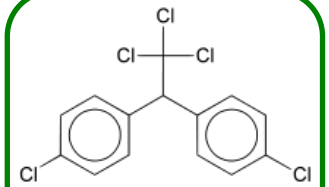
Neurotransmitters

(…)

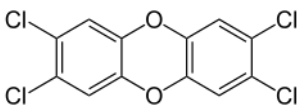
## Xenobiotics



Drugs



Pesticides



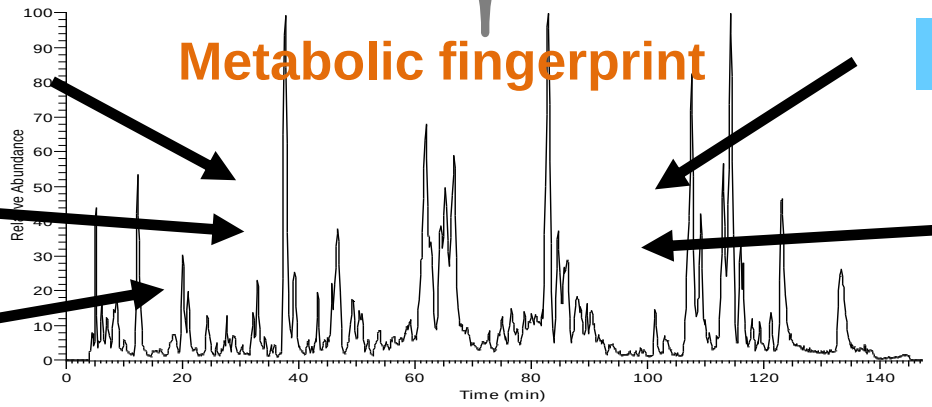
Pollutants

(…)

Food / drinking

Gut microbiota

Pathology



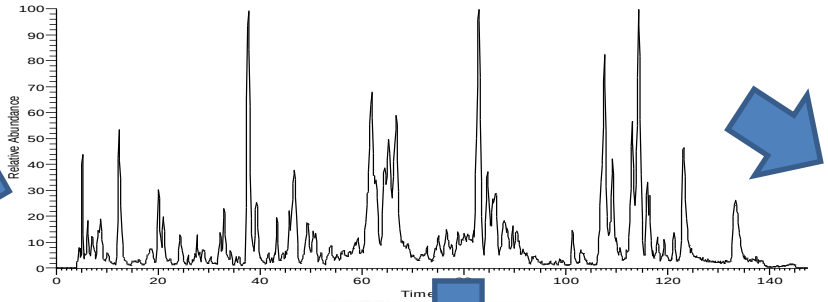

Central metabolism

Environment : xenobiotics (pollutants, drugs…)

# 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-API-MS

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

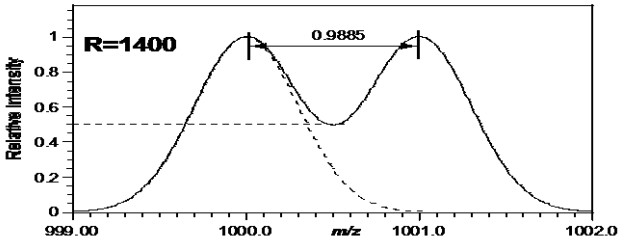
But:

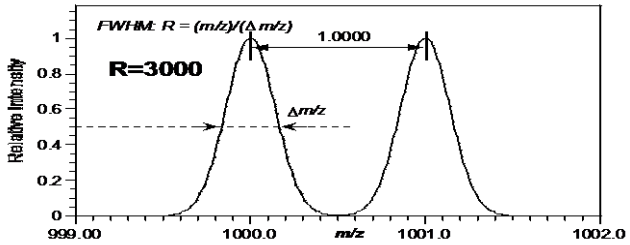
Poor inter-platform reproducibility

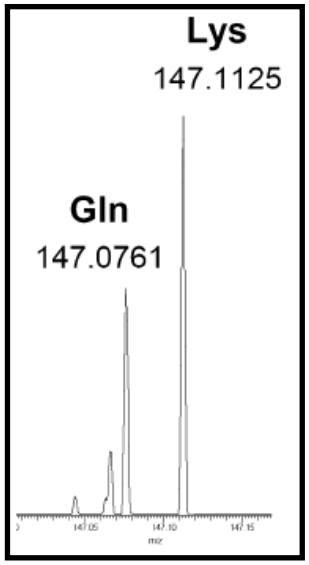
MetaboHUB
Infrastructure Nationale
en Métabolomique

# MS à haute résolution: détecter plus de métabolites et les identifier plus facilement

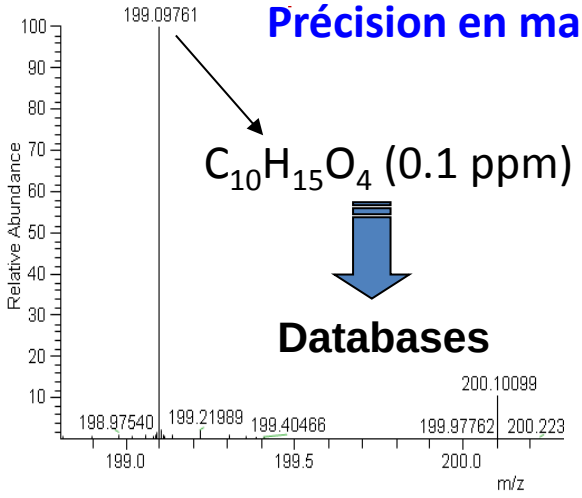
## Résolution en masse



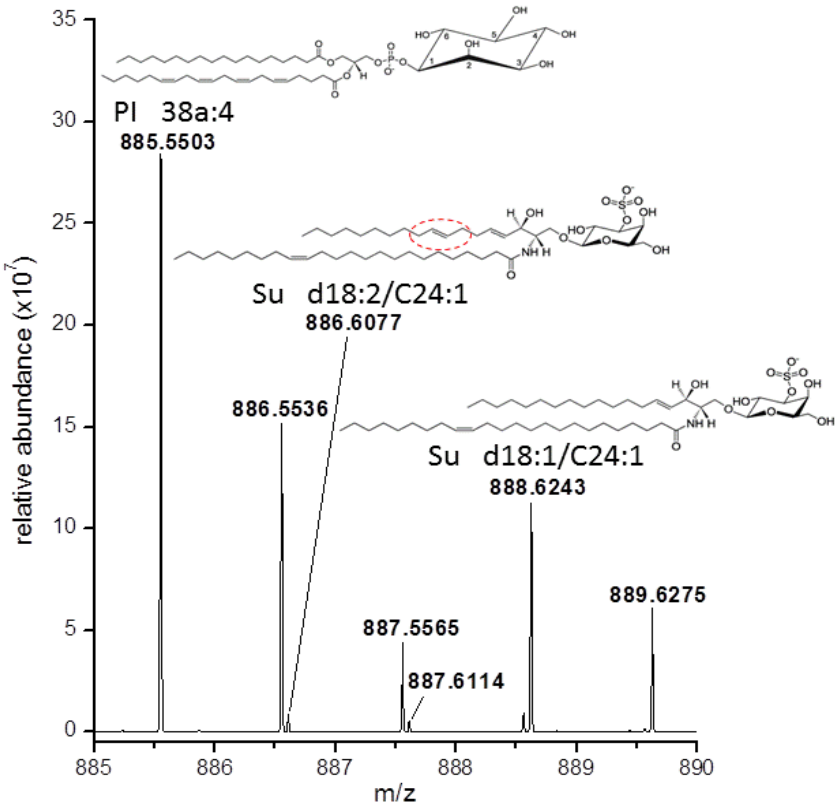




## Précision en masse




$C_{10}H_{15}O_4$ (0.1 ppm)

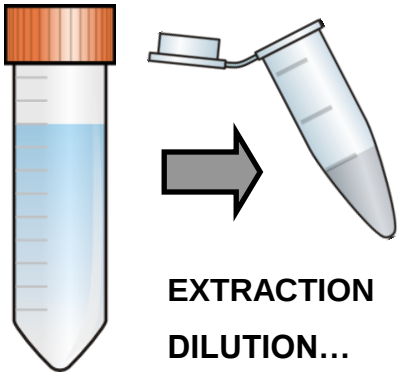
**Databases**



# Déroulement d'une analyse métabolomique

MetaboHUB
Infrastructure Nationale
en Métabolomique

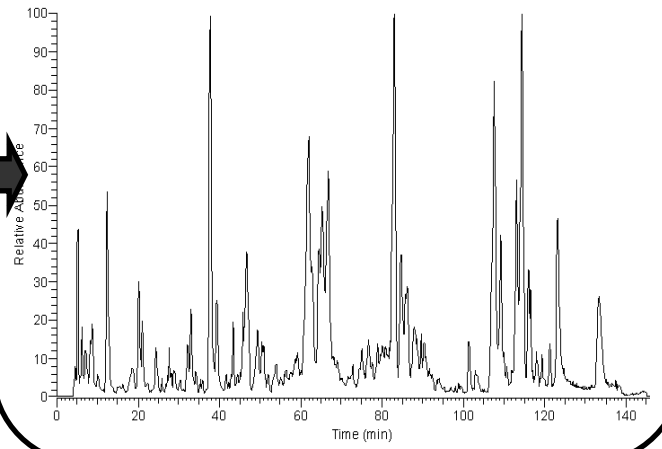
## PREPARATION DES ECHANTILONS



EXTRACTION

DILUTION…

## ACQUISITION DES EMPREINTES

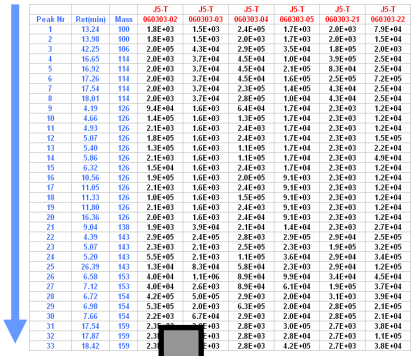



## TRAITEMENT DES DONNEES

### DETECTION AUTOMATIQUE DES SIGNAUX

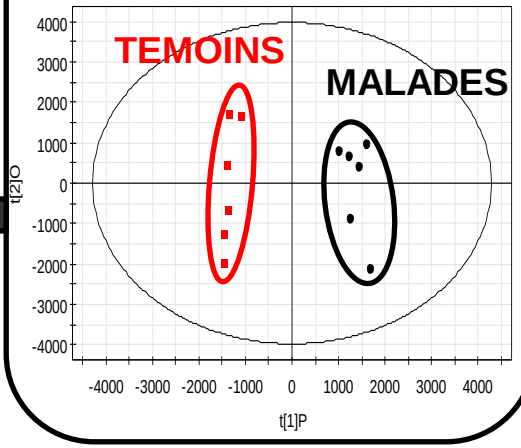
ECH.

VARIABLES (RT-MASSES)



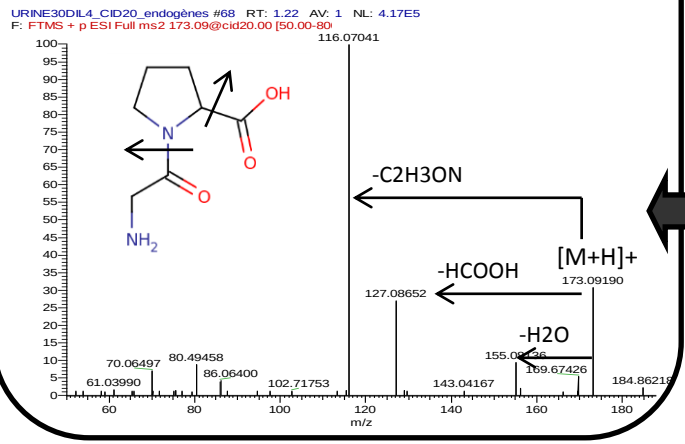
| Peak Nr | Ret(min) | Mass | J5-T 060303-02 | J5-T 060303-03 | J5-T 060303-04 | J5-T 060303-05 | J5-T 060303-21 | J5-T 060303-22 |
|---|---|---|---|---|---|---|---|---|
| 1 | 13.24 | 100 | 1.8E+03 | 1.5E+03 | 2.4E+05 | 1.7E+03 | 2.0E+03 | 7.9E+04 |
| 2 | 13.98 | 100 | 1.8E+03 | 1.5E+03 | 2.0E+03 | 1.7E+03 | 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 | 2.5E+03 | 2.8E+03 | 3.0E+05 | 2.7E+03 | 3.8E+04 |
| 32 | 17.87 | 159 | 2.3E+03 | 2.8E+03 | 2.8E+03 | 2.8E+04 | 2.7E+03 | 1.1E+05 |
| 33 | 18.42 | 159 | 2.3E+03 | 2.8E+03 | 4.2E+05 | 2.7E+03 | 2.7E+03 | 3.8E+04 |

### DETECTION AUTOMATIQUE DES SIGNAUX

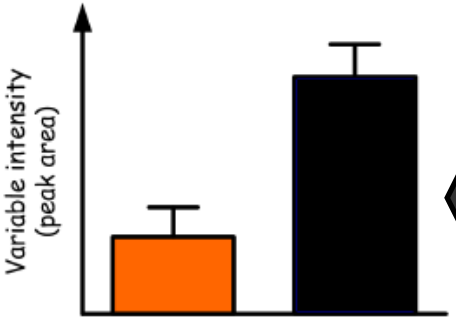



## IDENTIFICATION DES BIOMARQUEURS

### BASES DES DONNEES, MS/MS …




## CONFIRMATION ET QUANTIFICATION



MetaboHUB
Infrastructure Nationale
en Métabolomique

# Les différents types de bases de données en métabolomiques

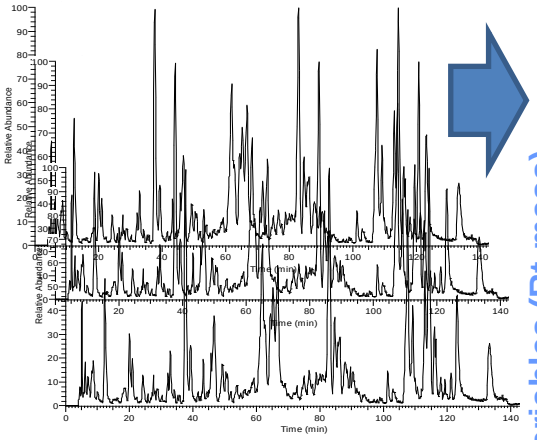
## Bases de données biochimiques/métaboliques

annotation

## Bases de données spectrales:

ESI/MS: annotation

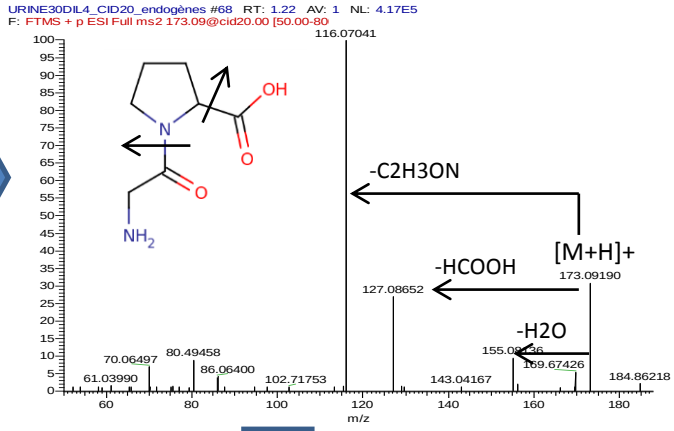
MS/MS: identification

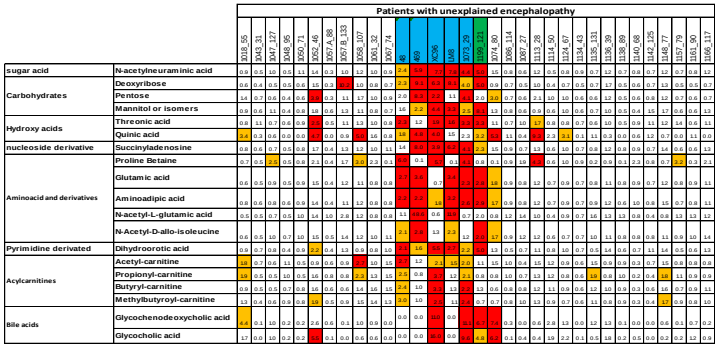



Samples

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 |




## Bases de données de profils métaboliques

MetaboHUB
Infrastructure Nationale
en Métabolomique

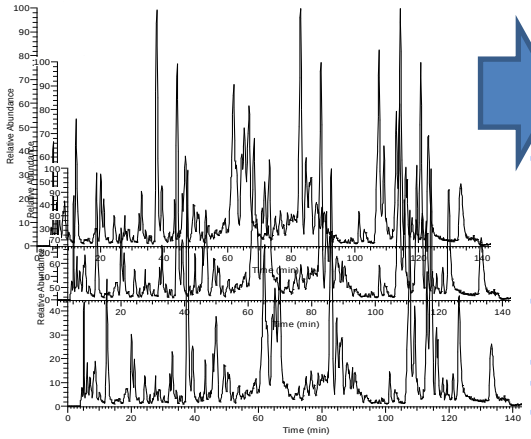
# Spectral databases

| database | | thematic | Conception / URL | Instrument |
|---|---|---|---|---|
| **NIST** | ⦿ | general | National institute for standard and technology (USA)<br>www.nist.gov/srd/nist1a.htm | GC/MS |
| **Fiehn Library** | ● | general | Fiehn Laboratory Univ California Davis – Genome center<br>http://fiehnlab.ucdavis.edu/Metabolite-Library-2007 | GC/MS |
| **Golm** | ● | plant | Max Planck Institute for Molecular Plant Physiology (Germany)<br>csbdb.mpimp-golm.mpg.de | GC/MS |
| **HMDB** | ● | human metabolites | Department of Computing Science, University of Alberta (Canada)<br>www.hmdb.ca/extrIndex.htm | NMR, API/MS/MS |
| **Lipidmaps** | ● | lipidomics | LIPID MAPS Bioinformatics Core (USA)<br>www.lipidmaps.org/data/index.html. | API-MS/MS |
| **Massbank** | ● | general | Keio university, university of Tokyo, Kyoto university, RIKEN plant Science center (Japan) and others<br>www.massbank.jp | API-MS/MS |
| **Metlin** | ● | human metabolites | Scripps Center for Mass Spectrometry<br>metlin.scripps.edu | API-MS/MS |
| **Brucker** | ○ | general | | NMRS |
| **Madison Metabolomic Consortium database** | ● | general | http://mmcd.nmrfam.wisc.edu/ | NMRS |

● free access, ⦿ partially free access, ○ licenced

MetaboHUB
Infrastructure Nationale
en Métabolomique

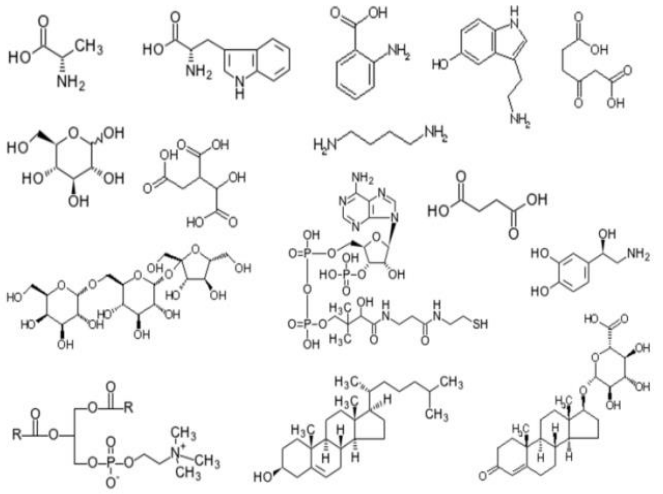
# Annotation of peak lists is required to help for metabolite identification




