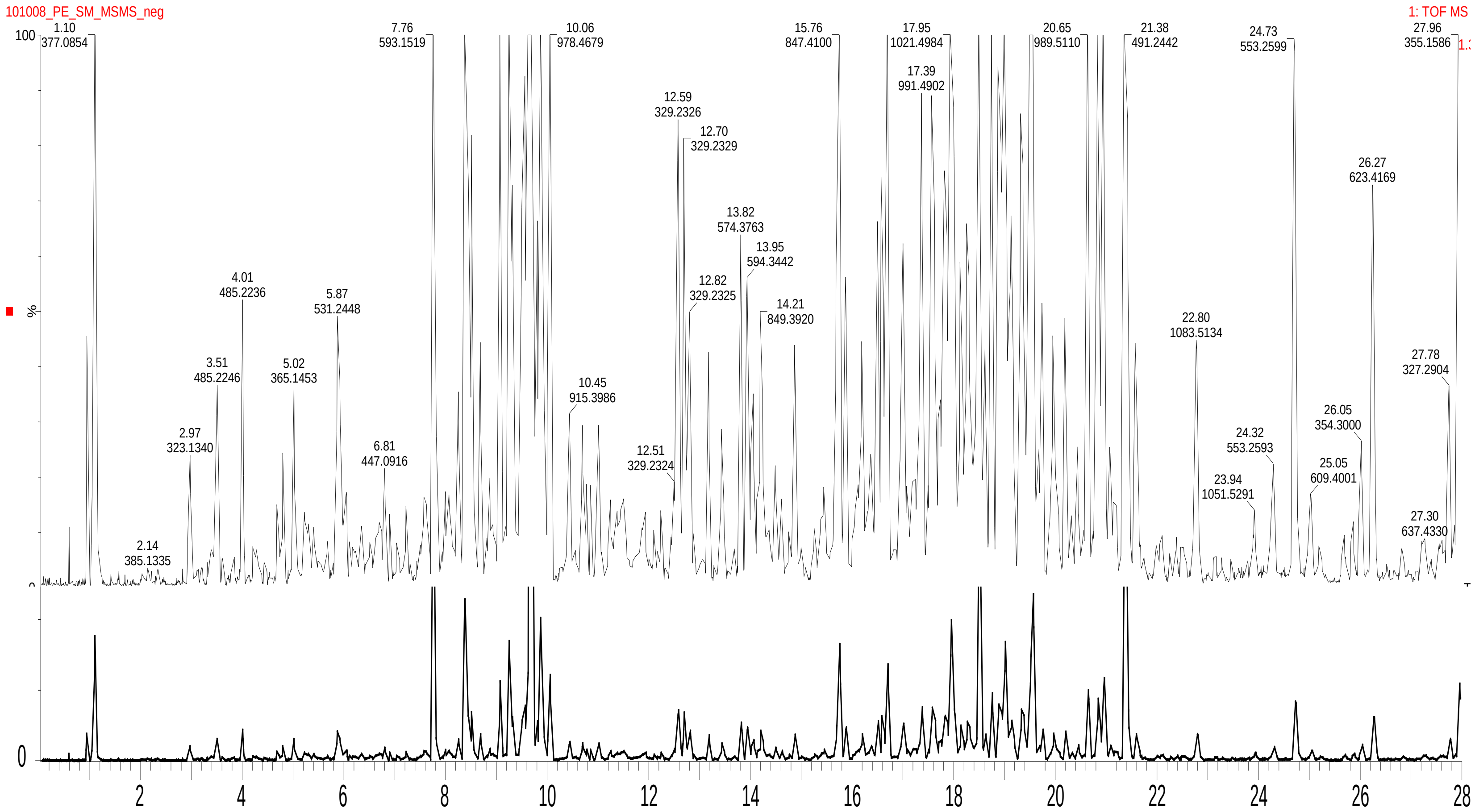
Déréplication (PARTIE I)

Automatisation de l’analyse des données brutes

Samuel Bertrand

1h

Pourquoi l’automatisation ?



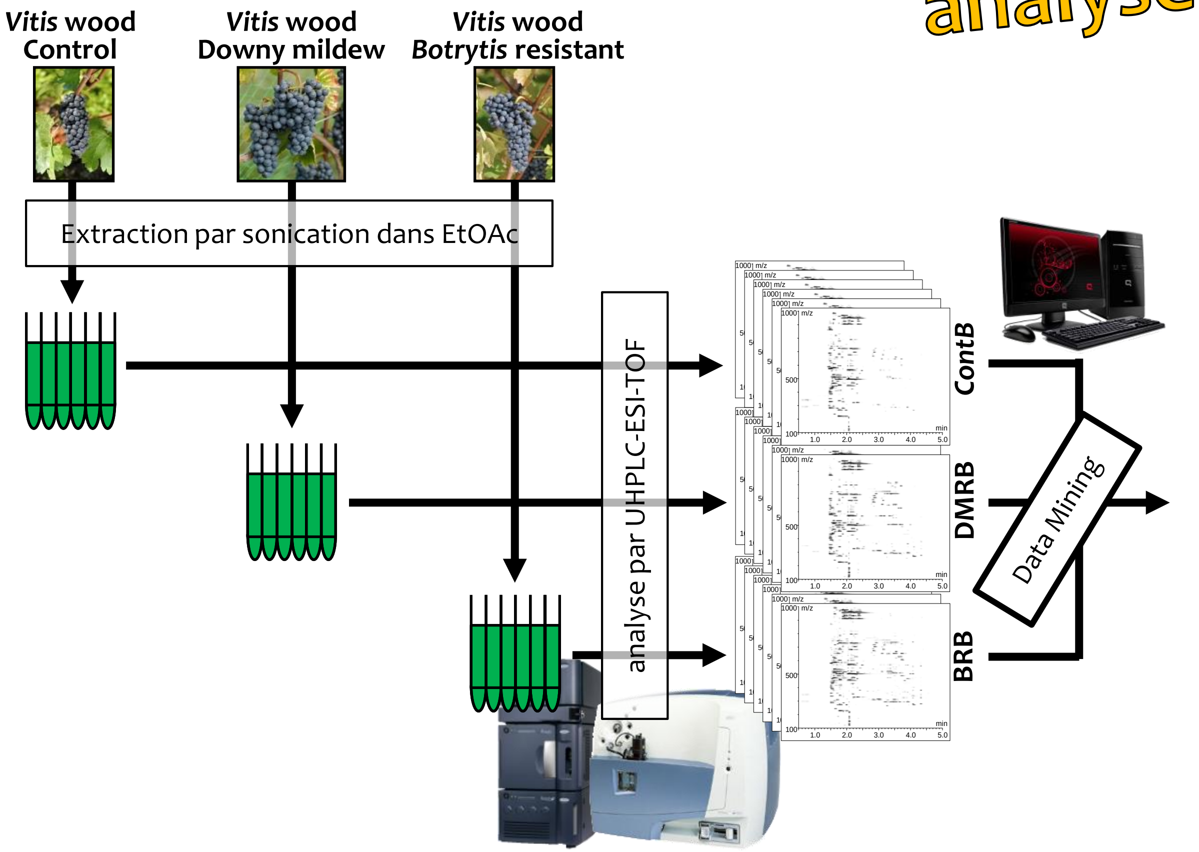
Besoins d’automatisation pour accélérer l’analyse de chromatogrammes

Al-Massarani *et al.* 2011 Phytochemistry 79: 129-140

Pourquoi l'automatisation ?