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…**

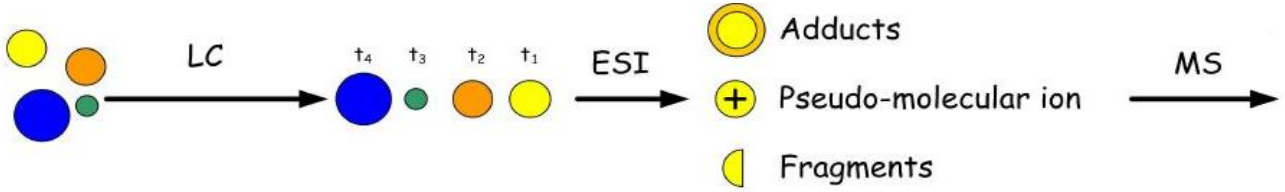
**…Few hundreds of metabolites ??**

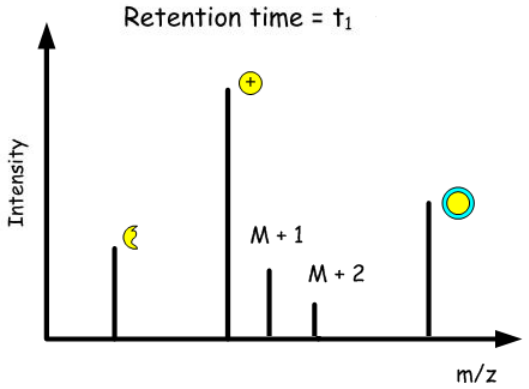
- **Chemical and biochemical databases: KEGG (**www.genome.jp/kegg**), Metlin (**www.metlin.scripps.edu**), HMDB (**www.hmdb.ca)
- **spectral databases**

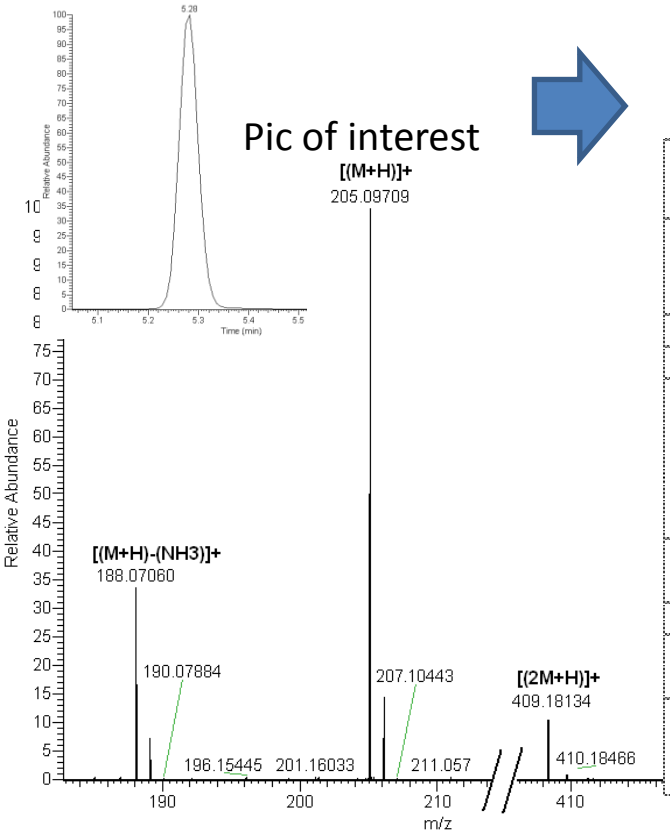
MetaboHUB
Infrastructure Nationale
en Métabolomique

# The relevance of a spectral database

## 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., PhD work, 2008-2011, Roux et al., Anal. Chem. 2012)**

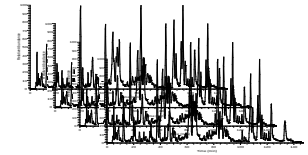
MetaboHUB
Infrastructure Nationale en Métabolomique

# Development of an ESI-mass spectral database for metabolomics : methodology

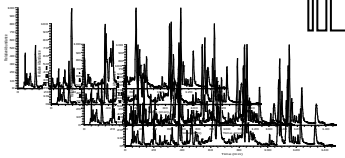
**Chemical library (reference compounds)**

**ESI mass spectral database**

- Gather together signals from the same metabolite
- $MS^n$ confirmation



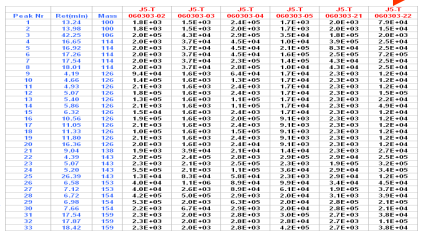
Control



Disease

Samples

Variables (Tr-masse)



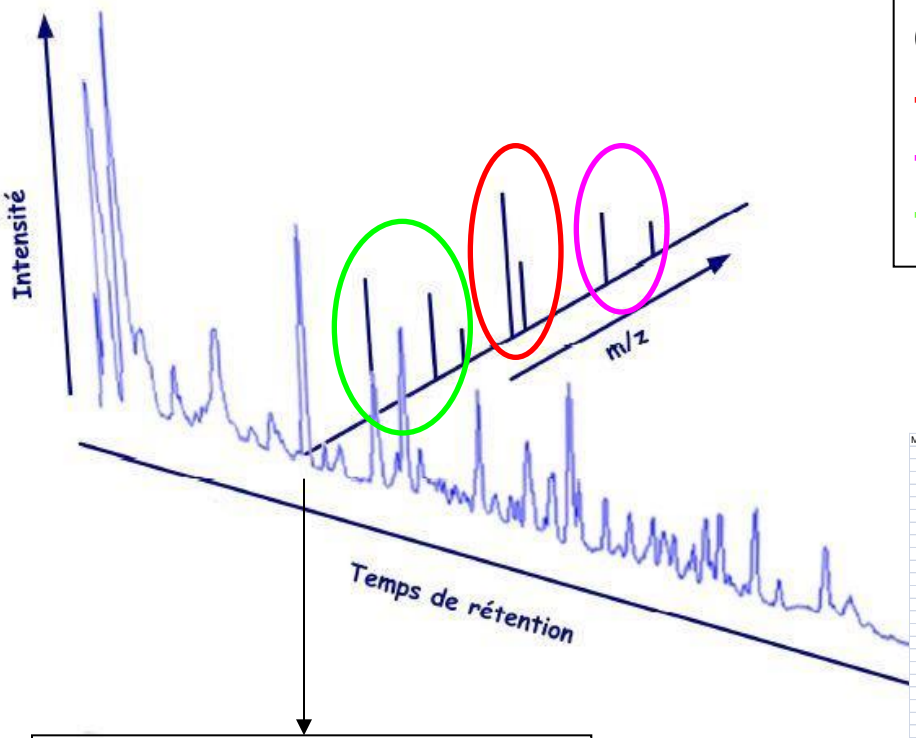
**Structure dataset with multivariable analysis**

**Annotated peaklist**

**Metabolomics relational database:**
- Kind of biofluid
- Physiological Factors
- Disease, toxicology

MetaboHUB
Infrastructure Nationale
en Métabolomique

# Analysis of reference compounds




One molecule = many ions :
- **Pseudo-molecular and isotopes**
- **Adducts**
- **Fragments**

**Annot. FIA***

| M/Z | Intensity | Relative | M/Z (theo) | Delta (ppm) | RDBE | Composition | Attribution | FIA | UPLC (C8) | UPLC (C18) | HSF5 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 72.04431 | 7.7E+04 | 0.82 | 72.04439 | -1.11 | 1.5 | C3 H6 O N | [(M+H)-(C6H10O3)-(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 90.05482 | 3.6E+05 | 3.78 | 90.05496 | -1.55 | 0.5 | C3 H8 O2 N | [(M+H)-(C6H10O3)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 116.034 | 2.7E+04 | 0.29 | 116.03422 | -0.22 | 2.5 | C7 H10 O N | [(M+H)-(C2H6O3)-(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 142.08585 | 1.0E+03 | 0.01 | 142.08626 | -2.85 | 2.5 | C7 H12 O2 N | [(M+H)-(C2H6O3)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 174.11217 | 1.6E+04 | 0.17 | 174.11247 | -0.3 | 1.5 | C8 H16 O3 N | [(M+H)-(HCOOH)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 184.09687 | 2.4E+03 | 0.03 | 184.09682 | 0.26 | 3.5 | C9 H14 O3 N | [(M+H)-2(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 202.10729 | 1.5E+04 | 0.16 | 202.10738 | -0.46 | 2.5 | C9 H16 O4 N | [(M+H)-(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| **220.11799** | **9.69E+06** | **100** | **220.11795** | **0.18** | **1.5** | **C9 H18 N O5** | **[(M+H)]+** | **0** | **1.80** | **4.77** | **11.17** |
| 221.12135 | 6.84E+05 | 7.06 | 221.121305 | 0.2 | 1.5 | C8 13C H18 N O5 | [(M+H)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 222.12166 | 4.59E+04 | 0.47 | 222.122194 | -2.4 | 1.5 | C9 H18 N O4 18O | [(M+H)]+ (18O) | 0 | 1.80 | 4.77 | 11.17 |
| 223.12604 | 2.26E+03 | 0.02 | 223.125549 | 2.2 | 1.5 | C8 13C H18 N O4 18O | [(M+H)]+ (13C+18O) | 0 | 1.80 | 4.77 | 11.17 |
| 238.12923 | 1.76E+03 | 0.02 | 238.128515 | 3 | 0.5 | C9 H20 N O6 | [(M+H)+(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 242.09992 | 6.61E+06 | 68.15 | 242.099895 | 0.1 | 1.5 | C9 H17 N Na O5 | [(M+Na)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 243.10331 | 4.44E+05 | 4.58 | 243.10325 | 0.25 | 1.5 | C8 13C H17 N Na O5 | [(M+Na)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 244.10363 | 1.79E+04 | 0.18 | 244.104139 | -2.09 | 1.5 | C9 H17 N Na O4 18O | [(M+Na)]+ (18O) | 0 | 1.80 | 4.77 | 11.17 |
| 258.07386 | 1.29E+05 | 1.33 | 258.073833 | 0.1 | 1.5 | C9 H17 K N O5 | [(M+K)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 259.07731 | 1.49E+03 | 0.02 | 259.077188 | 0.47 | 1.5 | C8 13C H17 K N O5 | [(M+K)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 264.082 | 1.68E+05 | 1.74 | 264.08184 | 0.61 | 1.5 | C9 H16 N Na2 O5 | [(M-H+2Na)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 265.08498 | 3.08E+03 | 0.03 | 265.085195 | -0.81 | 1.5 | C8 13C H16 N Na2 O5 | [(M-H+2Na)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 310.08654 | 1.10E+03 | 0.01 | 310.08732 | -2.52 | 1.5 | C10 H18 N Na2 O7 | [(M+Na)+(HCOONa)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 461.21113 | 6.45E+05 | 6.65 | 461.210569 | 1.22 | 2.5 | C18 H34 N2 Na O10 | [(2M+Na)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 462.21438 | 1.08E+05 | 1.11 | 462.213924 | 0.99 | 2.5 | C17 13C H34 N2 Na O10 | [(2M+Na)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 463.2174 | 9.76E+03 | 0.1 | 463.217279 | 0.26 | 2.5 | C16 13C2 H34 N2 Na O10 | [(2M+Na)]+ (13C2) | 0 | 1.80 | 4.77 | 11.17 |
| 483.19256 | 4.21E+04 | 0.43 | 483.192514 | 0.1 | 2.5 | C18 H33 N2 Na2 O10 | [(2M-H+2Na)]+ | 0 | 1.80 | 4.77 | 11.17 |

One molecule = one retention time obtained by LC/MS

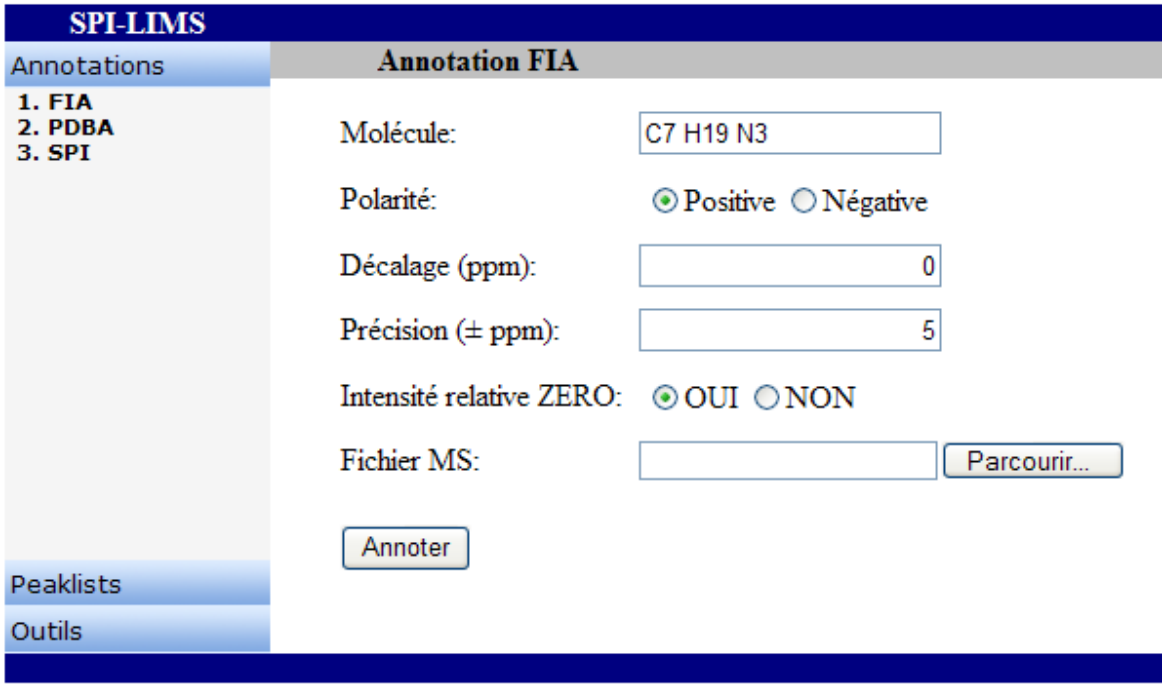
**ESI-mass spectral database**
→ **Annot. SPI***

* Tools of SPI-LIMS web Interface developed at CEA (CEA/DSV/GIPSI)

MetaboHUB
Infrastructure Nationale en Métabolomique

# Reference compounds analysis :

## FIA spectrum annotation




**« Annotation FIA »: to calculate precise *m/z* of potential ions for a given mass formula and compare them to experimental *m/z* :**

- **Isotopes: $^{13}C$, $^{34}S$, $^{18}O$…**
- **Adducts: Na, K, Cl…**
- **Fragments (in source CID)**