- **En métabolomique :**

Besoins d'automatisation pour analyser les chromatogrammes de façon homogène !



Vitis wood
Control

Vitis wood
Downy mildew

Vitis wood
Botrytis resistant

Extraction par sonication dans EtOAc

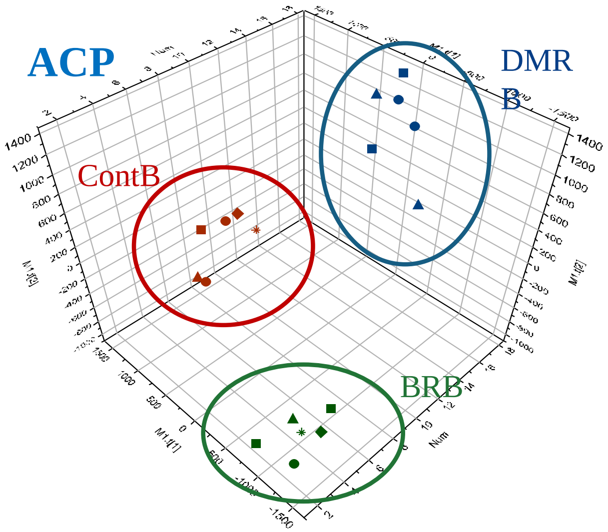
analyse par UHPLC-ESI-TOF

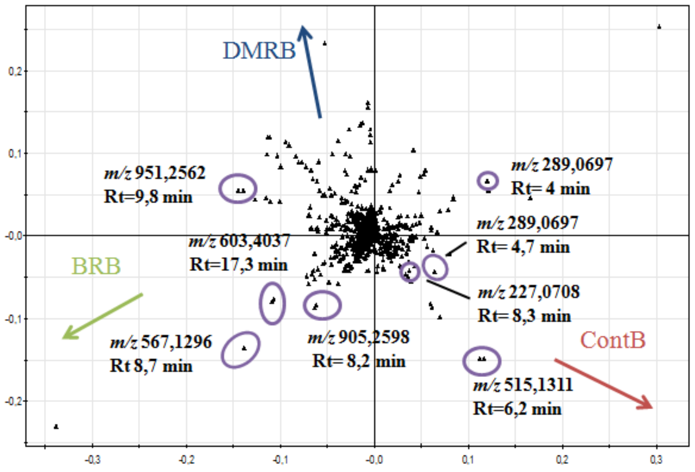
ContB

DMRB

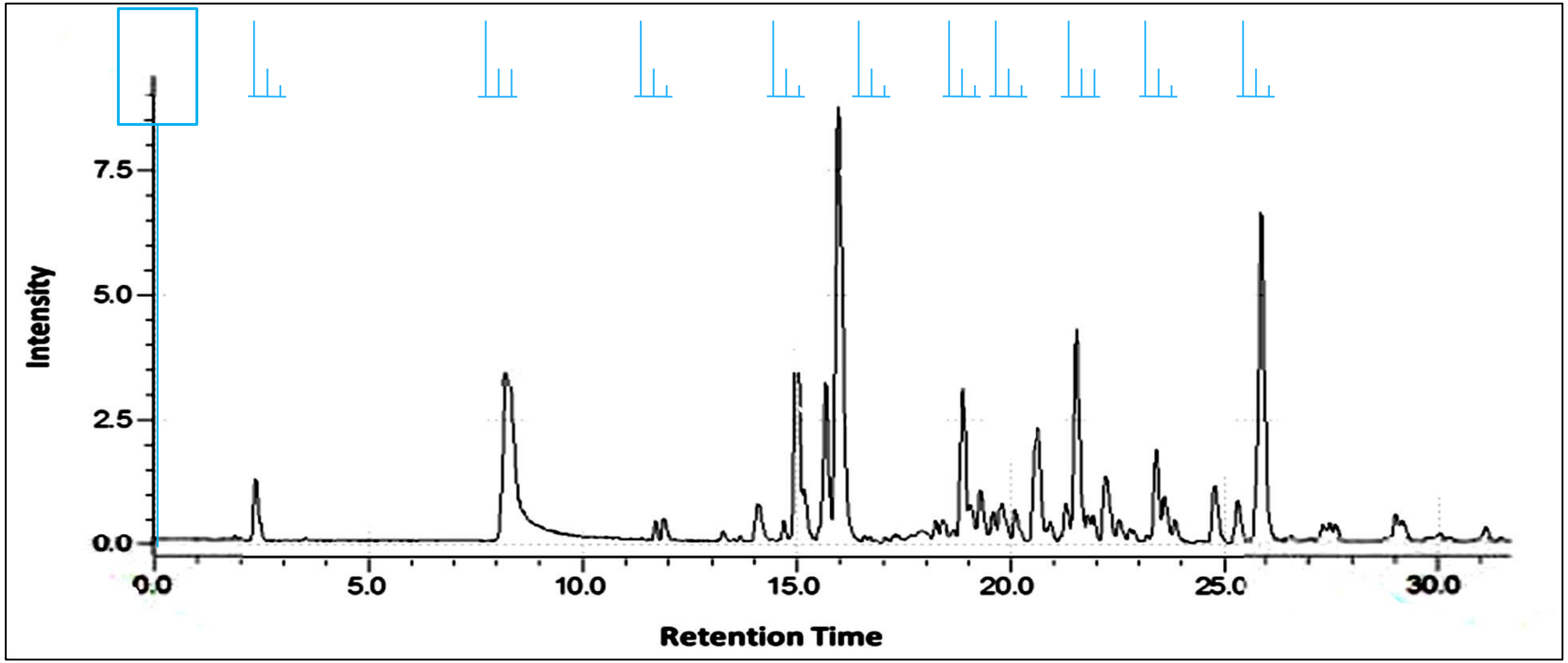
BRB

Data Mining

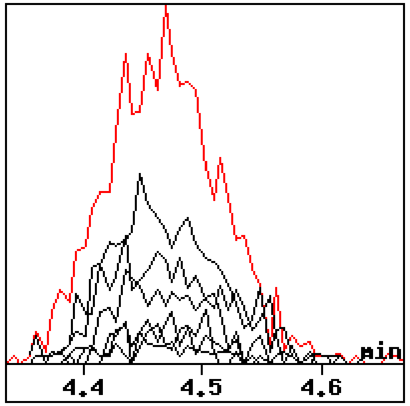




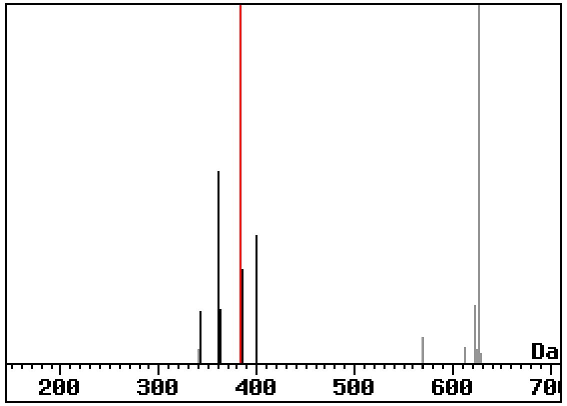
Détection automatique de pics



Regroupement des ions



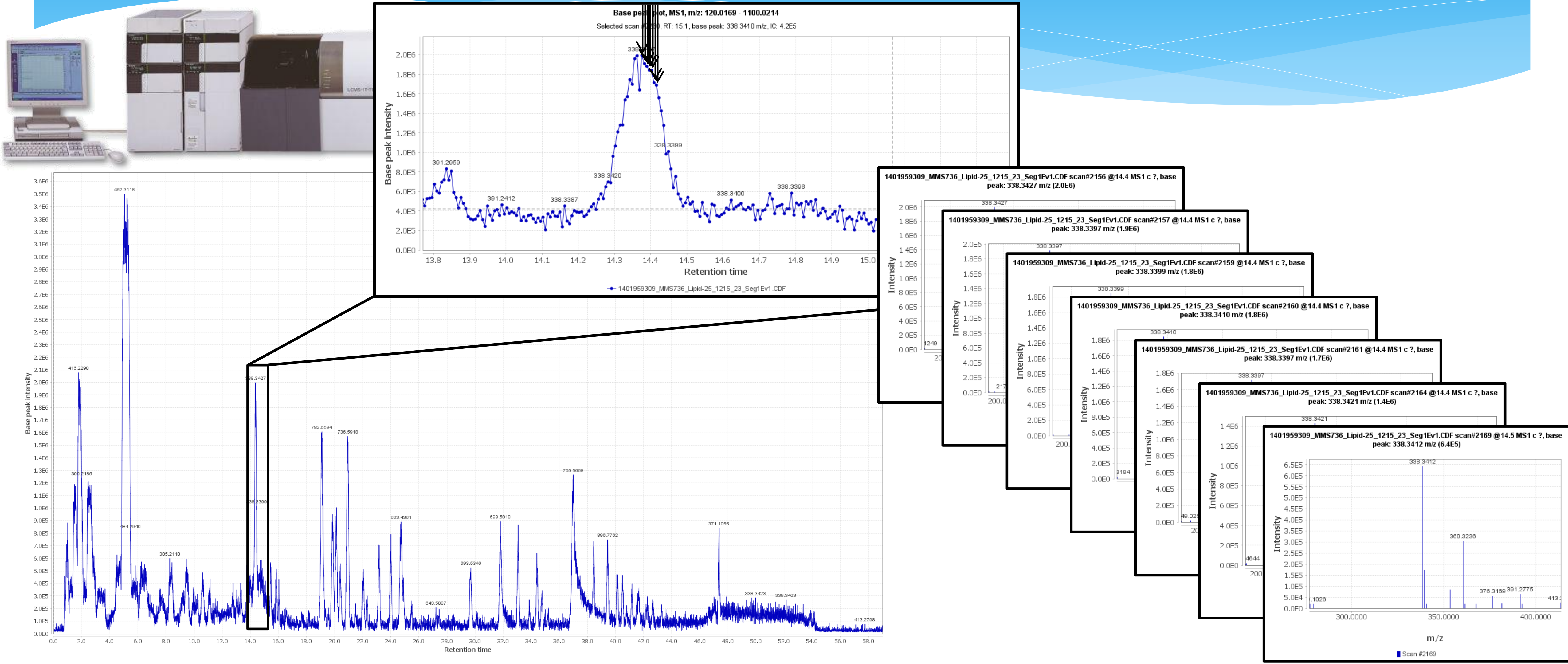
Pseudo-spectre



Détection automatique de pics

Ce qui peut nous paraitre facile ne l'est pas forcement pour un ordinateur !