# Reference compounds analysis

MetaboHUB
Infrastructure Nationale
en Métabolomique

## Pantothenic acid $[M+H]^+$

Fragments ions and their isotopes

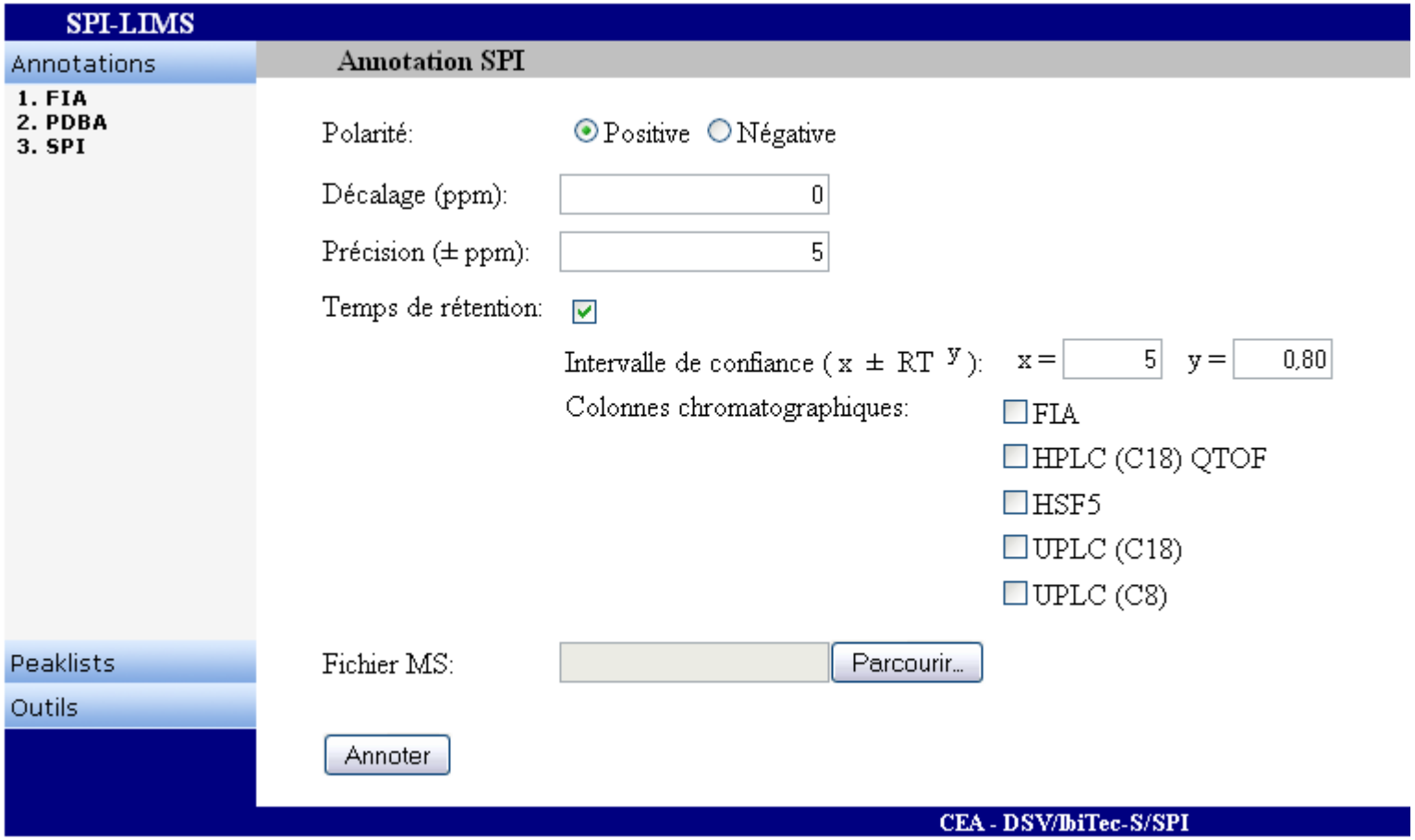
Pseudo-molecular ion and its isotopes

Adducts ions and their isotopes

| M/Z | Intensity | Relative | M/Z (theo) | Delta (ppm) | RDBE | Composition | Attribution | Retention time | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 72.04431 | 7.7E+04 | 0.82 | 72.04439 | -1.11 | 1.5 | C3 H6 O N | [(M+H)-(C6H10O3)-(H2O)]+ | 1.80 | 4.77 | 11.17 |
| 90.05482 | 3.6E+05 | 3.78 | 90.05496 | -1.55 | 0.5 | C3 H8 O2 N | [(M+H)-(C6H10O3)]+ | 1.80 | 4.77 | 11.17 |
| 116.034 | 2.7E+04 | 0.29 | 116.03422 | -0.22 | 2.5 | C7 H10 O N | [(M+H)-(C2H6O3)-(H2O)]+ | 1.80 | 4.77 | 11.17 |
| | | | 26 | -2.85 | 2.5 | C7 H12 O2 N | [(M+H)-(C2H6O3)]+ | 1.80 | 4.77 | 11.17 |
| | | | 47 | -0.3 | 1.5 | C8 H16 O3 N | [(M+H)-(HCOOH)]+ | 1.80 | 4.77 | 11.17 |
| | | | 82 | 0.26 | 3.5 | C9 H14 O3 N | [(M+H)-2(H2O)]+ | 1.80 | 4.77 | 11.17 |
| | | | 38 | -0.46 | 2.5 | C9 H16 O4 N | [(M+H)-(H2O)]+ | 1.80 | 4.77 | 11.17 |
| | | | **95** | **0.18** | **1.5** | **C9 H18 N O5** | **[(M+H)]+** | 1.80 | 4.77 | 11.17 |
| | | | 05 | 0.2 | 1.5 | C8 13C H18 N O5 | [(M+H)]+ (13C) | 1.80 | 4.77 | 11.17 |
| | | | 94 | -2.4 | 1.5 | C9 H18 N O4 18O | [(M+H)]+ (18O) | 1.80 | 4.77 | 11.17 |
| | | | 49 | 2.2 | 1.5 | C8 13C H18 N O4 18O | [(M+H)]+ (13C+18O) | 1.80 | 4.77 | 11.17 |
| 238.12923 | 1.76E+03 | 0.02 | 238.128515 | 3 | 0.5 | C9 H20 N O6 | [(M+H)+(H2O)]+ | 1.80 | 4.77 | 11.17 |
| 242.09992 | 6.61E+06 | 68.15 | 242.099895 | 0.1 | 1.5 | C9 H17 N Na O5 | [(M+Na)]+ | 1.80 | 4.77 | 11.17 |
| 243.10331 | 4.44E+05 | 4.58 | 243.10325 | 0.25 | 1.5 | C8 13C H17 N Na O5 | [(M+Na)]+ (13C) | 1.80 | 4.77 | 11.17 |
| 244.10363 | 1.79E+04 | 0.18 | 244.104139 | -2.09 | 1.5 | C9 H17 N Na O4 18O | [(M+Na)]+ (18O) | 1.80 | 4.77 | 11.17 |
| 258.07386 | 1.29E+05 | 1.33 | 258.073833 | 0.1 | 1.5 | C9 H17 K N O5 | [(M+K)]+ | 1.80 | 4.77 | 11.17 |
| 259.07731 | 1.49E+03 | 0.02 | 259.077188 | 0.47 | 1.5 | C8 13C H17 K N O5 | [(M+K)]+ (13C) | 1.80 | 4.77 | 11.17 |
| 264.082 | 1.68E+05 | 1.74 | 264.08184 | 0.61 | 1.5 | C9 H16 N Na2 O5 | [(M-H+2Na)]+ | 1.80 | 4.77 | 11.17 |
| 265.08498 | 3.08E+03 | 0.03 | 265.085195 | -0.81 | 1.5 | C8 13C H16 N Na2 O5 | [(M-H+2Na)]+ (13C) | 1.80 | 4.77 | 11.17 |
| 310.08654 | 1.10E+03 | 0.01 | 310.08732 | -2.52 | 1.5 | C10 H18 N Na2 O7 | [(M+Na)+(HCOONa)]+ | 1.80 | 4.77 | 11.17 |
| | | | 69 | 1.22 | 2.5 | C18 H34 N2 Na O10 | [(2M+Na)]+ | 1.80 | 4.77 | 11.17 |
| | | | 24 | 0.99 | 2.5 | C17 13C H34 N2 Na O10 | [(2M+Na)]+ (13C) | 1.80 | 4.77 | 11.17 |
| | | | 79 | 0.26 | 2.5 | C16 13C2 H34 N2 Na O10 | [(2M+Na)]+ (13C2) | 1.80 | 4.77 | 11.17 |
| | | | 14 | 0.1 | 2.5 | C18 H33 N2 Na2 O10 | [(2M-H+2Na)]+ | 1.80 | 4.77 | 11.17 |

MetaboHUB
Infrastructure Nationale
en Métabolomique

# Reference compounds analysis :
# Mass spectrum annotation



MetaboHUB
Infrastructure Nationale en Métabolomique

# Reference compounds analysis: annotation using CEA spectral database

## Pantothenic acid $[M+H]^+$

| M/Z | Intensity | Relative | M/Z (theo) | Delta (ppm) | RDBE | Composition | Attribution | FIA | UPLC (C8) | UPLC (C18) | HSF5 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 72.04431 | 7.7E+04 | 0.82 | 72.04439 | -1.11 | 1.5 | C3 H6 O N | [(M+H)-(C6H10O3)-(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 90.05482 | 3.6E+05 | 3.78 | 90.05496 | -1.55 | 0.5 | C3 H8 O2 N | [(M+H)-(C6H10O3)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 116.034 | 2.7E+04 | 0.29 | 116.03422 | -0.22 | 2.5 | C7 H10 O N | [(M+H)-(C2H6O3)-(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 142.08585 | 1.0E+03 | 0.01 | 142.08626 | -2.85 | 2.5 | C7 H12 O2 N | [(M+H)-(C2H6O3)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 174.11217 | 1.6E+04 | 0.17 | 174.11247 | -0.3 | 1.5 | C8 H16 O3 N | [(M+H)-(HCOOH)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 184.09687 | 2.4E+03 | 0.03 | 184.09682 | 0.26 | 3.5 | C9 H14 O3 N | [(M+H)-2(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 202.10729 | 1.5E+04 | 0.16 | 202.10738 | -0.46 | 2.5 | C9 H16 O4 N | [(M+H)-(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| **220.11799** | **9.69E+06** | **100** | **220.11795** | **0.18** | **1.5** | **C9 H18 N O5** | **[(M+H)]+** | **0** | **1.80** | **4.77** | **11.17** |
| 221.12135 | 6.84E+05 | 7.06 | 221.121305 | 0.2 | 1.5 | C8 13C H18 N O5 | [(M+H)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 222.12166 | 4.59E+04 | 0.47 | 222.122194 | -2.4 | 1.5 | C9 H18 N O4 18O | [(M+H)]+ (18O) | 0 | 1.80 | 4.77 | 11.17 |
| 223.12604 | 2.26E+03 | 0.02 | 223.125549 | 2.2 | 1.5 | C8 13C H18 N O4 18O | [(M+H)]+ (13C+18O) | 0 | 1.80 | 4.77 | 11.17 |
| 238.12923 | 1.76E+03 | 0.02 | 238.128515 | 3 | 0.5 | C9 H20 N O6 | [(M+H)+(H2O)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 242.09992 | 6.61E+06 | 68.15 | 242.099895 | 0.1 | 1.5 | C9 H17 N Na O5 | [(M+Na)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 243.10331 | 4.44E+05 | 4.58 | 243.10325 | 0.25 | 1.5 | C8 13C H17 N Na O5 | [(M+Na)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 244.10363 | 1.79E+04 | 0.18 | 244.104139 | -2.09 | 1.5 | C9 H17 N Na O4 18O | [(M+Na)]+ (18O) | 0 | 1.80 | 4.77 | 11.17 |
| 258.07386 | 1.29E+05 | 1.33 | 258.073833 | 0.1 | 1.5 | C9 H17 K N O5 | [(M+K)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 259.07731 | 1.49E+03 | 0.02 | 259.077188 | 0.47 | 1.5 | C8 13C H17 K N O5 | [(M+K)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 264.082 | 1.68E+05 | 1.74 | 264.08184 | 0.61 | 1.5 | C9 H16 N Na2 O5 | [(M-H+2Na)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 265.08498 | 3.08E+03 | 0.03 | 265.085195 | -0.81 | 1.5 | C8 13C H16 N Na2 O5 | [(M-H+2Na)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 310.08654 | 1.10E+03 | 0.01 | 310.08732 | -2.52 | 1.5 | C10 H18 N Na2 O7 | [(M+Na)+(HCOONa)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 461.21113 | 6.45E+05 | 6.65 | 461.210569 | 1.22 | 2.5 | C18 H34 N2 Na O10 | [(2M+Na)]+ | 0 | 1.80 | 4.77 | 11.17 |
| 462.21438 | 1.08E+05 | 1.11 | 462.213924 | 0.99 | 2.5 | C17 13C H34 N2 Na O10 | [(2M+Na)]+ (13C) | 0 | 1.80 | 4.77 | 11.17 |
| 463.2174 | 9.76E+03 | 0.1 | 463.217279 | 0.26 | 2.5 | C16 13C2 H34 N2 Na O10 | [(2M+Na)]+ (13C2) | 0 | 1.80 | 4.77 | 11.17 |
| 483.19256 | 4.21E+04 | 0.43 | 483.192514 | 0.1 | 2.5 | C18 H33 N2 Na2 O10 | [(2M-H+2Na)]+ | 0 | 1.80 | 4.77 | 11.17 |

Human urines

| RT | MZ | # | Masse | Composé | Composition | Attribution | UPLC (C18) |
|---|---|---|---|---|---|---|---|
| 4.72 | 202.107624 | 1 | 202.10738 | pantothenic acid | C9 H16 O4 N | [(M+H)-(H2O)]+ | 4.77 |
| 4.73 | 90.054915 | 1 | 90.05496 | pantothenic acid | C3 H8 O2 N | [(M+H)-(C6H10O3)]+ | 4.77 |
| 4.74 | 220.117542 | 1 | 220.11795 | pantothenic acid | C9 H18 N O5 | [(M+H)]+ | 4.77 |
| 4.74 | 221.12168 | 1 | 221.121305 | pantothenic acid | C8 13C H18 N O5 | [(M+H)]+ (13C) | 4.77 |
| 4.74 | 222.122088 | 1 | 222.122194 | pantothenic acid | C9 H18 N O4 18O | [(M+H)]+ (18O) | 4.77 |
| 4.74 | 242.100637 | 1 | 242.099895 | pantothenic acid | C9 H17 N Na O5 | [(M+Na)]+ | 4.77 |
| 4.74 | 258.075314 | 1 | 258.073833 | pantothenic acid | C9 H17 K N O5 | [(M+K)]+ | 4.77 |

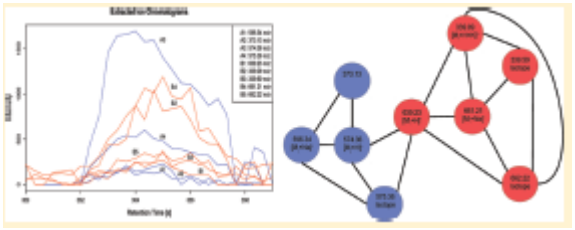
MetaboHUB
Infrastructure Nationale
en Métabolomique

# Tools for annotation and metabolite identification, and data on biofluids begin to be published

analytical chemistry

**CAMERA: An Integrated Strategy for Compound Spectra Extraction and Annotation of Liquid Chromatography/Mass Spectrometry Data Sets**

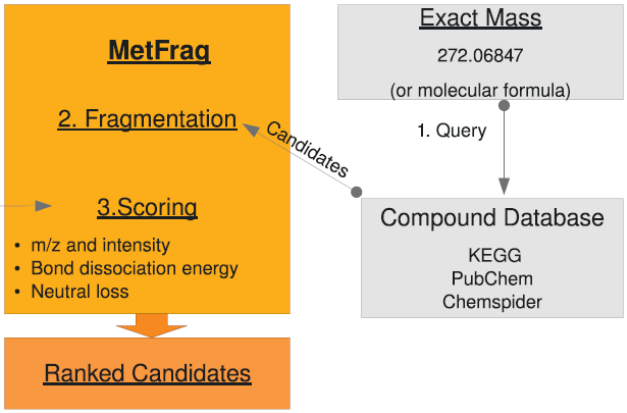
Carsten Kuhl,*,† Ralf Tautenhahn,‡ Christoph Böttcher,† Tony R. Larson,§ and Steffen Neumann*,†



BMC Bioinformatics

**In silico fragmentation for computer assisted identification of metabolite mass spectra**

Sebastian Wolf[1*], Stephan Schmidt[1], Matthias Müller-Hannemann[2], Steffen Neumann[1]




analytical chemistry

**Annotation of the Human Adult Urinary Metabolome and Metabolite Identification Using Ultra High Performance Liquid Chromatography Coupled to a Linear Quadrupole Ion Trap-Orbitrap Mass Spectrometer**

Aurelie Roux,† Ying Xu,† Jean-François Heilier,‡,§ Marie-Françoise Olivier,† Eric Ezan,† Jean-Claude Tabet,⊥ and Christophe Junot*,†

analytical chemistry

**Evaluation of Coupling Reversed Phase, Aqueous Normal Phase, and Hydrophilic Interaction Liquid Chromatography with Orbitrap Mass Spectrometry for Metabolomic Studies of Human Urine**

Tong Zhang,*,† Darren J. Creek,‡,§ Michael P. Barrett,‡ Gavin Blackburn,† and David G. Watson†

Journal of proteome research

**The Footprints of Gut Microbial–Mammalian Co-Metabolism**

Xiaojiao Zheng,†,† Guoxiang Xie,† Aihua Zhao,† Linjing Zhao,† Chun Yao,† Norman H. L. Chiu,§ Zhanxiang Zhou,† Yuqian Bao,‖ Weiping Jia,*,‖ Jeremy K. Nicholson,⊥ and Wei Jia*,†

**HMDB 3.0—The Human Metabolome Database in 2013**

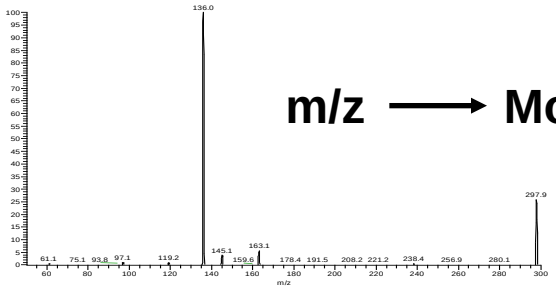
*Nucleic Acids Research, 2013, Vol. 41, Database issue D801–D807*
*doi:10.1093/nar/gks1065*

**Biochemical data on ~ 40000 metabolites**
**Spectral data (RMN, MS)**

MetaboHUB
Infrastructure Nationale
en Métabolomique

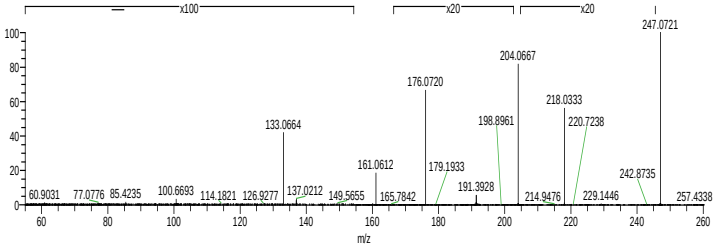
# Formal identification of metabolites often requires several complementary analytical tools

## Mass spectrometry



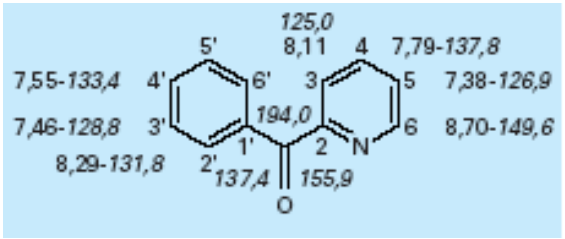

**m/z ⟶ Molecular mass**

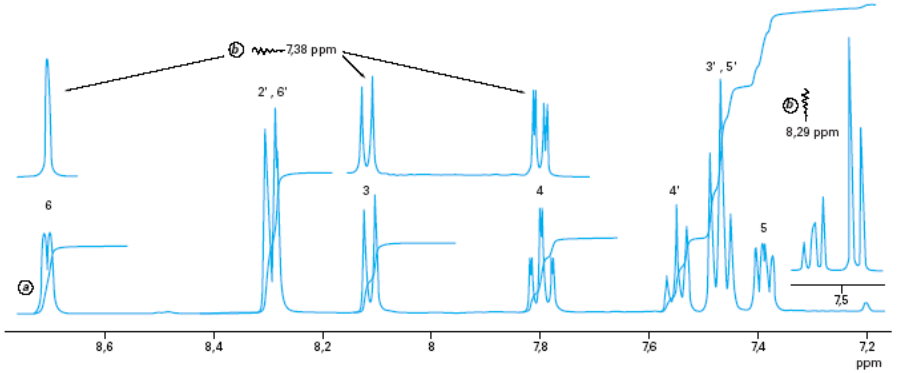
## MS/MS, MSⁿ




**Structural information**

## NMR spectroscopy




**To discriminate between isomers**

MetaboHUB
Infrastructure Nationale
en Métabolomique

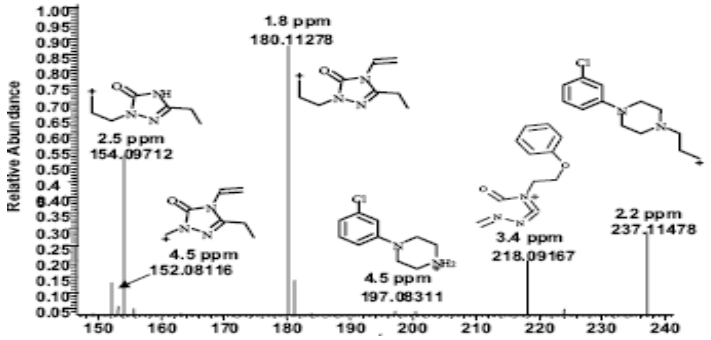
# L'identification des métabolites:

1. Spectre de masse
   - Isoler $[M+H]^+$ ou $[M-H]^-$
   - composition élémentaire ($C_xH_yO_z$)
   - Isotopes ($^{13}C$, $^{18}O$, $^{34}S$…)

→ Annotation

2. Bases de données

3. Spectres de fragmentation (CID)




4. Expériences complémentaires (échange H/D, RMN)

5. Confirmation
   - Synthèse chimique
   - Analyse LC/MS comparative

# Metabolite identification

Metabolomics (2007) 3:211–221
DOI 10.1007/s11306-007-0082-2

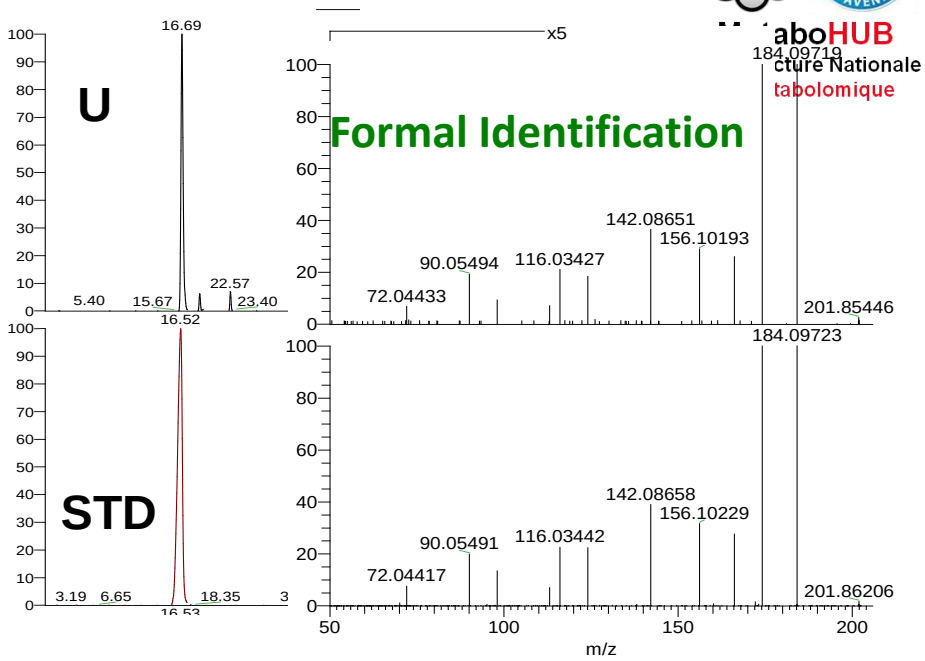
ORIGINAL ARTICLE

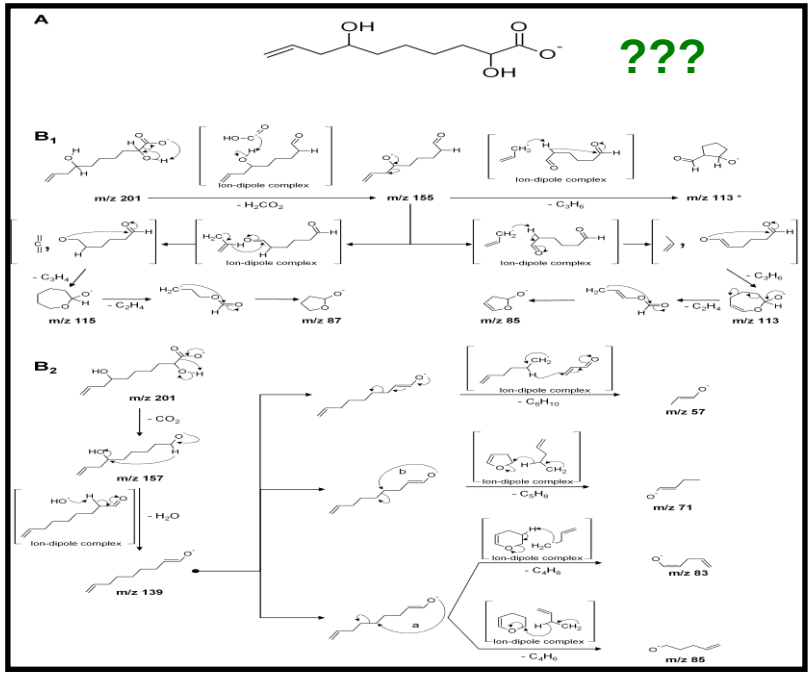
## Proposed minimum reporting standards for chemical analysis

**Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI)**

Lloyd W. Sumner · Alexander Amberg · Dave Barrett · Michael H. Beale · Richard Beger · Clare A. Daykin · Teresa W.-M. Fan · Oliver Fiehn · Royston Goodacre · Julian L. Griffin · Thomas Hankemeier · Nigel Hardy · James Harnly · Richard Higashi · Joachim Kopka · Andrew N. Lane · John C. Lindon · Philip Marriott · Andrew W. Nicholls · Michael D. Reily · John J. Thaden · Mark R. Viant

1. Identified compounds (see below).
2. Putatively annotated compounds (e.g. without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries).
3. Putatively characterized compound classes (e.g. 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.