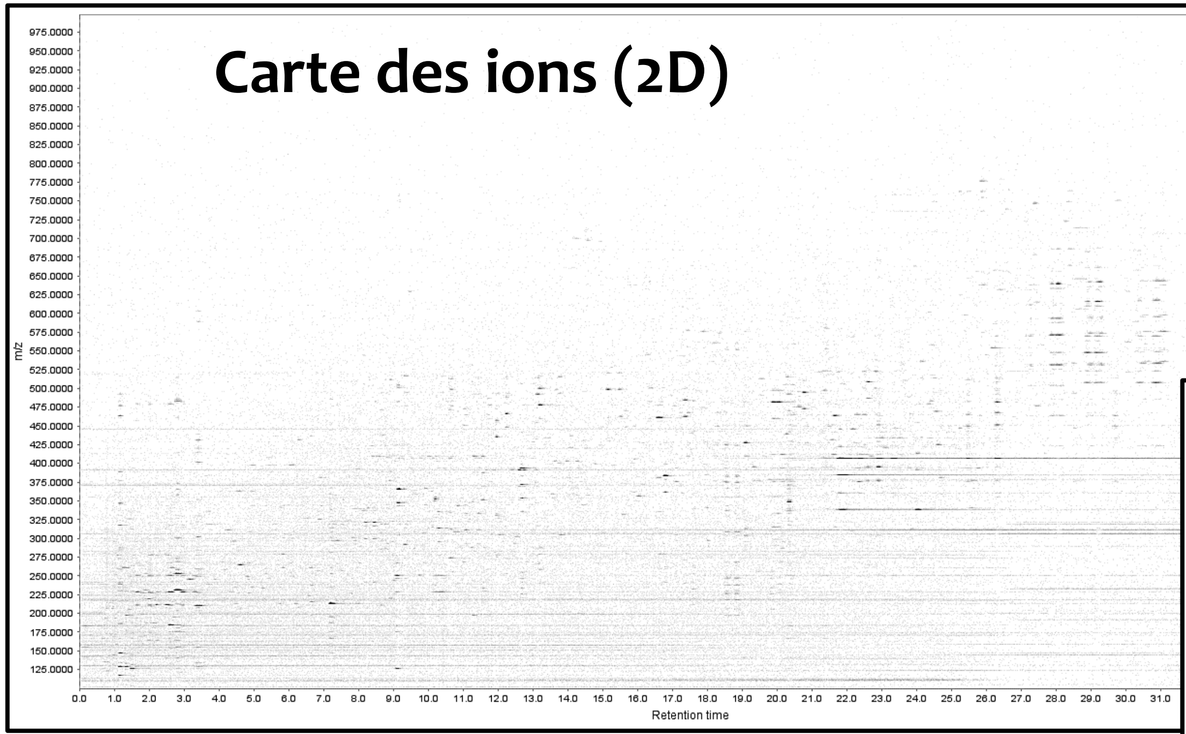
Les données LC-MS

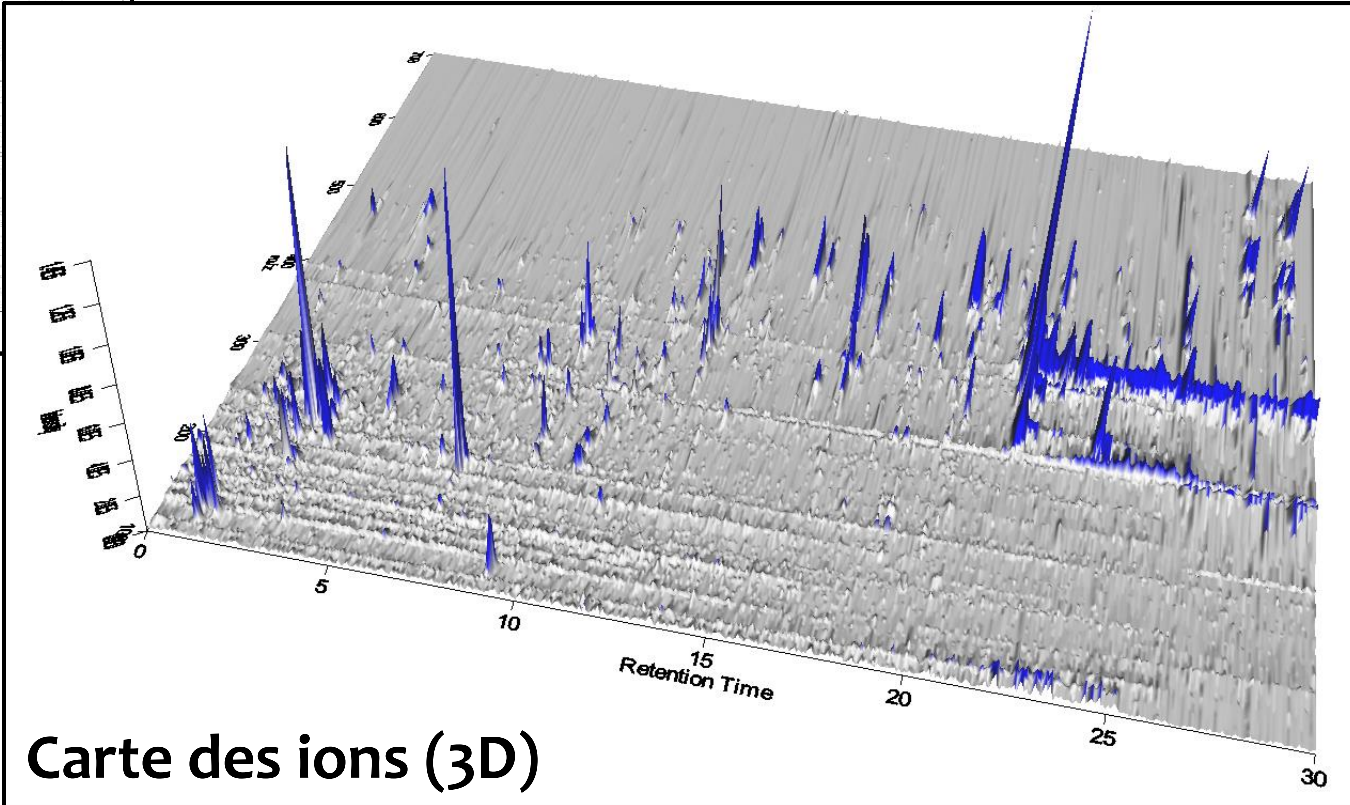


]

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<mzXML xmlns="http://sashimi.sourceforge.net/schema_revision/mzXML_2.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://sashimi.sourceforge.net/schema_revision/mzXML_2.0 http://sashimi.sourceforge.net/schema_revision/mzXML_2.0/mzXML_idx_2.0.xsd">
        <msRun scanCount="7801" startTime="PT0.00000S" endTime="PT52.00000S">
                <parentFile fileName="file://C:\LabSolutions\Data\Samuel\Matthieu_2014\MMS1090_Lipid6-fast_1808_5.lcd" fileType="RAWData"></parentFile>
                <msInstrument>
                        <msManufacturer category="msManufacturer" value="Shimadzu"></msManufacturer>
                        <msModel category="msModel" value="LCMS-IT-TOF"></msModel>
                        <msIonisation category="msIonisation" value="ESI"></msIonisation>
                        <msMassAnalyzer category="msMassAnalyzer" value="IT-TOF"></msMassAnalyzer>
                        <msDetector category="msDetector" value="MCP"></msDetector>
                        <software type="acquisition" name="LCMSsolution" version="3.00"></software>
                </msInstrument>
                <dataProcessing>
                        <software type="conversion" name="TTFLDataExport" version="1"></software>
                </dataProcessing>
                <scan num="1" msLevel="1" peaksCount="215" polarity="+" retentionTime="PT0.00S" lowMz="145" highMz="759" basePeakMz="338.340" basePeakIntensity="78976">
                        <peaks precision="32" byteOrder="network" pairOrder="m/z-
int">QxEc9AAAAABDER/fRk4AAEMRItEAAAAAQxUC/wAAAABDFQX3Rk4AAEMVCOgAAAAAQyUcVwAAAABDJR92Rk4AAEMlIo8AAAAAQzEKnwAAAABDMQ3gRk4AAEMxERoAAAAAQzYzogAAAABDNjbpRk4AAEM2OjAAAAAAQ0APcAAAAABDQBLKRk4AAENAFisAAAAAQ0crX
QAAAABDRy7MRpsAAENHMjoAAAAAQ0ocdwAAAABDSh/sRk4AAENKI2EAAAAAQ0oPpAAAAABDTRMZRk4AAENNFpRGTgAAQ0oaDwAAAABDVinhAAAAAENWLXBGTgAAQ1Yw/wAAAABDVx9PAAAAAENXIt5GTgAAQ1cmbQAAAABDWRLxAAAAAENZFodGTgAAQ1kaHEdOAABDW