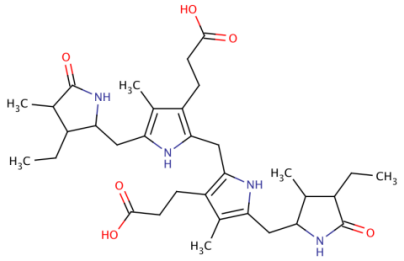


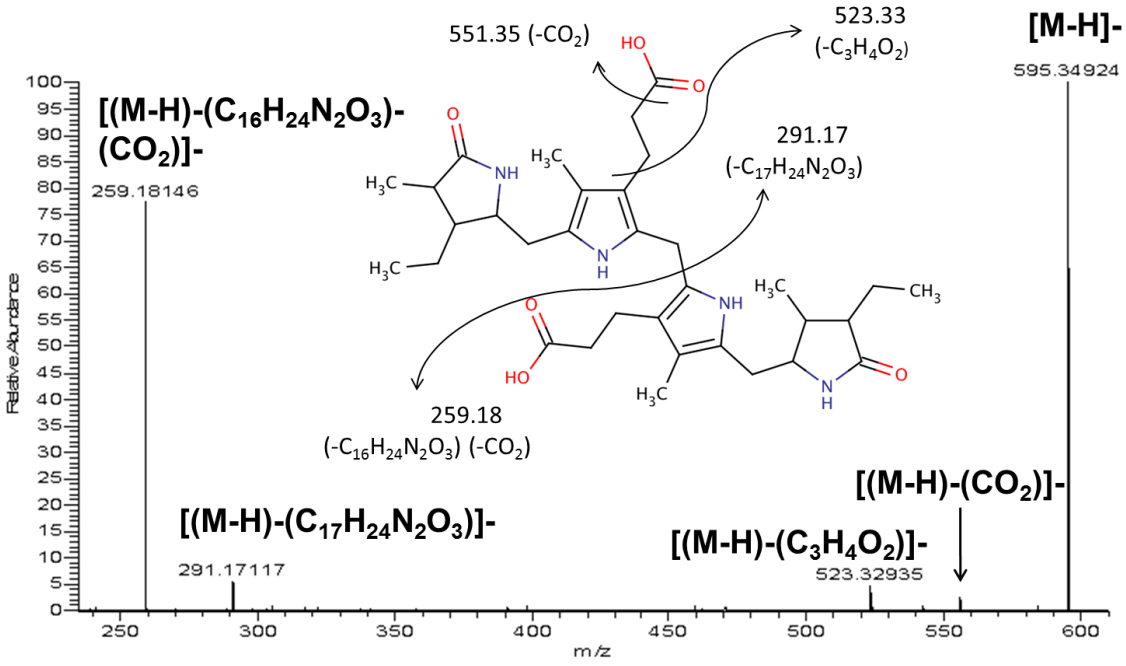



MetaboHUB
Infrastructure Nationale
en Métabolomique

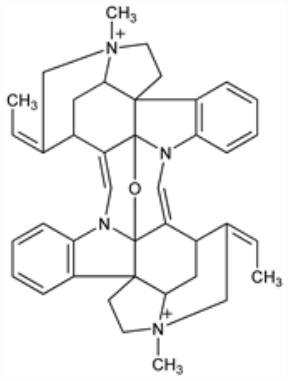
| XCMS output | | | CAMERA output | | | Inter-sample correlation | Public database annotation |
|---|---|---|---|---|---|---|---|
| Variable number | m/z | Retention time (min) | isotopes | adduct | pcgroup | | |
| 1806 | 303.1443 | 9.33 | ** | ** | 531 | NA | ** |
| 4663 | 593.3312 | 9.34 | [681][M]+ | ** | 512 | NA | L-Urobilin |
| 4668 | 594.3368 | 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]+** | ** | **394** | **0.98** | ** |
| 4878 | 631.3256 | 9.40 | ** | **[M+Cl]-** | **394** | **0.96** | ** |
| 3797 | 481.2789 | 9.46 | ** | ** | 552 | NA | GPCho(10:0/4:0) / GPCho(12:0/2:0) |
| 2763 | 381.1910 | 9.53 | ** | ** | 627 | NA | ** |
| 3834 | 485.1792 | 9.61 | ** | ** | 544 | NA | Rutaevin / Nafenopinglucuronide |
| 1255 | 253.1440 | 9.67 | ** | ** | 556 | NA | ** |

## L-Urobilinogen



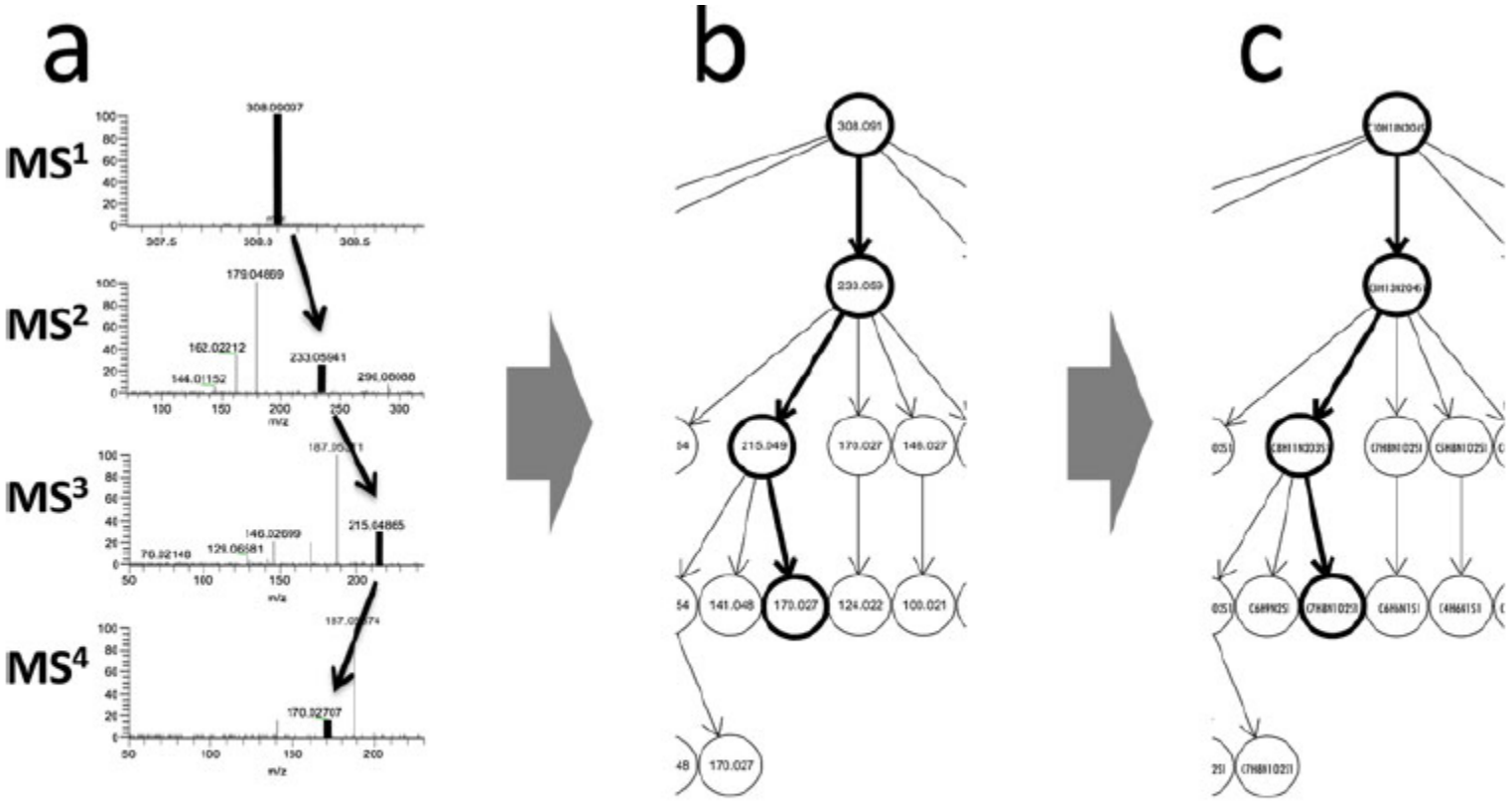



## C-Curarine



MetaboHUB
Infrastructure Nationale
en Métabolomique

# Automated analysis of multistage MS ($MS^n$): Spectral trees




d

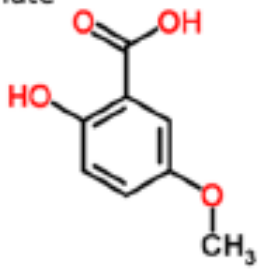
$$C_{10}H_{18}N_3O_6S \Rightarrow C_8H_{13}N_2O_4S \Rightarrow C_8H_{11}N_2O_3S \Rightarrow C_7H_8NO_2S$$

(Kasper PT, Rapid Commun. Mass Spectrom., 2012)

MetaboHUB
Infrastructure Nationale
en Métabolomique

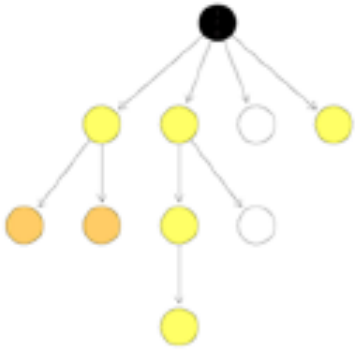
# Automated analysis of multistage MS ($MS^n$): Spectral trees

Id_7: 5-Methoxysalicylate

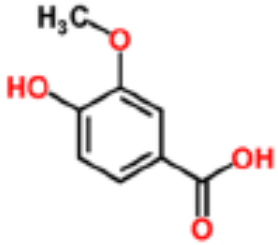



Tree of fragments

Tree of losses

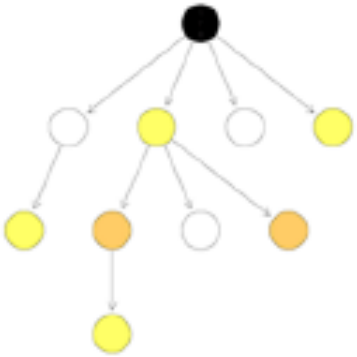


Id_8: Vanillic acid




Tree of fragments

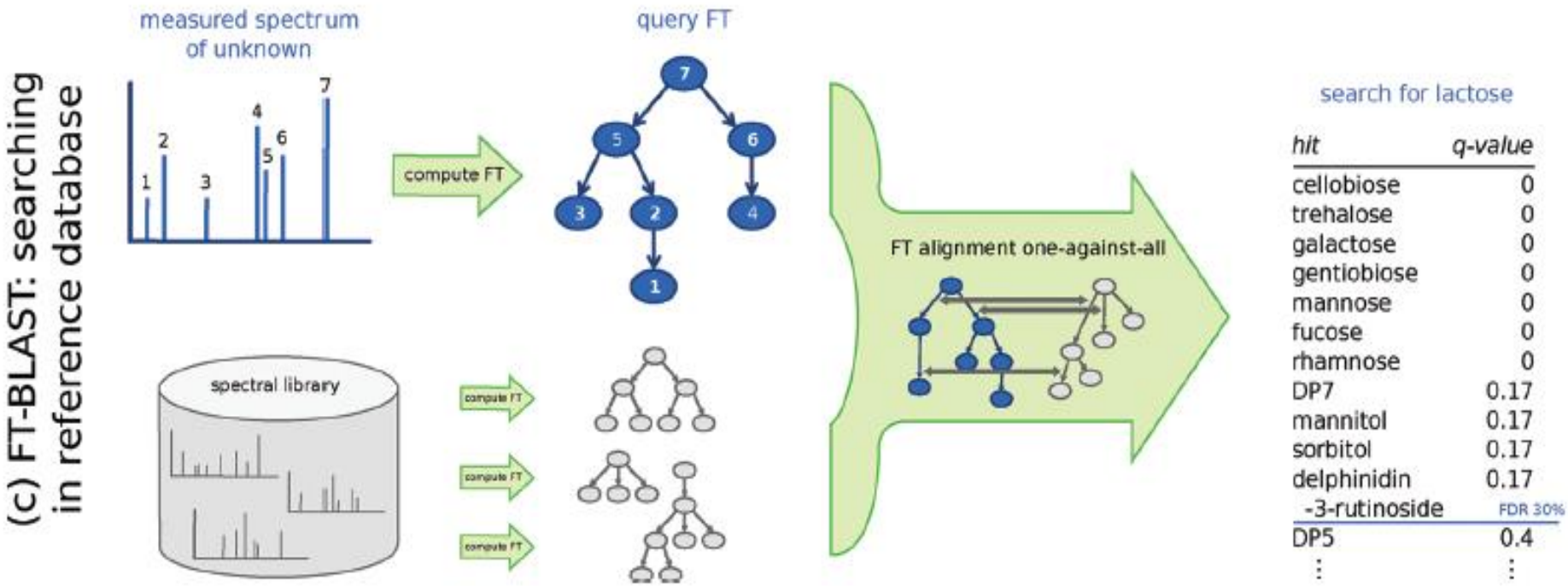
Tree of losses



(Rojas-Cherto M., Anal. Chem., 2012)

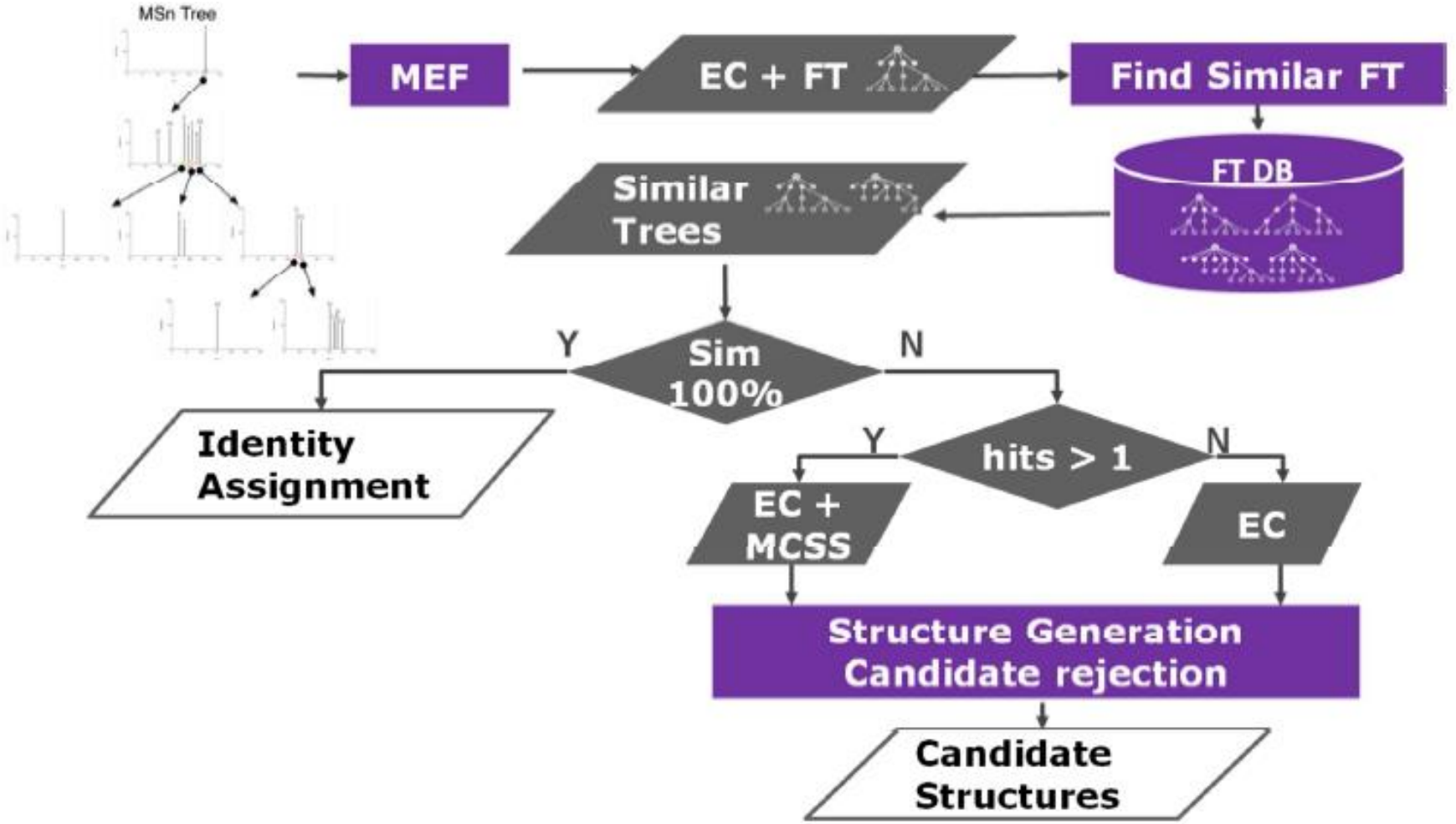
MetaboHUB
Infrastructure Nationale
en Métabolomique

# Automated analysis of multistage MS ($MS^n$): Spectral trees




(Rasche F, Anal. Chem., 2012)

# Automated analysis of multistage MS ($MS^n$): Spectral trees




(Peironcely J, Anal. Chem., 2013)

MetaboHUB
Infrastructure Nationale
en Métabolomique

# *Mise en place d'une méthode d'étalonnage pour la construction d'une base de données MS/MS en science métabolomique*

***F. Ichou**, D. Lesage, C. Junot, J-C. Tabet*

RFMF
Réseau Français de Métabolomique et Fluxomique

IBiSA
Infrastructures Biologie Santé et Agronomie

Travail actuellement poursuivi et coordonné par **R. Cole** dans le cadre de l'infrastructure MetaboHUB