R2rRwCAAENZIUEAAAAAQ2EGoQAAAABDYQpKRk4AAENhDeoAAAAAQ2EjygAAAABDYSdsRk4AAENhKw8AAAAAQ2QqnwAAAABDZC5PRpsAAENkMfhGTgAAQ2Q1oUZOAABDZDlLAAAAAENpKhYAAAAAQ2ktxUZOAABDaTF8Rk4AAENpNTIAAAAAQ2obSgAAAABDah8ARk4AAEN
qIrYAAAAAQ3QqkgAAAABDdC5jRs4AAEN0Mi1GTgAAQ3Q19wAAAABDeijoAAAAAEN6LL9GTgAAQ3owlgAAAABDhBEqAAAAAEOEEyNGTgAAQ4QVGwAAAABDhZL1AAAAAEOFlPFGTgAAQ4WW7UZOAABDhZjoAAAAAEOLj40AAAAAQ4uRlkZOAABDi5OcRpsAAEOLlaVGmwAAQ4uX
qwAAAABDjRgKAAAAAEONGhNGzgAAQ40cHEbOAABDjR40AAAAAEONIDEAAAAAQ40iOkZOAABDjSRGRpsAAEONJk8AAAAAQ5OemwAAAABDk6CxRk4AAEOToscAAAAAQ5WW8AAAAABDlZkJRk4AAEOVmyMAAAAAQ5YXHgAAAABDlhk6Rk4AAEOWG1QAAAAAQ5cgkwAAAABD
lyKwRk4AAEOXJMkAAAAAQ5gTpgAAAABDmBXCRpsAAEOYF+IAAAAAQ5kamQAAAABDmRy5Rk4AAEOZHtlGTgAAQ5kg+UZOAABDmSMZRk4AAEOZJTkAAAAAQ5ybzQAAAABDnJ3zRk4AAEOcoBoAAAAAQ6R+NQAAAABDpIBpRk4AAEOkgpxGTgAAQ6SE0AAAAABDqScRAAAAAE
OpKUtHNAAAQ6kriEeaQABDqS3CRzQAAEOpMABGTgAAQ6kyOgAAAABDsJyBAAAAAEOwnshGTgAAQ7ChEEbOAABDsKNaRk4AAEOwpaEAAAAAQ7QmIQAAAABDtChvRk4AAEO0Kr1GTgAAQ7QtCwAAAABDuaT+AAAAAEO5p1VGTgAAQ7mpqgAAAABDvSauAAAAAEO9KQlGTgAA
Q70rZwAAAABDw510AAAAAEPDn9hGmwAAQ8OiQEc0AABDw6SlR5pAAEPDpw1HjYAAQ8OpdUZOAABDw6vaAAAAAEPEHy4AAAAAQ8QhlkabAABDxCP+Rs4AAEPEJmZGTgAAQ8QozgAAAABDxSWRAAAAAEPFJ/xGTgAAQ8UqZEabAABDxSzMRk4AAEPFLzgAAAAAQ8krKQAAAAB
DyS2bRk4AAEPJMAkAAAAAQ8wjLAAAAABDzCWhRk4AAEPMKBMAAAAAQ86d4wAAAABDzqBbRpsAAEPOotRGzgAAQ86lSUabAABDzqfBAAAAAEPPHkUAAAAAQ88gvkZOAABDzyM2Rk4AAEPPJbIAAAAAQ9ImXwAAAABD0ijeRk4AAEPSK10AAAAAQ9QgvgAAAABD1CM9Rk4AAE
PUJb8AAAAAQ9h3eQAAAABD2Hn/Rk4AAEPYfIQAAAAAQ/I9pQAAAABD8kBSRk4AAEPyQv4AAAAARAFT5wAAAABEAVVHRk4AAEQBVqkAAAAARAFYCwAAAABEAVltRk4AAEQBWs0AAAAARAsNxAAAAABECw8zRk4AAEQLEKAAAAAARAxMPgAAAABEDE2vRk4AAEQMTx9GTgAAR
AxQjgAAAABEEO9xAAAAAEQQ8OhGTgAARBDyXgAAAABEFyv8AAAAAEQXLXpGTgAARBcu+AAAAABEIdkMAAAAAEQh2pdGTgAARCHcJEbOAABEId2vAAAAAEQiG3sAAAAARCIdBkZOAABEIh6SAAAAAEQl3zkAAAAARCXgyUZOAABEJeJbAAAAAEQrS7QAAAAARCtNS0ZOAABEK07j
AAAAAEQu3sIAAAAARC7gXUZOAABELuH4AAAAAEQ9zFAAAAAARD3N/UZOAABEPc+pAAAAAA==</peaks>
                </scan>
                <scan num="3" msLevel="1" peaksCount="225" polarity="+" retentionTime="PT0.40S" lowMz="149" highMz="1412" basePeakMz="391.286" basePeakIntensity="92288">