## *Les bases de données*

- Majoritairement en ionisation par électron (E.I)
  - ✓ En E.I : beaucoup de fragmentations, absence parfois de l'ion moléculaire $M^{+\cdot}$ et spectres de masse reproductibles
  - ✓ Nombreuses bases de données et de tailles importantes
  - ✓ Exemple :

| NIST | 220.000 spectres E.I |
|---|---|

- En API : Peu de fragmentations, peu reproductible et interférence de la matrice
  - Conséquence : bases de données MS peu fiables
  - Généralement bases de données MS/MS
    - ✓ Problème de reproductibilité en CID à un régime de basse énergie
    - ✓ 2 approches utilisées

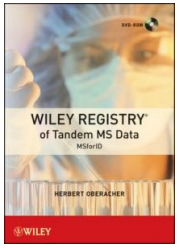
# Les différentes approches de base de données MS/MS:

1 - Approche non-standardisée

2 - Approche standardisée

Principe reposant sur l'enregistrement de multiples empreintes CID

- ***H.Oberacher et M. Pavlic[7]***
  - ✓ Analyses à 10 énergies de collision
  - ✓ Elimination de l'ion parent
    - Grande variation de l'abondance
  - ✓ Amélioration de l'algorithme de matching




*'MSFor ID library'*

✓ Autres exemples :

MassBank **Test** hmp METLIN

[7]H.Oberacher, M.Pavlic. J. Mass Spectrom. 2009, 44, 485

# Les différentes approches de bases de données MS/MS:

1 - Approche non-standardisée

2 - Approche standardisée

Etalonnage avec un composé

- ***P. Marquet et al[8]***
  - ✓ Glafenine
- ***C.Hopley and T. Bristow[9]***
  - ✓ Reserpine

Available online at www.sciencedirect.com
SCIENCE DIRECT
Clinical Biochemistry 38 (2005) 362–372
CLINICAL BIOCHEMISTRY
ELSEVIER

Review

LC-MS/MS systematic toxicological analysis: Comparison of MS/MS spectra obtained with different instruments and settings

Robert Jansen, Gérard Lachatre, Pierre Marquet*

Department of Pharmacology-Toxicology, University Hospital, 2 av. Martin-Luther-King, 87042 Limoges cedex, France

Received 18 April 2004; received in revised form 16 July 2004; accepted 5 November 2004
Available online 21 December 2004

CLINICAL BIOCHEMISTRY

RAPID COMMUNICATIONS IN MASS SPECTROMETRY
*Rapid Commun. Mass Spectrom.* 2008; 22: 1779–1786
Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/rcm.3545

RCM

**Towards a universal product ion mass spectral library – reproducibility of product ion spectra across eleven different mass spectrometers**

**Chris Hopley[1]*, Tony Bristow[2], Anneke Lubben[3], Alec Simpson[4], Elaine Bull[4], Katerina Klagkou[5], Julie Herniman[6] and John Langley[6]**

[1]LGC, Queens Road, Teddington TW11 0LY, UK
[2]Astra Zeneca, Macclesfield SK10 2NA, UK
[3]Department of Chemistry, University of Bath, Bath BA2 7AY, UK
[4]GlaxoSmithKline, Gunnelswood Road, Stevenage SG1 2NY, UK
[5]ThermoElectron, Stafford House, Boundary Way, Hemel Hempstead HP2 7GE, UK
[6]School of Chemistry, University of Southampton, Southampton SO17 1BJ, UK

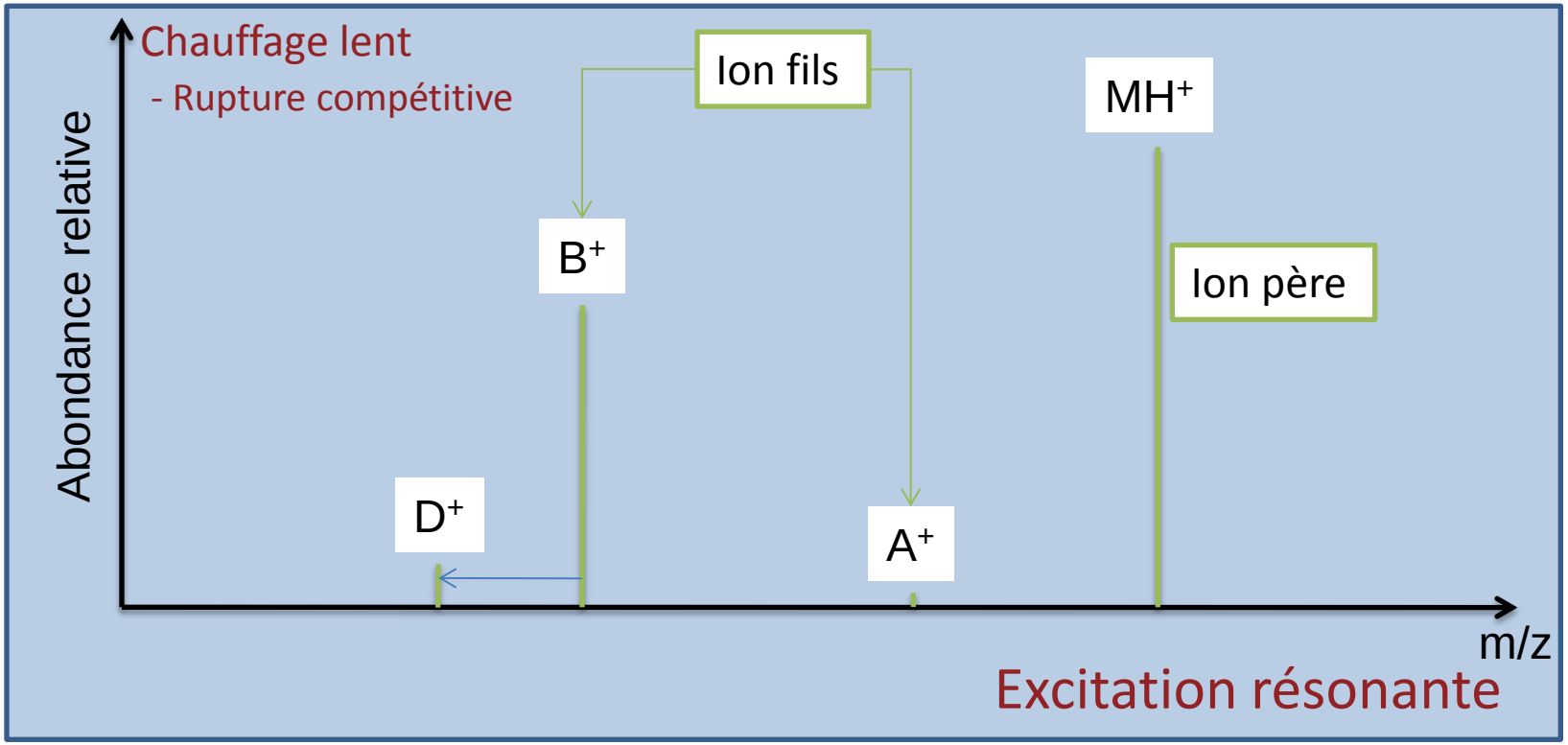
Received 8 October 2007; Revised 20 March 2008; Accepted 20 March 2008

Rapid Communications in Mass Spectrometry

[8]P. Marquet et al, Clin. biochem. 2005, 38, 362
[9]C.Hopley, T. Bristow, Rap. Com. Mass Spectrom. 2008, 22, 1779

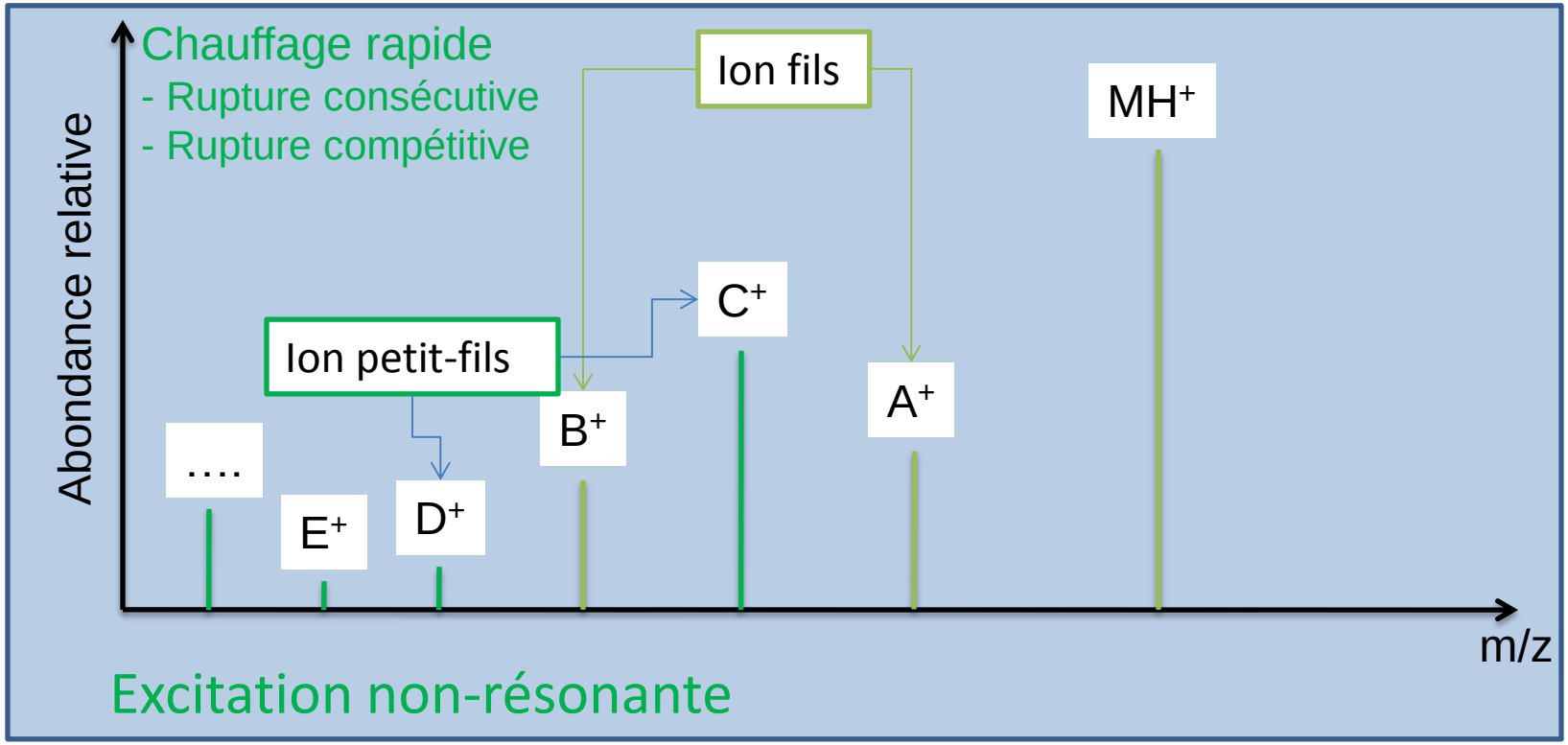
# Type d'excitation

Instruments à piégeage d'ions (Piège LTQ, 3D)



# Type d'excitation

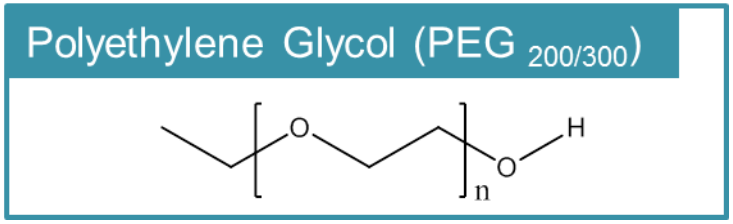
Instruments à faisceau d'ions (TQ, QTOF,…)



# *Procédure d'étalonnage*



- **Ajuster les ratios des abondances relatives des ions produits par la fragmentation du Polyéthylène glycol (PEG) pour avoir des spectres CID comparables à un spectre de référence**




- **Instruments à piégeage d'ions (Résonant) :**
  - ✓Choisir la même énergie de collision normalisée (15%, 20% et 25%)
  - ✓Ajuster le temps d'activation
- **Instruments à faisceau d'ions (Non-résonant) :**
  - ✓ Choisir trois énergies de collision
  - ✓ Ajuster la pression de la cellule de collision

# Objectifs

## Evaluation de la procédure :

### Reproductibilité inter-laboratoire sur le même type d'instruments

✓Est-il possible d'utiliser des spectres de référence enregistrés dans d'autres laboratoires pour identifier les métabolites ?

### Comparaison inter-instrumentale

✓Les spectres CID provenant de différents instruments donnent-ils les mêmes informations ?

### Quelles sont les limites de la méthode d'étalonnage ?

✓Doit-on construire une base de données MS/MS spécifique à chaque type d'instruments (tandems à faisceau d'ions ou à piégeage d'ions)?

MetaboHUB
Infrastructure Nationale
en Métabolomique

# *Validation de la procédure*

- 19 composés :
  - ✓ Panel représentatif de la diversité des problèmes rencontrés en métabolomique
- 9 laboratoires partenaires
- 13 instruments (QTOF, LTQ, Orbitrap)
- Mise en situation avec un logiciel commercial (SmileMS)

MetaboHUB
Infrastructure Nationale
en Métabolomique

# Panel de composés

| Composé | Masse molaire (g/mol) | Polarité | Caractéristique des composés |
|---|---|---|---|
| Caféine | 194 | Pos | **Groupe A.** Peu de voies de dissociations compétitives (< 2 ions fragments dans les pièges à ions) |
| Myricetine | 318 | Pos&Neg | |
| Xanthosine | 284 | Pos&Neg | |
| Inosine | 268 | Pos&Neg | |
| 5-Methoxyindoleacetate | 205 | Pos | |
| Acide Indoleacrylique | 187 | Neg | **Groupe B.** 2-3 ions fragments |
| Acide 12-hydroxydodecanoïque | 216 | Neg | |
| Acide Xanthurenique | 205 | Neg | |
| Dimethoate | 229 | Pos | |
| Naringénine | 272 | Pos | |
| Tyrosine | 181 | Pos&Neg | |
| Panthenol | 205 | Pos&Neg | **Groupe C.** Beaucoup de voies de dissociations (> 3 ions fragments) |
| Acide Cis-aconitique | 174 | Neg | |
| Acide Trans-aconitique | 174 | Neg | |
| Cystathionine | 222 | Pos&Neg | |
| Acetamiprid | 222 | Pos&Neg | |
| Arginine | 174 | Pos&Neg | |
| Glutathion | 307 | Pos&Neg | |
| Carnosine | 226 | Pos&Neg | |

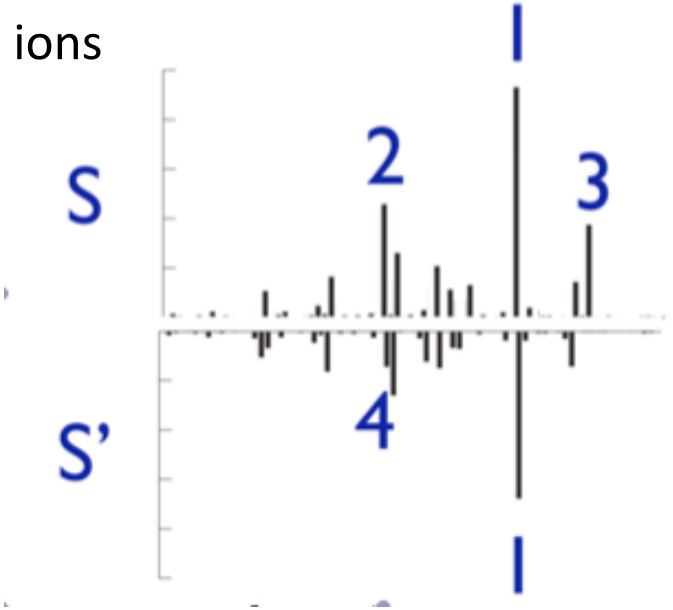
MetaboHUB
Infrastructure Nationale
en Métabolomique

# Logiciel SmileMS

## Originalité de l'algorithme X-Rank

- Comparaison par rapport aux m/z et indirectement aux abondances relatives des ions
- Classification des fragments par ordre d'intensité des pics
- Compatibilité Inter-instrumentale
  - Extraction par protéowizards des différents formats constructeurs

Convertion sous .mzXML, .mgf




Pierre-Alain Binz and co-workers, Anal. Chem. 2009; 81, 7604

MetaboHUB
Infrastructure Nationale
en Métabolomique

# Principe de l'expérience

Comparaison :

- ✓ Inter-QTOF
- ✓ Inter-LTQ
- ✓ Inter-Orbitrap
- ✓ Inter-instrumentale

Deux principaux résultats :

- Score d'identification
  - ✓ Score dépendant de la quantité d'informations contenues dans le spectre CID
  - ✓ Nette baisse du score en négatif sauf pour les molécules ayant beaucoup d'ions fragments
    - Mode négatif plus pauvre en fragments
- Ecart-type

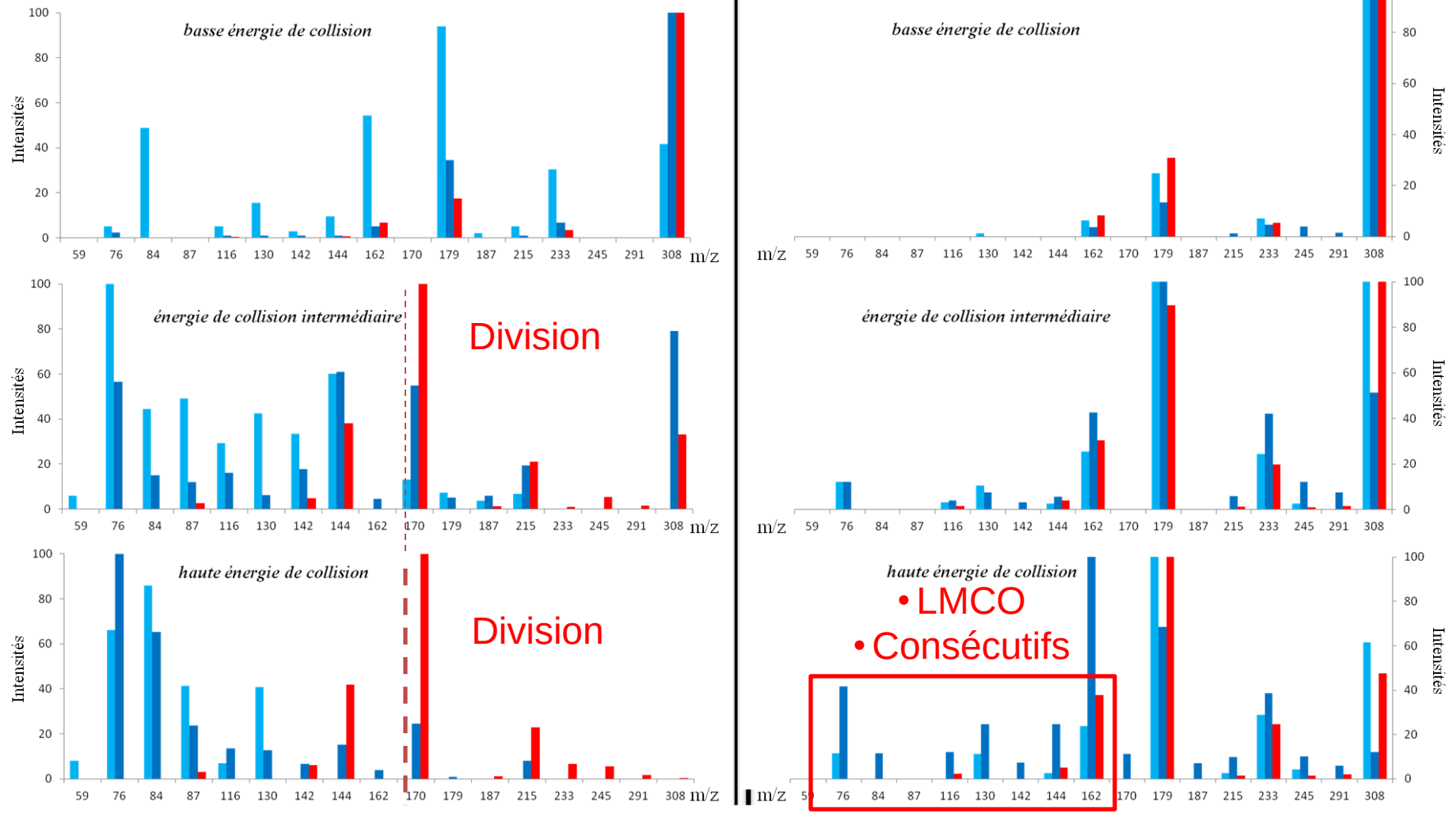
# Résultats inter-instrumentaux

MetaboHUB Infrastructure Nationale en Métabolomique

Glutathion

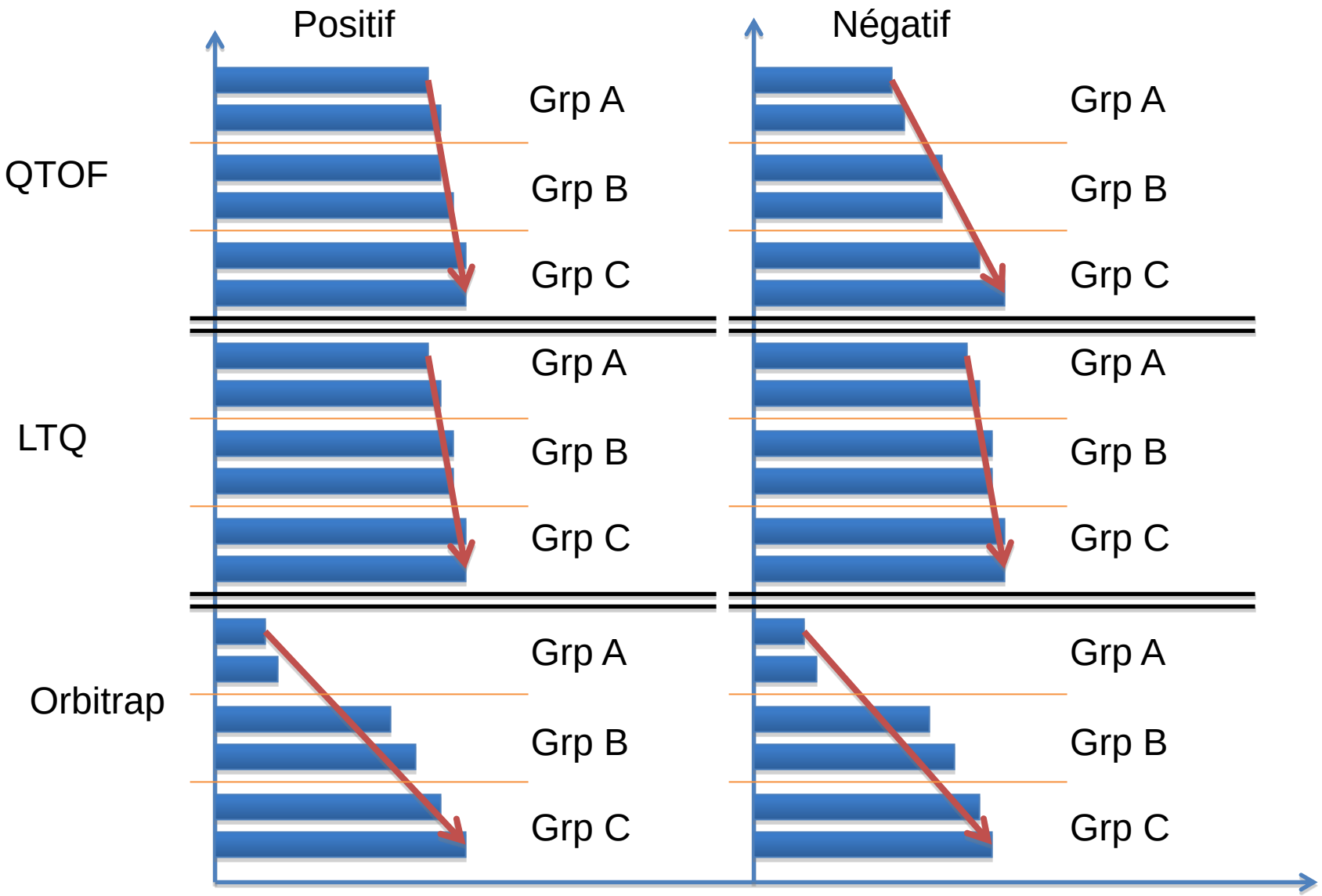
Avant standardisation

Après standardisation



MetaboHUB
Infrastructure Nationale
en Métabolomique

# Quantité d'informations

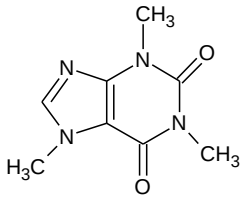


MetaboHUB
Infrastructure Nationale
en Métabolomique

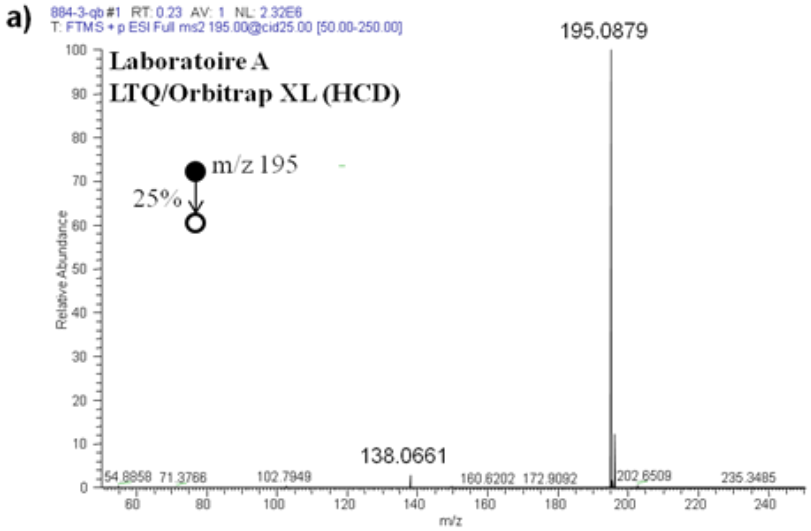
# Résultats inter-Orbitrap

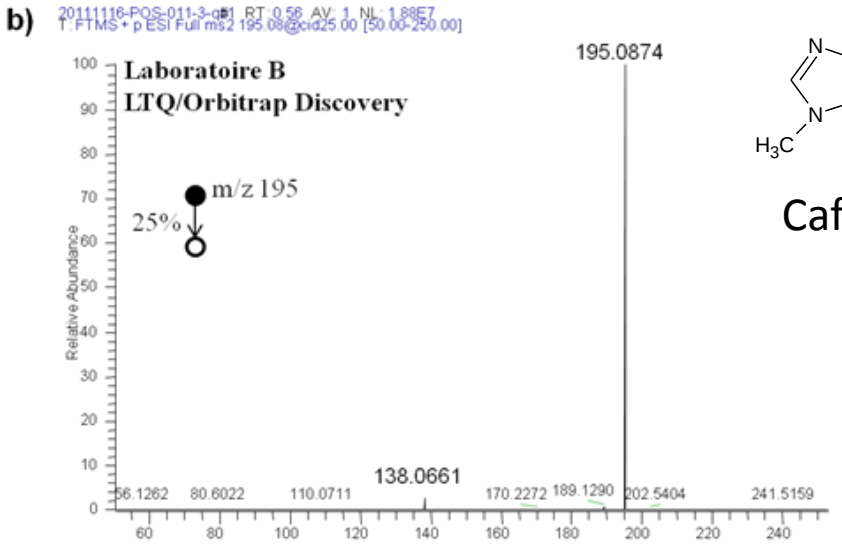
Score moyen d'identification et écart-types proches de 0

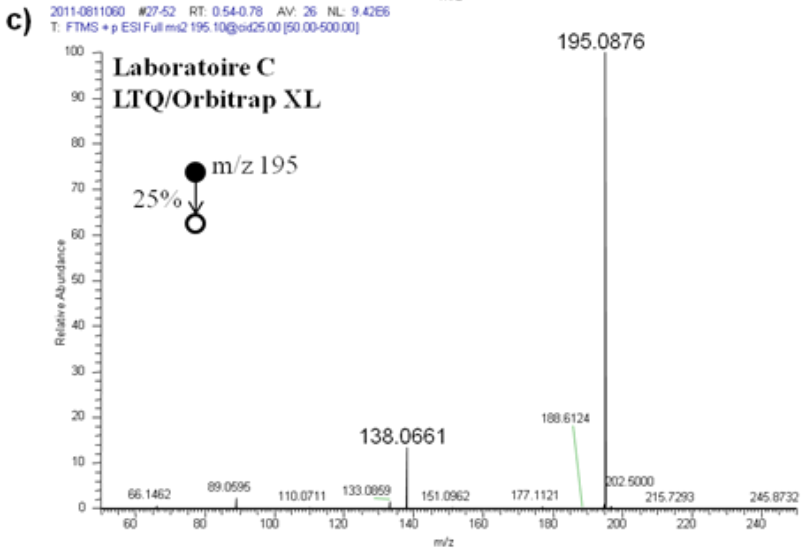
En mode positif

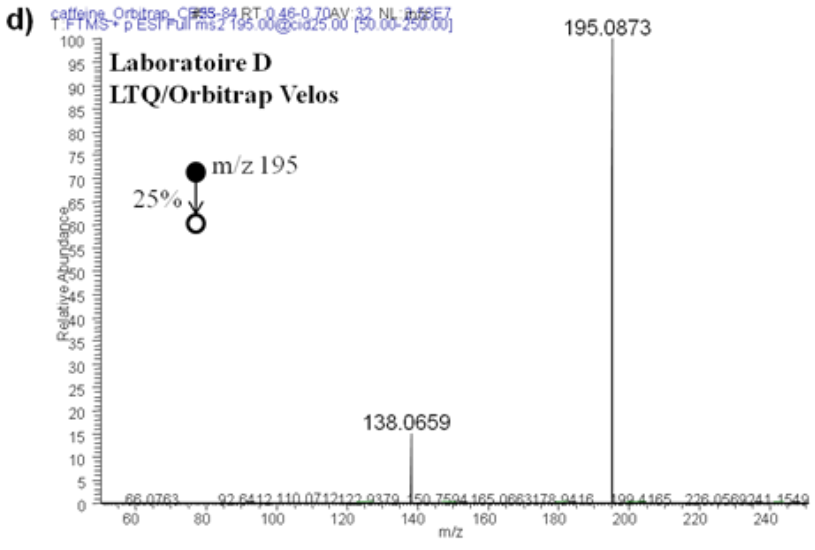



Caféine



MetaboHUB
Infrastructure Nationale
en Métabolomique

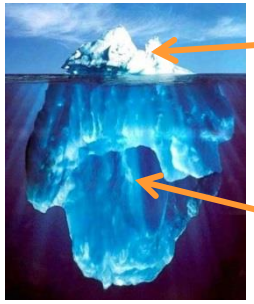
# Conclusion

## *Procédure de standardisation*

- Meilleure reproductibilité inter-laboratoires et inter-instrumentale
- Intérêts de la procédure d'étalonnage
  - ✓ Contrôle de l'énergie interne
  - ✓ Réduction des ruptures consécutives pour les tandems à faisceau d'ions
  - ✓ Chaque instrument peut être une référence
- 2 Sources de déviation dans les scores
  - ✓ Les molécules avec peu de fragmentations
  - ✓ Les tandems à faisceau d'ions (contrôle énergétique plus difficile)

## *Perspectives*

- ✓ Validation et construction de la librairie
- ✓ Projet MétaboHUB



Standards commerciaux

Composés issus de matrices biologiques,…

MetaboHUB
Infrastructure Nationale
en Métabolomique

# Several methods have to implemented in order to achieve an optimal metabolome coverage

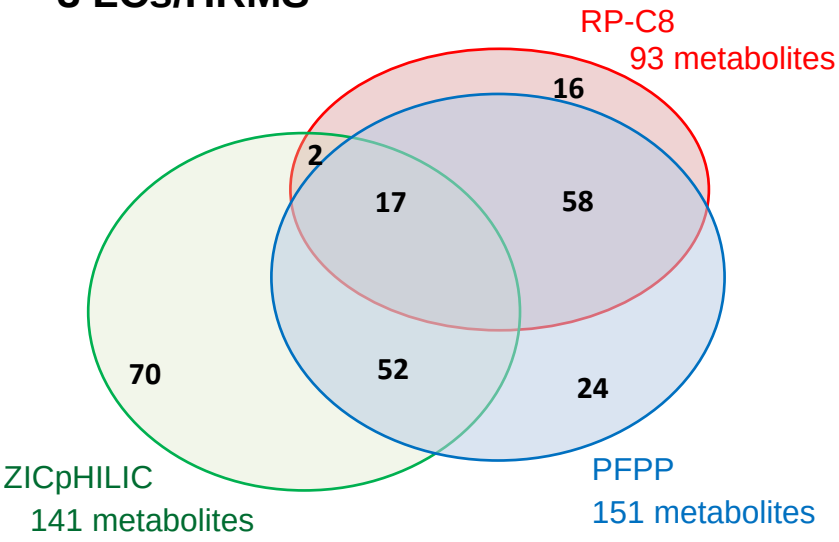
dx.doi.org/10.1021/ac2030738 | *Anal. Chem.* 2012, 84, 1994–2001

## Evaluation of Coupling Reversed Phase, Aqueous Normal Phase, and Hydrophilic Interaction Liquid Chromatography with Orbitrap Mass Spectrometry for Metabolomic Studies of Human Urine

Tong Zhang,*,† Darren J. Creek,‡,§ Michael P. Barrett,‡ Gavin Blackburn,† and David G. Watson†

**270 metabolites identified in human plasma**

**3 LCs/HRMS**

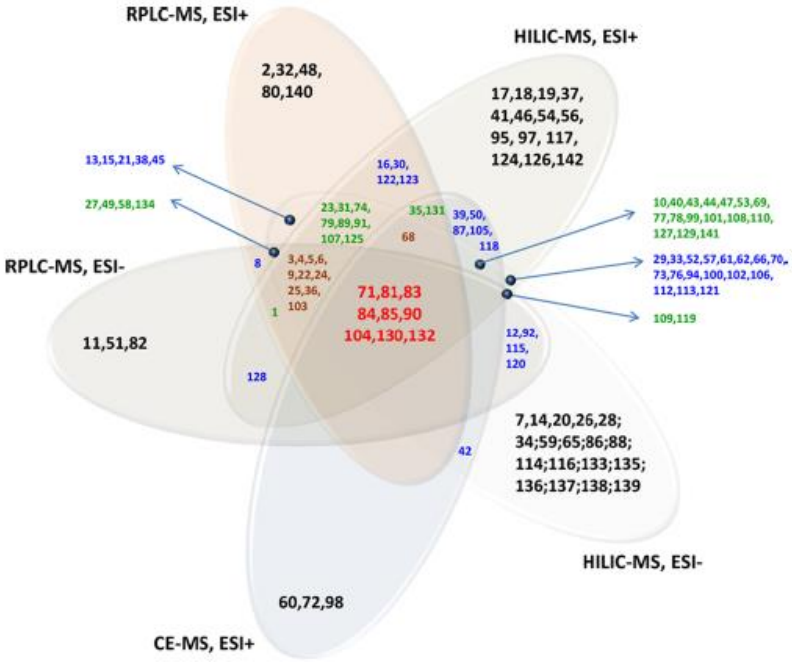



**(Boudah et al., J. Chromatogr. B, 2014)**

dx.doi.org/10.1021/ac300586m | *Anal. Chem.* 2012, 84, 6963–6972

## Systematic Evaluation of Extraction Methods for Multiplatform-Based Metabotyping: Application to the *Fasciola hepatica* Metabolome

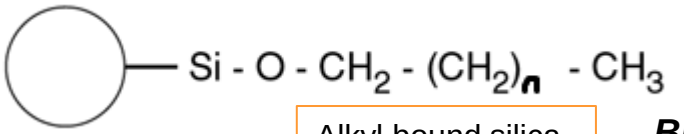
Jasmina Saric,† Elizabeth J. Want,† Urs Duthaler,‡ Matthew Lewis,† Jennifer Keiser,‡ John P. Shockcor,§ Gordon A. Ross,# Jeremy K. Nicholson,† Elaine Holmes,*,† and Marina F. M. Tavares*,†,∥



# The is a need to add other separative dimensions to HRMS

## Reverse Phase Liquid Chromatography

Classic



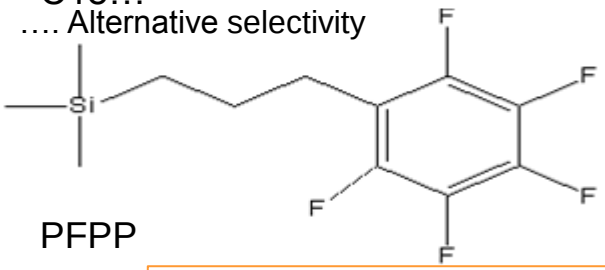

Alkyl bound silica

C8,
C18…

Retention mechanism :
- **Hydrophobic interactions**

***Bile acids, Steroid hormones…***

…. Alternative selectivity




PFPP

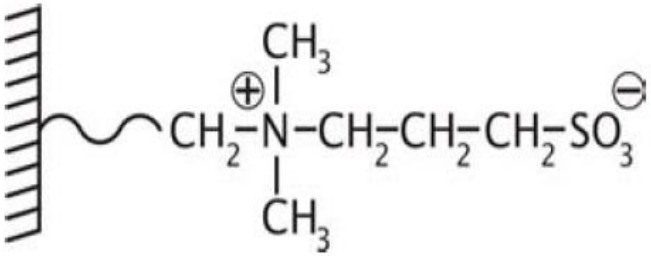
PentaFluoroPhenylPropyl bound silica

Retention mechanism :
-**π–π interactions** for phenyl-based compounds
-**dipole–dipole** interactions for halo/polar compounds

***Amino acids , cyclic compounds…***

Most widely used columns:
-aqueous and organic solvents compatibility
-robustness
-well defined retention mechnisms
-continuous manufacturer improvement

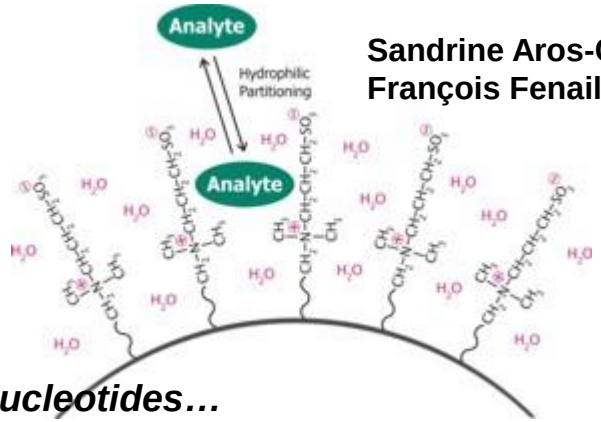
## Hydrophilic Interaction Liquid Chromatography




ZIC HILIC

Zwitterionic sulfobetaine bound silica

Retention mechanism :
-**hydrophilic partitioning** from the eluent to the enriched-water layer
-**electrostatic interactions** with either positive and negative charges

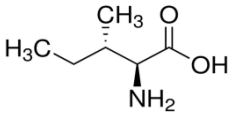



***Carboxylic acids, Sugars, Nucleotides…***

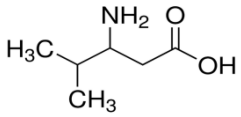
**Sandrine Aros-Calt**
**François Fenaille**