                        <peaks precision="32" byteOrder="network" pairOrder="m/z-
int">QxUABgAAAABDFQL/RlAAAEMVBfdGUAAAQxUI6AAAAABDHyJAAAAAAEMfJUxGUAAAQx8oXwAAAABDMN1qAAAAAEMw4KRGUAAAQzDj3QAAAABDMQdmAAAAAEMxCp9GUAAAQzEN4AAAAABDNinOAAAAAEM2LRVGUAAAQzYwW0ZQAABDNjOiAAAAAEM9rQEAAAA
AQz2wVUZQAABDPbOvAAAAAENPHeYAAAAAQ08hYkZQAABDTyTjAAAAAENPKF9GUAAAQ08r4QAAAABDWRLxAAAAAENZFodGmwAAQ1kaHEcBAABDWR2rRpsAAENZIUFGUAAAQ1kk1gAAAABDWiMfAAAAAENaJrVGUAAAQ1oqSgAAAABDZCqfAAAAAENkLk9GmwAAQ2Qx+EcB
AABDZDWhRpsAAENkOUtGmwAAQ2Q8+gAAAABDaSoWAAAAAENpLcVGUAAAQ2kxfAAAAABDb72yAAAAAENvwXVGUAAAQ2/FOQAAAABDdSzMAAAAAEN1MJ1GUAAAQ3U0ZwAAAABDeiURAAAAAEN6KOhHAQAAQ3osv0cBAABDejCWAAAAAEN7LkgAAAAAQ3syIEZQAABDezX3
AAAAAEOFkPkAAAAAQ4WS9UZQAABDhZTxRpsAAEOFlu0AAAAAQ4aXXAAAAABDhplYRlAAAEOGm1cAAAAAQ4eOwgAAAABDh5DBRlAAAEOHksAAAAAAQ4gc0AAAAABDiB7PRlAAAEOIINEAAAAAQ4uPjQAAAABDi5GWRlAAAEOLk5xGzwAAQ4uVpUZQAABDi5erAAAAAEONEesAA
AAAQ40T9EZQAABDjRYBAAAAAEONGA0AAAAAQ40aE0abAABDjRwcRlAAAEONHigAAAAAQ40gMQAAAABDjSI6RlAAAEONJEYAAAAAQ46IAwAAAABDjooPRlAAAEOOjBwAAAAAQ5gTpgAAAABDmBXCRlAAAEOYF+IAAAAAQ5kcuQAAAABDmR7ZRlAAAEOZIPlGzwAAQ5kjGUZQAA
BDmSU5AAAAAEOfmhYAAAAAQ5+cQEZQAABDn55pAAAAAEOmnZQAAAAAQ6afy0ZQAABDpqICAAAAAEOpJNMAAAAAQ6knEUbPAABDqSlLR2gAAEOpK4hHp0AAQ6ktwkd0AABDqTAAAAAAAEOpqKcAAAAAQ6mq5EbPAABDqa0iRlAAAEOpr1wAAAAAQ7CDawAAAABDsIWyRlA
AAEOwh/kAAAAAQ7CcgQAAAABDsJ7IRlAAAEOwoRBGmwAAQ7CjWgAAAABDsSDsAAAAAEOxIzNGUAAAQ7ElfQAAAABDtCPTAAAAAEO0JiFGUAAAQ7Qob0ZQAABDtCq9RwEAAEO0LQtGmwAAQ7QvWQAAAABDtSiNAAAAAEO1Kt5GUAAAQ7UtLAAAAABDtyrrAAAAAEO3LTxGUA
AAQ7cvkAAAAABDt60bAAAAAEO3r3BGUAAAQ7exxAAAAABDwZguAAAAAEPBmpJGUAAAQ8Gc9wAAAABDw510AAAAAEPDn9hGzwAAQ8OiQEenQABDw6SlR7RAAEPDpw1HGoAAQ8OpdQAAAABDxCGWAAAAAEPEI/5GzwAAQ8QmZkabAABDxCjOAAAAAEPEpeYAAAAAQ8SoTkZ
QAABDxKq2AAAAAEPFJ/wAAAAAQ8UqZEZQAABDxSzMRlAAAEPFLzgAAAAAQ8axowAAAABDxrQORlAAAEPGtnoAAAAAQ8spSwAAAABDyyvARpsAAEPLLjIAAAAAQ8wjLAAAAABDzCWhRlAAAEPMKBNGUAAAQ8wqiAAAAABDzKU/AAAAAEPMp7RGUAAAQ8yqKQAAAABDzptrAA