MetaboHUB
Infrastructure Nationale
en Métabolomique

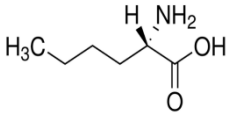
**To discriminate between isomer species**



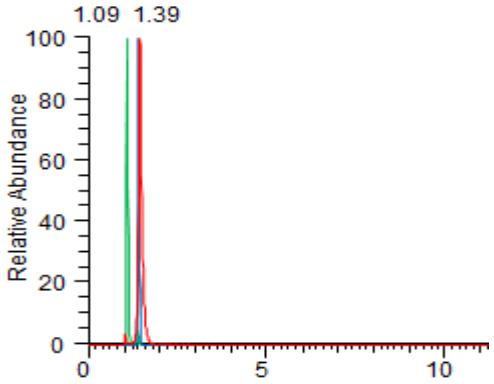

Isoleucine



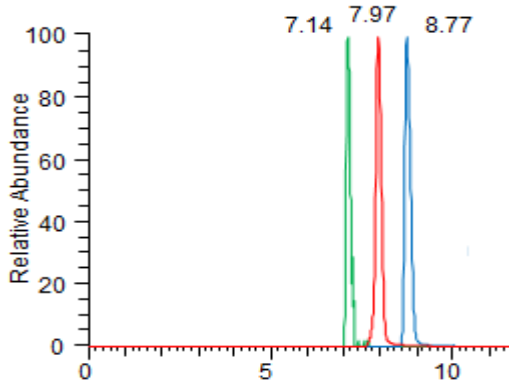

β Leucine



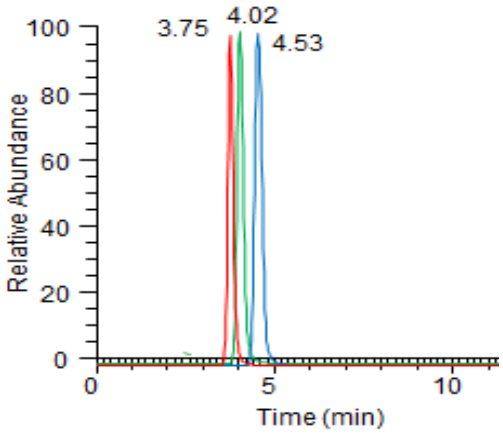

Norleucine




RP-C8

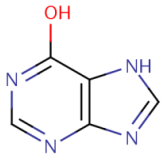


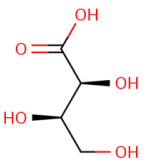

PFPP



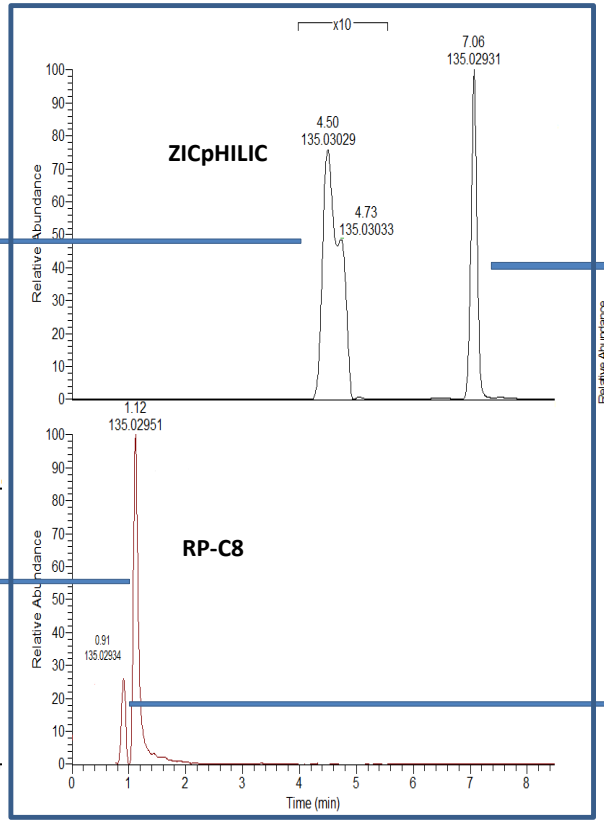

ZICpHILIC

MetaboHUB
Infrastructure Nationale en Métabolomique

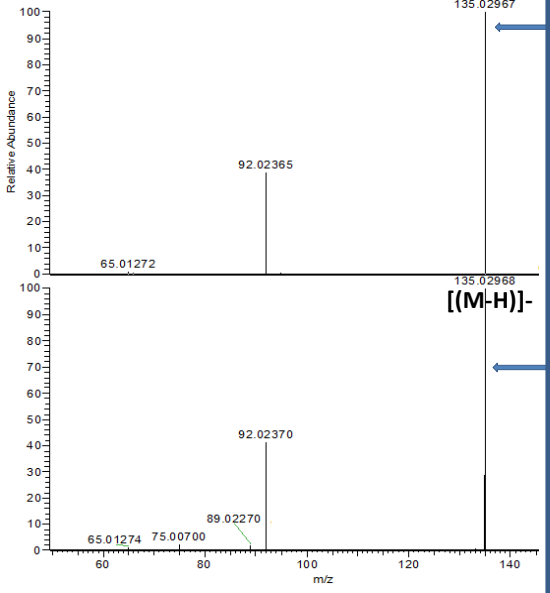




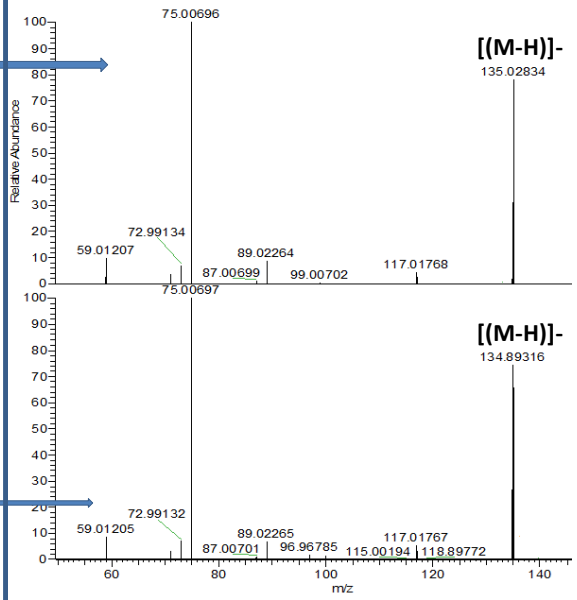
## LC/MS analyses of human plasma



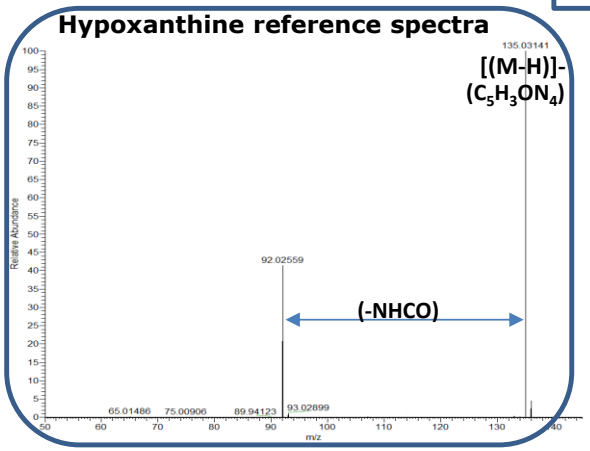

## LC/MS/MS analyses



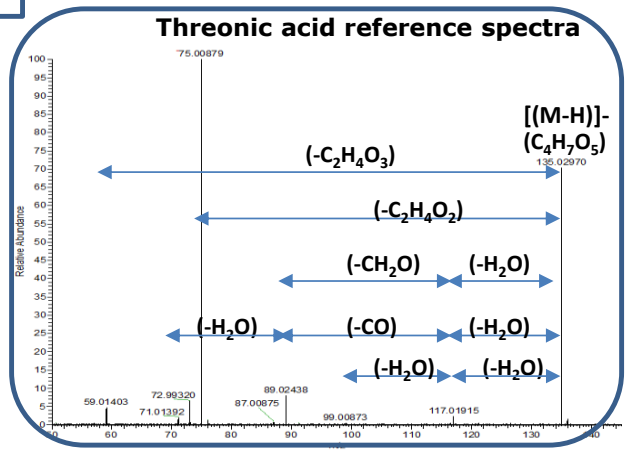

## LC/MS/MS analyses




## Hypoxanthine reference spectra

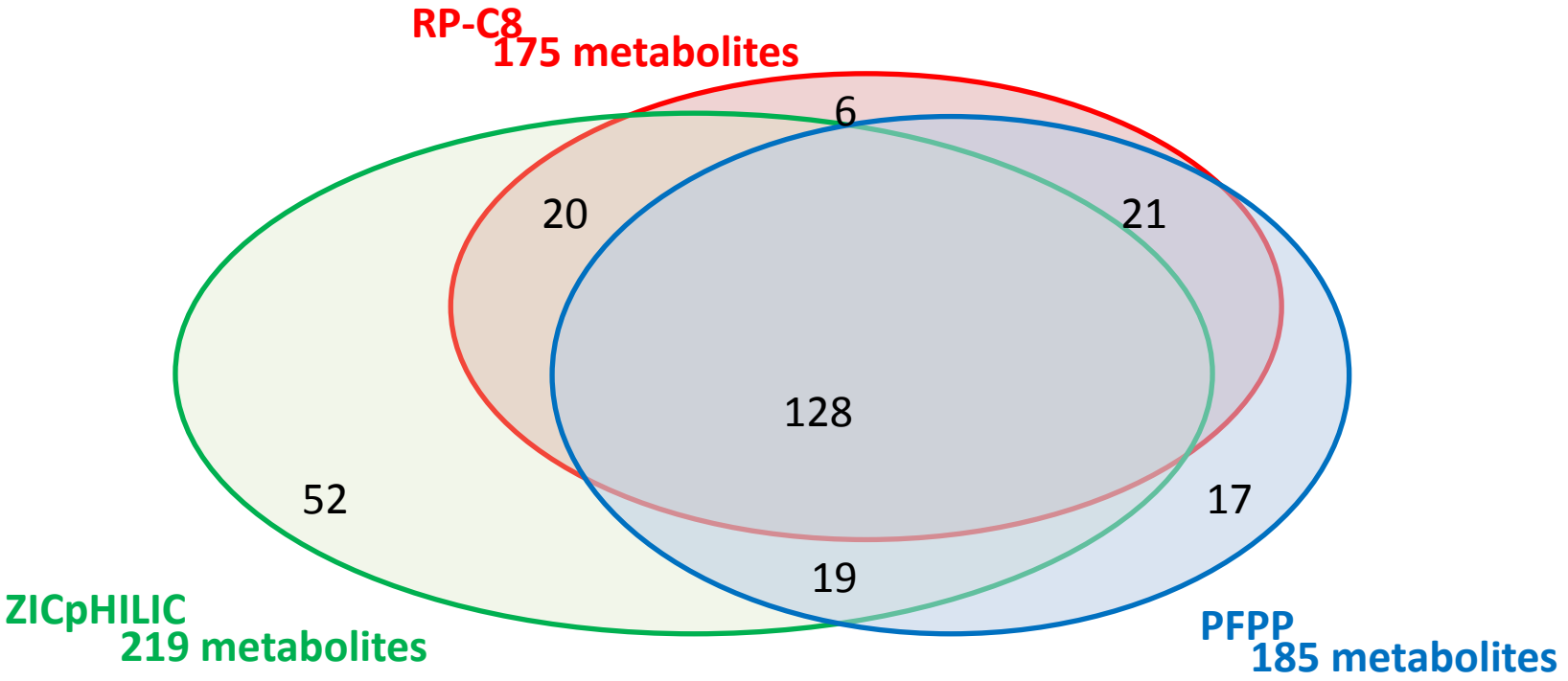



## Threonic acid reference spectra



MetaboHUB
Infrastructure Nationale
en Métabolomique

# Annotation of the human serum metabolome using LC/HRMS




-**266** metabolites were distributed over **209** distinct accurate masses

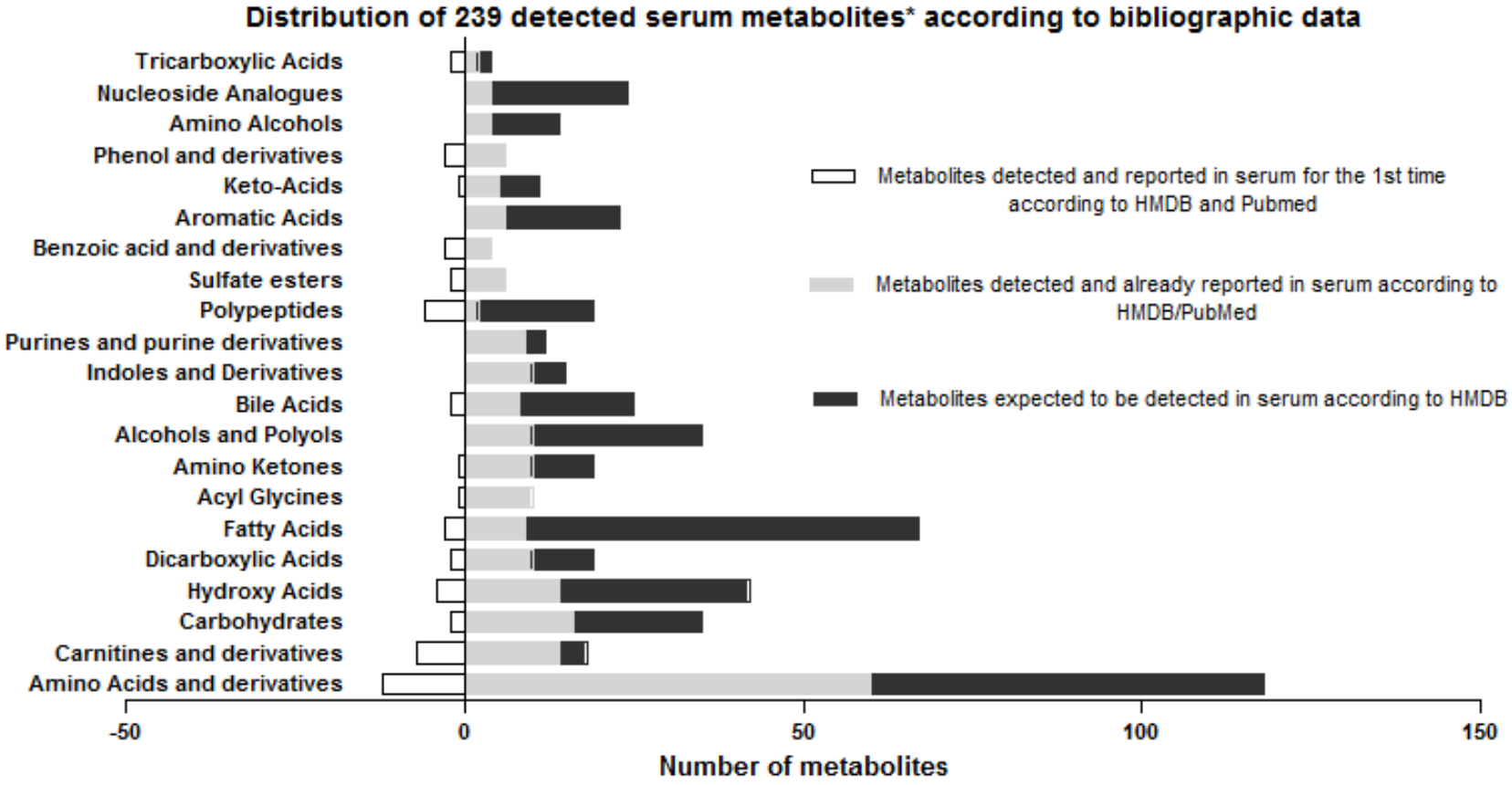
-ESI(-): HILIC extended metabolome coverage 50% and the number of retained metabolites (*k*>1) 75%.
-ESI(+): PFPP conditions improved the retention up to almost 200% of metabolites detected.

-using HILIC (ESI-) and PFPP (ESI+) ensure the detection of 243 out of 266 metabolites in serum samples.

**(Boudah et al., J. Chromatogr. B, 2014)**

# Annotation of the human serum metabolome using LC/HRMS

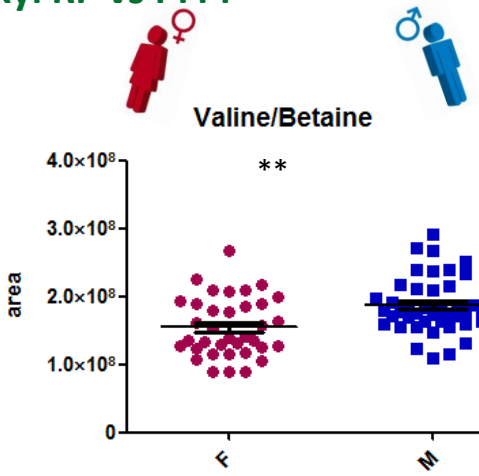
MetaboHUB
Infrastructure Nationale
en Métabolomique

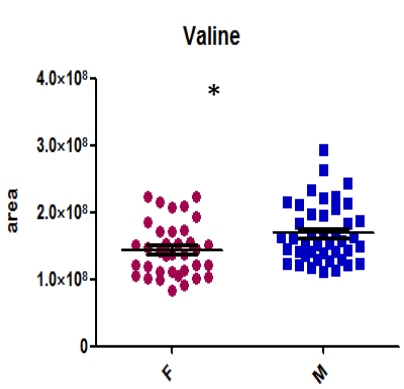


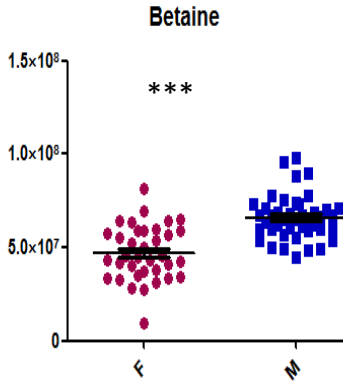
MetaboHUB
Infrastructure Nationale en Métabolomique

# Multiplatform strategy: toward a comprehensive assessment of metabolomic profiles

## Alkyl RP vs PFPP

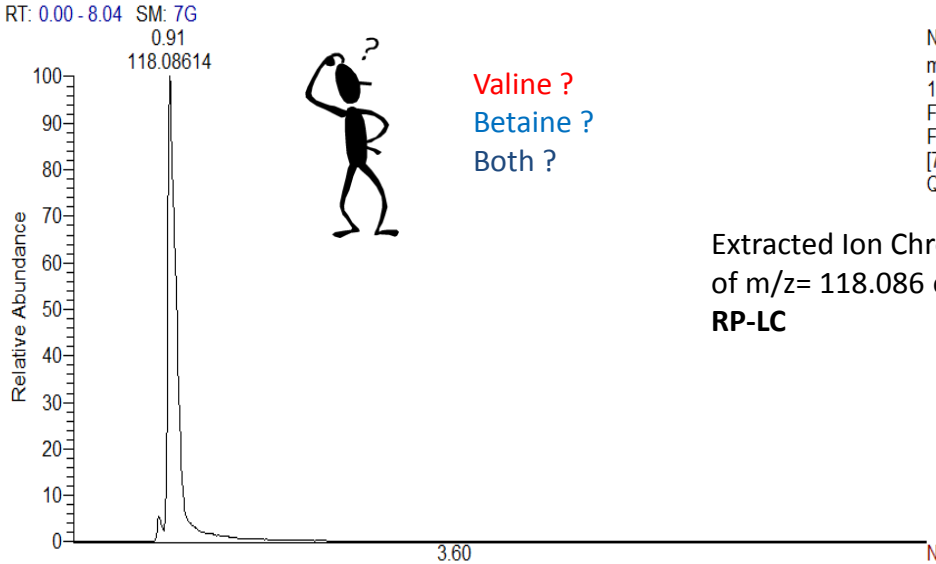


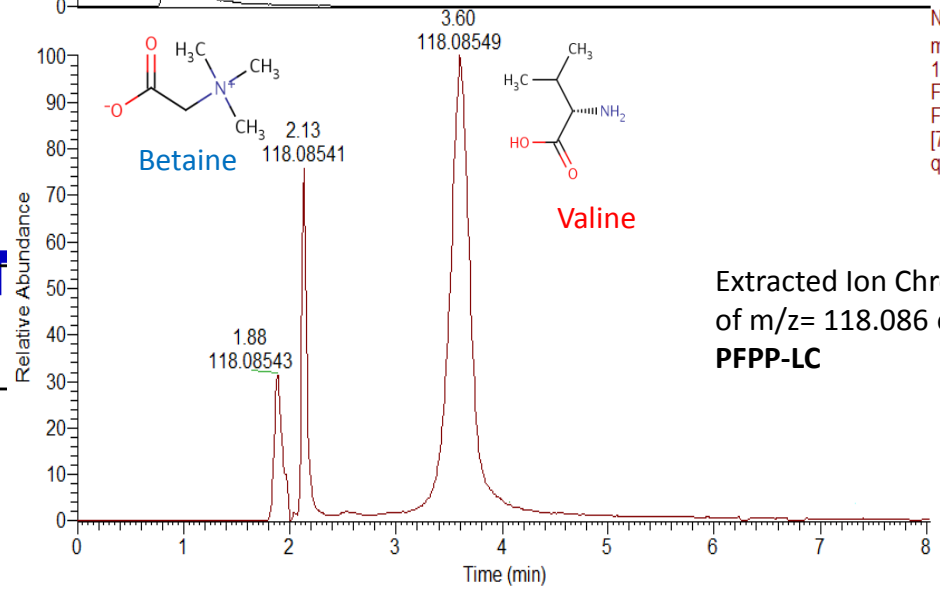




Proteinogenic amino acids

Detoxification reaction : liver, kidney




Extracted Ion Chromatogram of m/z= 118.086 obtained in **RP-LC**




Extracted Ion Chromatogram of m/z= 118.086 obtained in **PFPP-LC**

MetaboHUB
Infrastructure Nationale
en Métabolomique

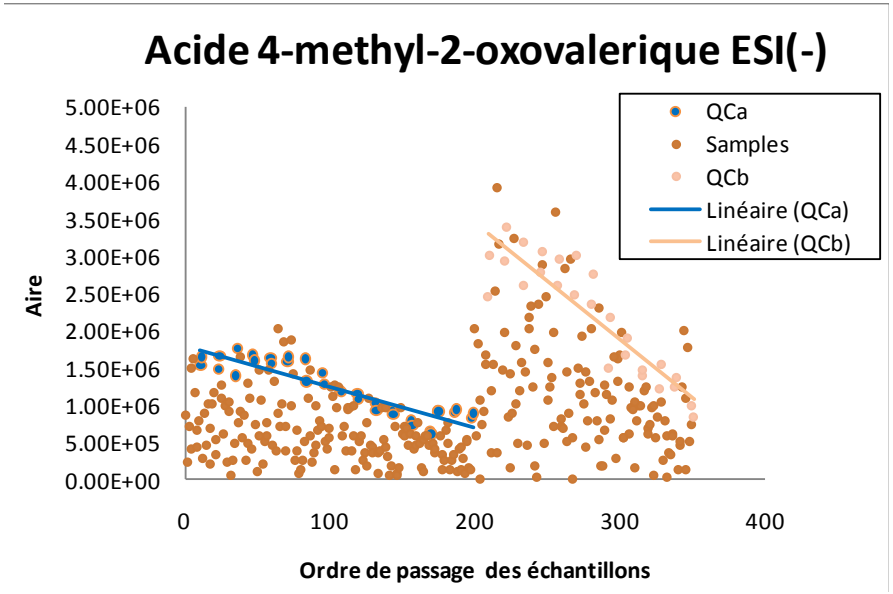
# Toward databases of metabolic profiles: it is required to control analytical biais through design of experiment

- Randomization is required
- Batches of ~ 100 injections
- Blanck and QC samples

**Some «reference» protocols are available:**

**Human plasma: Dunn WB, Nat. Protocols, 2011**

**Human urines: Want E., Nat. Protocols, 2010**




*Pool d'échantillons bio préparés à 3 dilutions et injectés en n=3 en vue du traitement des données*

QC-pool + EI_dil-8
QC-pool + EI_dil-8
QC-pool + EI_dil-8

QC-pool + EI_dil-4
QC-pool + EI_dil-4
QC-pool + EI_dil-4

QC-pool + EI_dil-2
QC-pool + EI_dil-2
QC-pool + EI_dil-2

blanc
QC-pool
QC-pool + EI

*10 échantillons bio*

Ech bio
Ech bio
Ech bio
...
Ech bio
Ech bio

blanc
QC-pool
QC-pool + EI

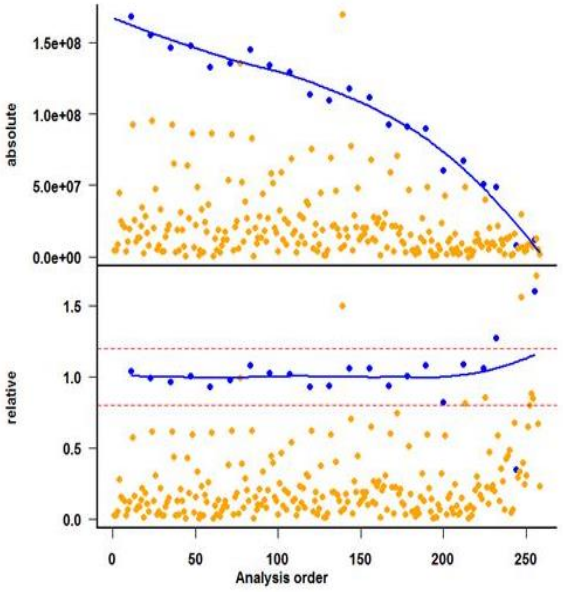
MetaboHUB
Infrastructure Nationale en Métabolomique

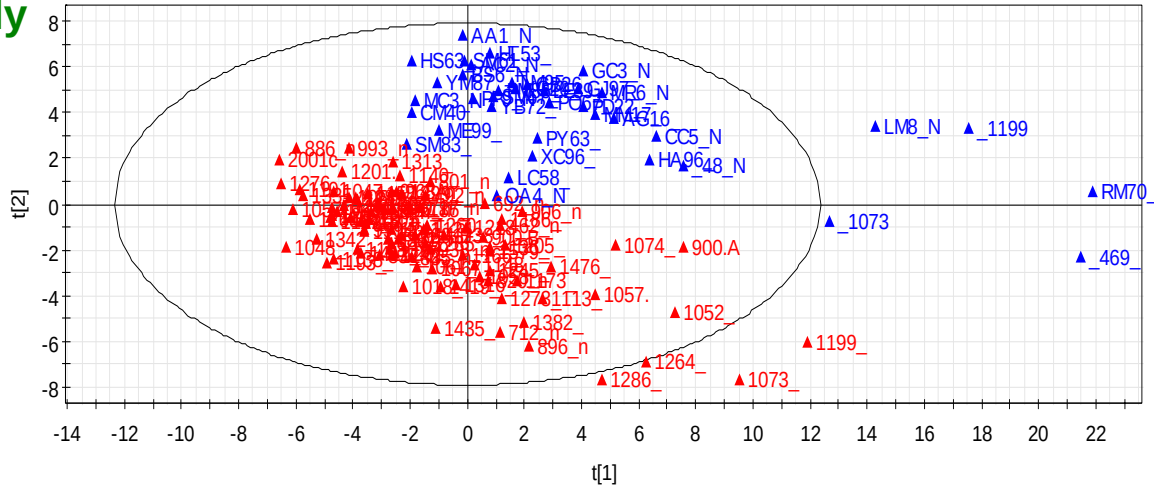
# Toward databases of metabolic profiles: it is required to control analytical biais through design of experiment

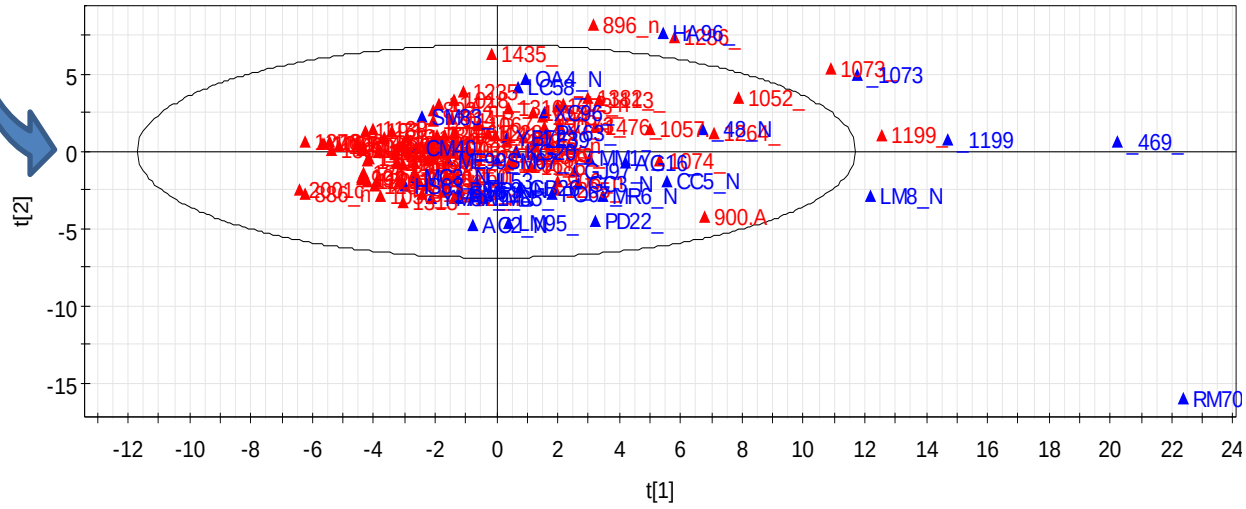
**LOESS: Low Order non linear locally Estimated Smoothing Function**

**Cleveland, J. Am. Stat. Assoc. 1979**

**Dunn WB, Nat. Protoc. 2011**



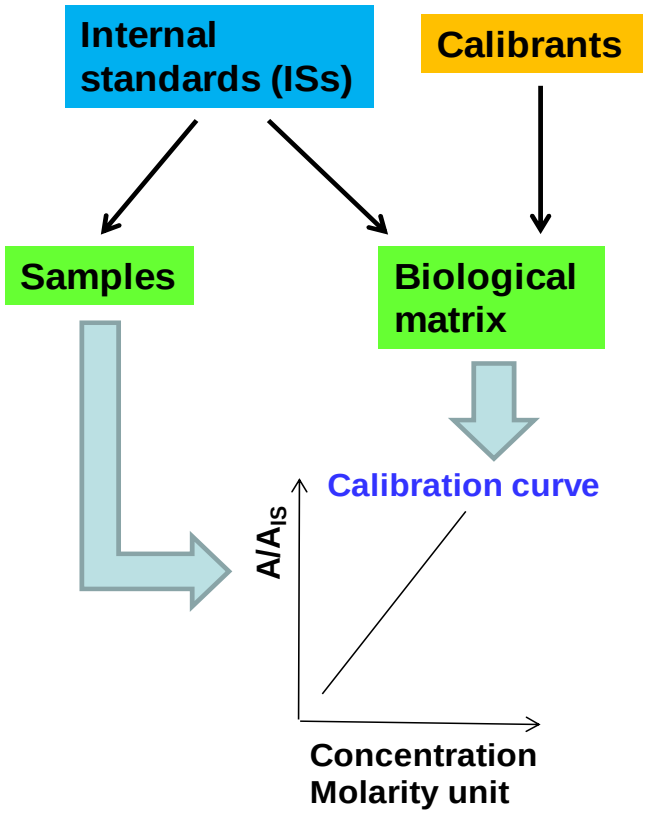



MetaboHUB
Infrastructure Nationale en Métabolomique

# Quantitative metabolite profiling for large scale studies

## Absolute quantification of metabolites




PLoS GENETICS

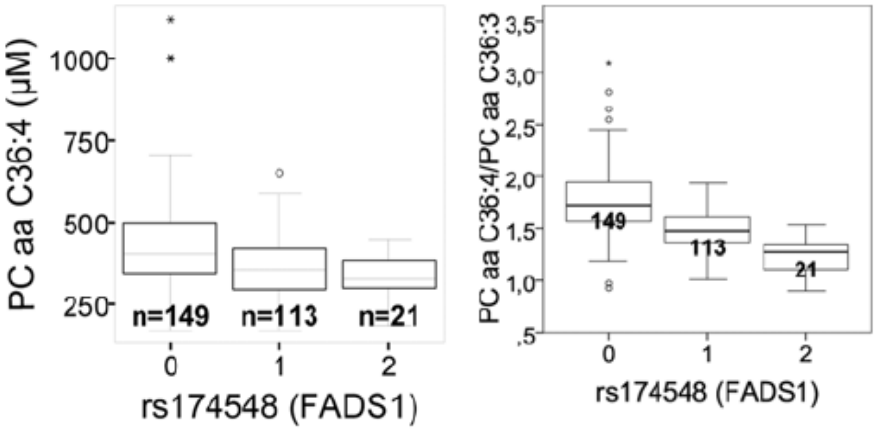
**Genetics Meets Metabolomics: A Genome-Wide Association Study of Metabolite Profiles in Human Serum**

Christian Gieger[1,2], Ludwig Geistlinger[1], Elisabeth Altmaier[3,4], Martin Hrabé de Angelis[5,6], Florian Kronenberg[7], Thomas Meitinger[8,9], Hans-Werner Mewes[3,10], H.-Erich Wichmann[1,2], Klaus M. Weinberger[11], Jerzy Adamski[5,6], Thomas Illig[1], Karsten Suhre[3,4]*

**Quantitative measurement of 363 metabolites in 284 serum samples**

**«Genetically determined metabotypes»**

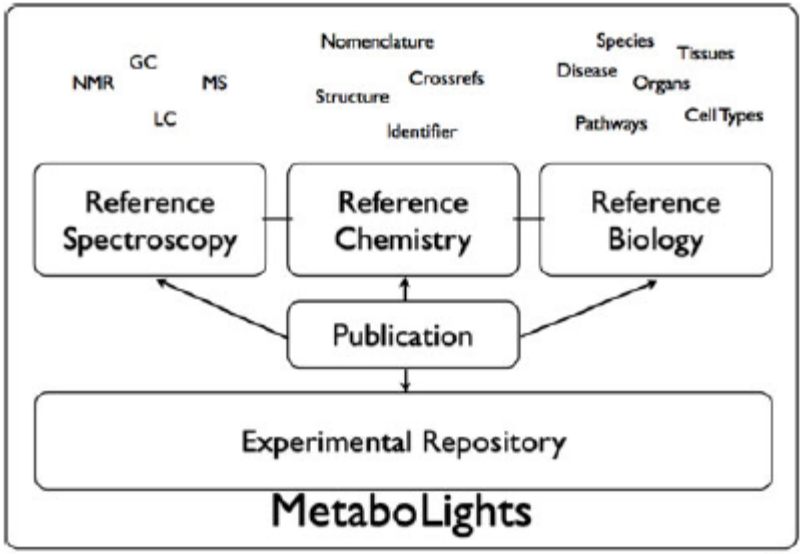
**Polymorphism in the *FADS1* (fatty acid delta 5 desaturase) gene**



Metabolomics (2012) 8:757–760
DOI 10.1007/s11306-012-0462-0

# MetaboLights: towards a new COSMOS of metabolomics data management

Christoph Steinbeck · Pablo Conesa · Kenneth Haug · Tejasvi Mahendraker · Mark Williams · Eamonn Maguire · Philippe Rocca-Serra · Susanna-Assunta Sansone · Reza M. Salek · Julian L. Griffin



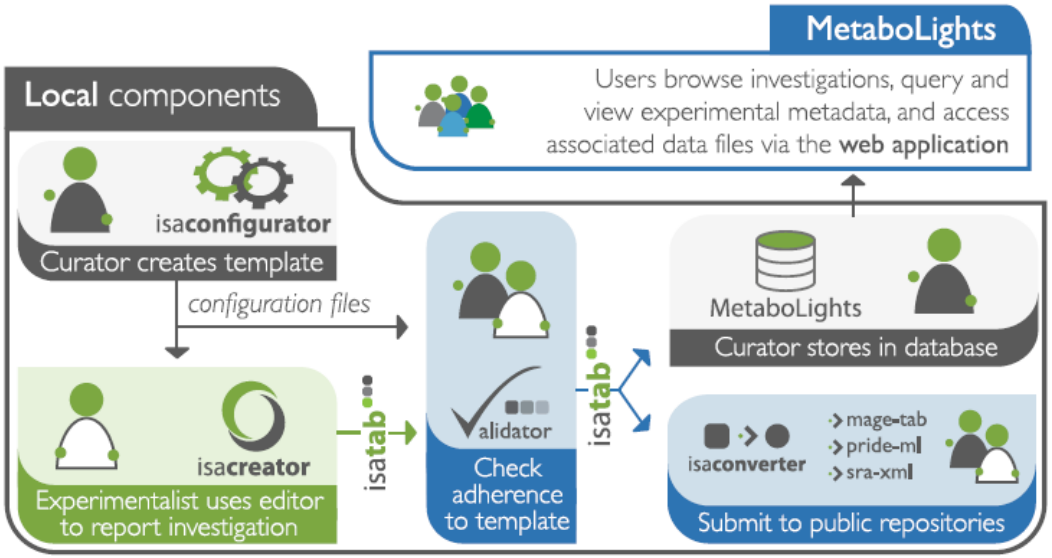




Fig. 2 MetaboLights data submission workflow

# Conclusion

**Apport significatif de la FT-MS dans la détection des métabolites (séparation des ions isobares) et dans leur identification (annotation/élucidation structurale).**

**Disponibilité de différents types de bases de données (bases spectrales de HRMS, de MS/MS, bases de données biochimiques et métabolomique). Les outils de normalisation et de quantification rendent possible la construction de bases de données de profils métaboliques.**

**Un des principaux enjeux: le partage des données concernant les composés inconnus. Pour cela, nécessité de standardisation des spectres MS/MS.**

**Très haute résolution (Orbitrap, FTICR) versus efficacité des acquisitions données dépendantes avec Q-TOF??**

MetaboHUB
Infrastructure Nationale en Métabolomique

# Remerciements


## CEA/LEMM
Jean-François Heilier
Alexandra Lafaye
Céline Ducruix
Erwan Werner
Geoffrey Madalinski
Emmanuel Godat
Jérôme Cotton
Ying Xu
Aurélie Roux
Eric Ezan
Benoit Colsch
Samia Boudah

## CEA/iRCM
Paul-Henri Roméo
Dhouha Darghouth
Lydie Oliveira
Bérengère Koehl
Marie-Françoise Olivier

## Laboratoire de Chimie Structurale Organique et Biologique (UPMC)
Jean-Claude Tabet,
Richard Cole
Denis Lesage, Sandra Alves
Estelle Paris, Farid Ichou

## Profilomic
Céline Ducruix, Jérôme Cotton, Simon Broudin, Fanny Leroux, Bruno Corman, Stéphanie Oursel, Victor Sabarly, Denis Desoubzdanne

## G.H. Pitié-Salpétrière
Frédéric Sedel
Maria del Mar Amador
Fanny Mochel
Foudil Lamari

## CEA/Programmes Transversaaux de Technologies pour la santé et Tox Nuc
Jacques Grassi, Marie-Thérèse Ménager

## CEA/DRT/LIST/LOAD
Antoine Souloumiac, Etienne Thévenot,