AAAEPOneNGUAAAQ86gW0bPAABDzqLURpsAAEPOpUlGUAAAQ86nwQAAAABDzyC+AAAAAEPPIzZGUAAAQ88lsgAAAABD0aUOAAAAAEPRp4pGUAAAQ9GqBQAAAABD2KTwAAAAAEPYp3ZGUAAAQ9ip/wAAAABD3JIFAAAAAEPclJFGUAAAQ9yXIQAAAABD4JUBAAAAAEPgl5BG
UAAAQ+CaIwAAAABD4Se7AAAAAEPhKk1GUAAAQ+Es4EZQAABD4S92AAAAAEPn5WYAAAAAQ+foA0ZQAABD5+qfAAAAAEP7oRMAAAAAQ/ujzUZQAABD+6aDAAAAAEP+kBoAAAAAQ/6S10ZQAABD/pWURlAAAEP+mFEAAAAARACylwAAAABEALP4RlAAAEQAtVgAAAAARAlTawAA
AABECVTWRlAAAEQJVkQAAAAARAxGfwAAAABEDEfuRlAAAEQMSV4AAAAARBWx1AAAAABEFbNQRlAAAEQVtM0AAAAARCsMLgAAAABEKw3ERlAAAEQrD10AAAAARCuLRwAAAABEK4zfRlAAAEQrjnUAAAAARLCXoQAAAABEsJjERlAAAESwmegAAAAA</peaks>
                </scan>
```

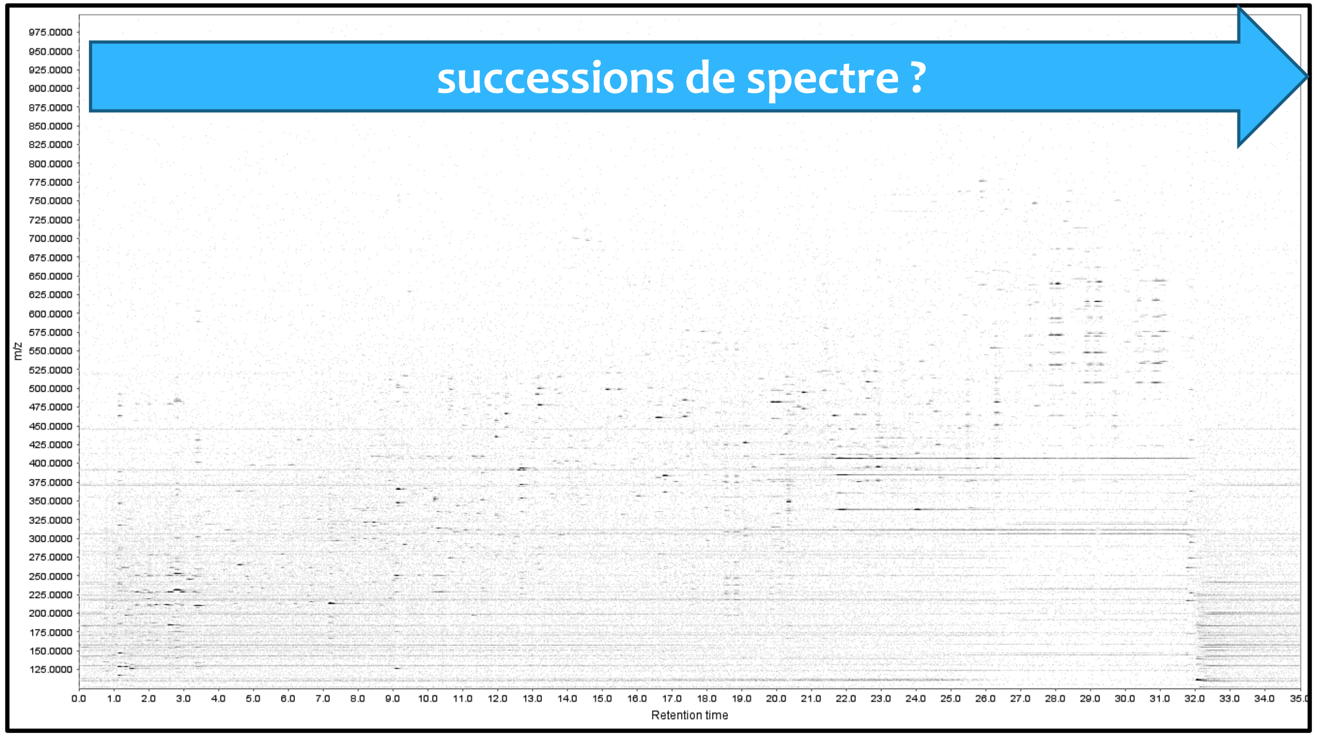
Les données LC-MS


Carte des ions (2D)


Carte des ions (3D)

La détection automatique de pics

- **Comment traiter des successions de spectre ?**



successions de spectre ?

La détection automatique de pics

- Comment traiter des successions de spectre ?

- Divers logiciels existants :
  - R/XCMS ou XCMS online
  - MZmine 2.0
  - MetaboAnalyst (en ligne)
  - OpenMS
  - OpenChrom
  - Mass++
  - …
  - Logiciels constructeurs: MassHunter (Agilent), MarkerLinx (Waters), Progenesis QI (Nonlinear Dynamics, Waters), IntelliXtract (ACD)…

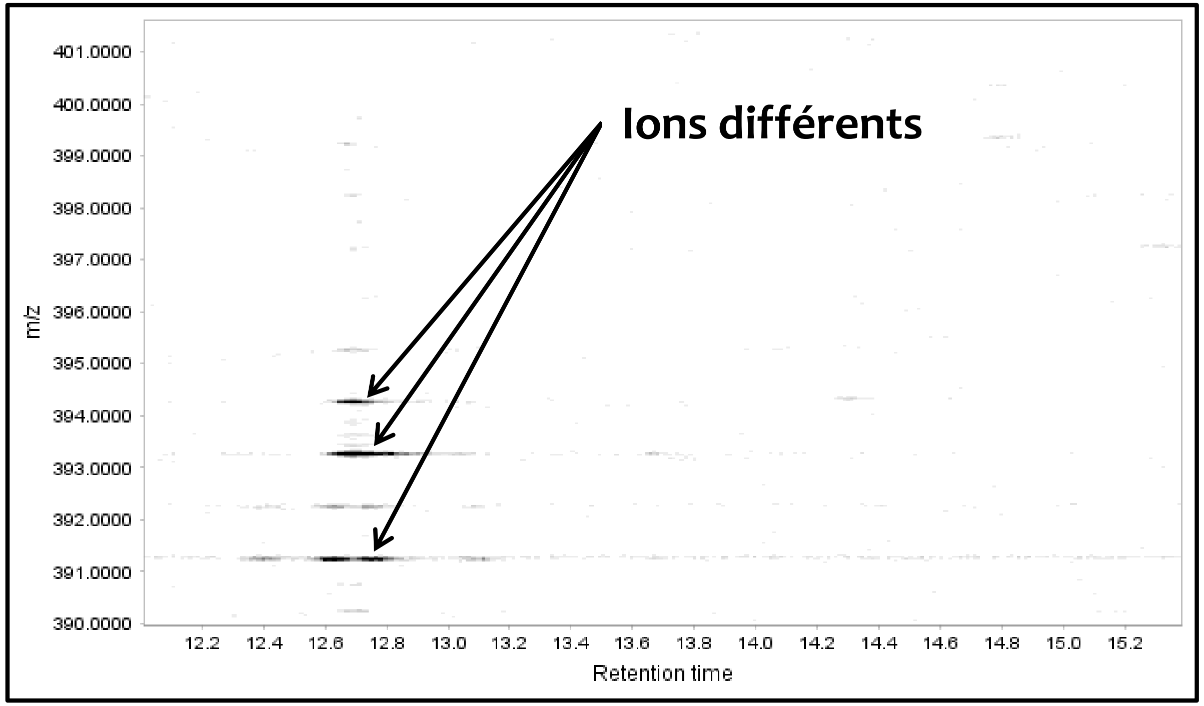
}

Gratuit
Compatible avec les format ouverts (cdf, mzXML)

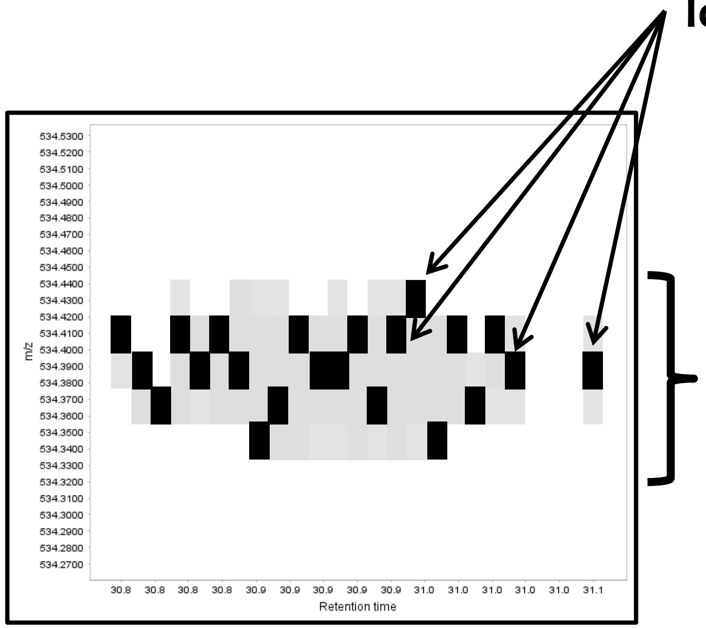
La détection automatique de pics

- **Les contraintes ?**

1. la différence de masse entre deux ions



Ions différents



Ions différents

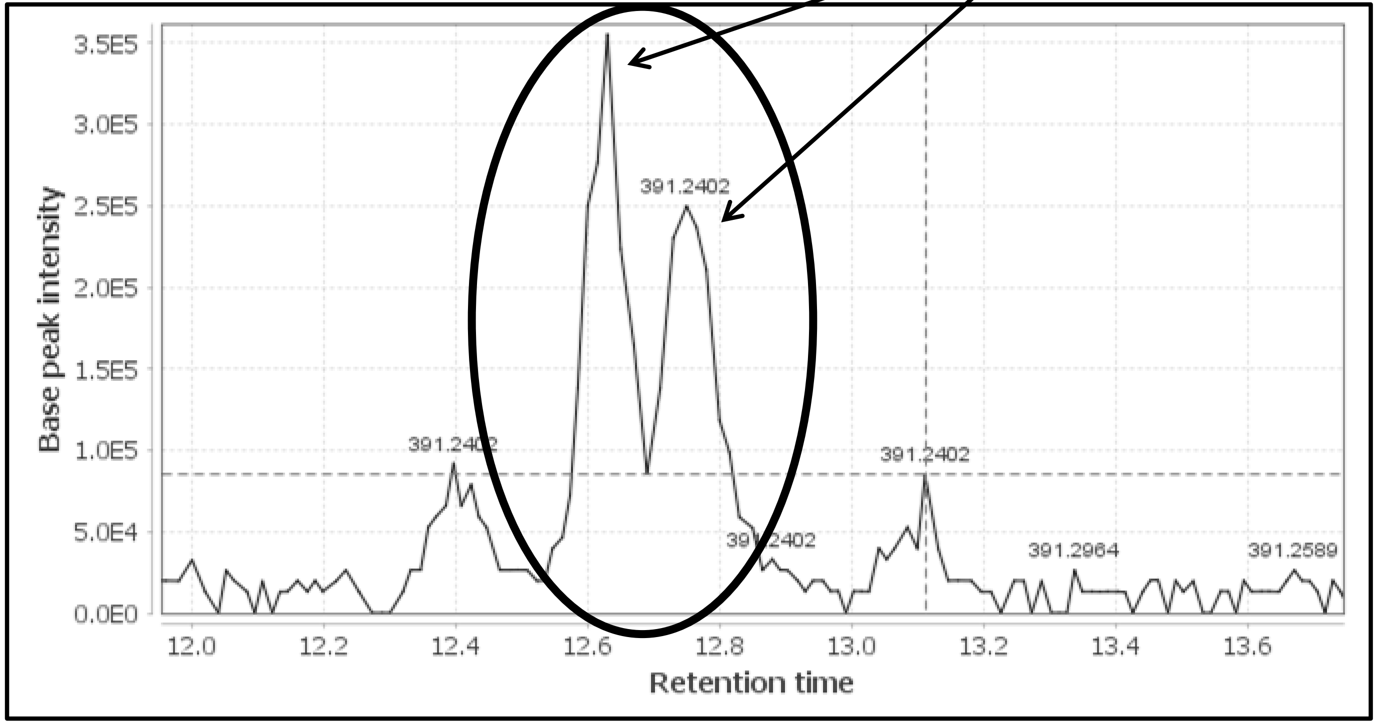
Fonction du spectromètre

La détection automatique de pics

- Les contraintes ?

2. la séparation des pics

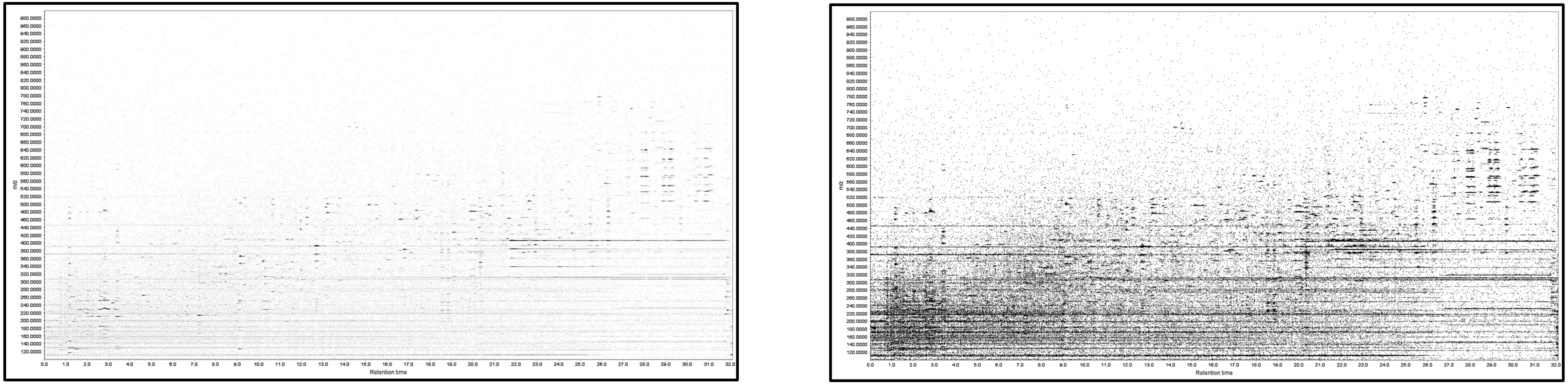
Deux pics ?



La détection automatique de pics

- **Les contraintes ?**

3. la présence de bruit

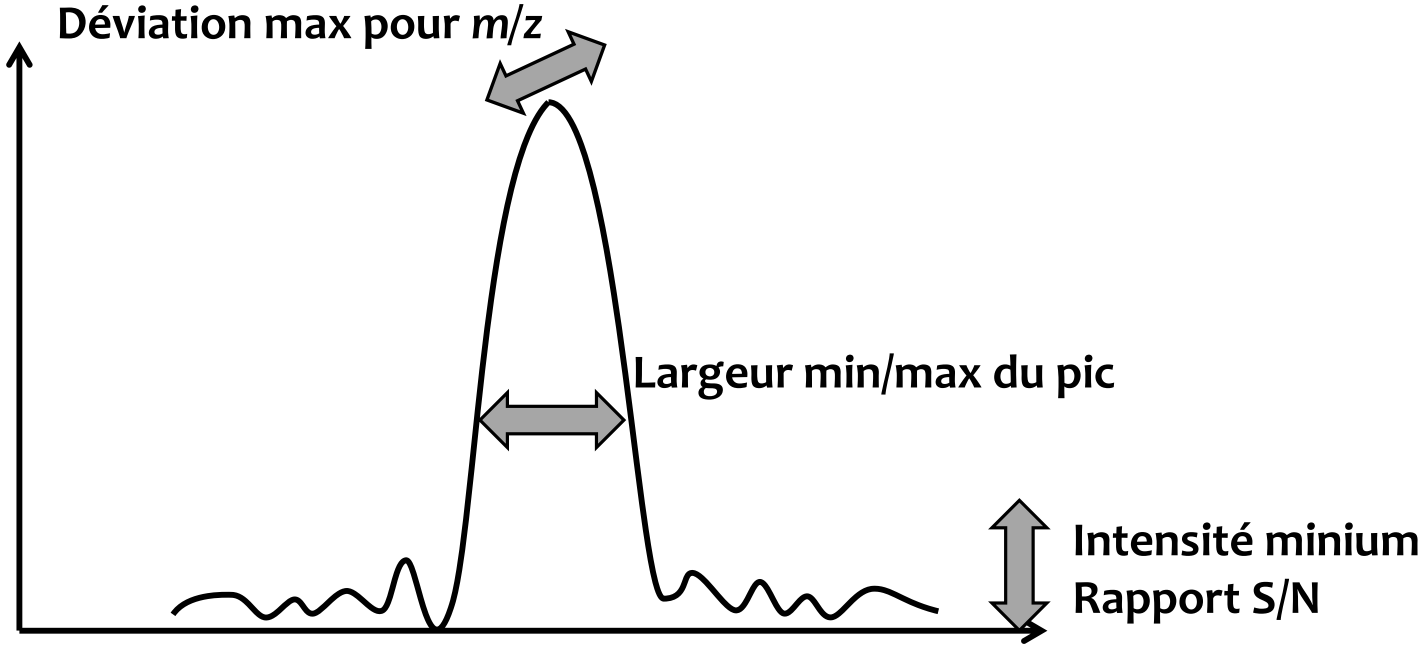


Même données

La détection automatique de pics

- Les contraintes ?

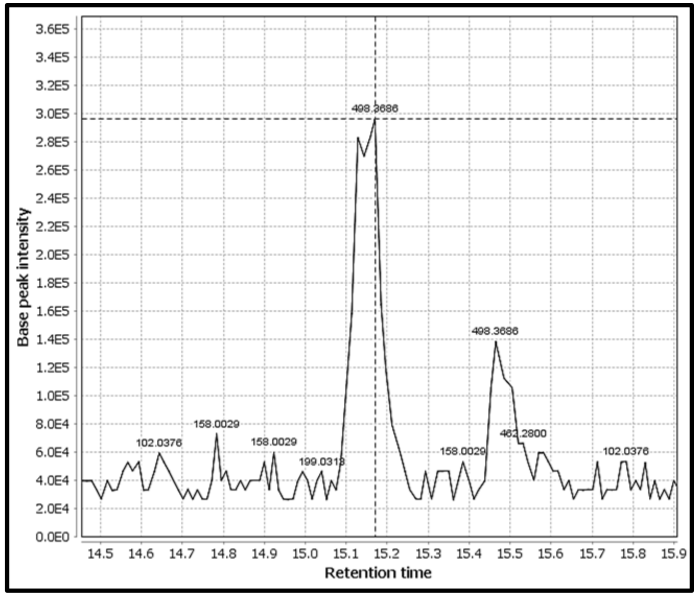
  1. la différence de masse entre deux ions
  2. la séparation des pics
  3. la présence de bruit

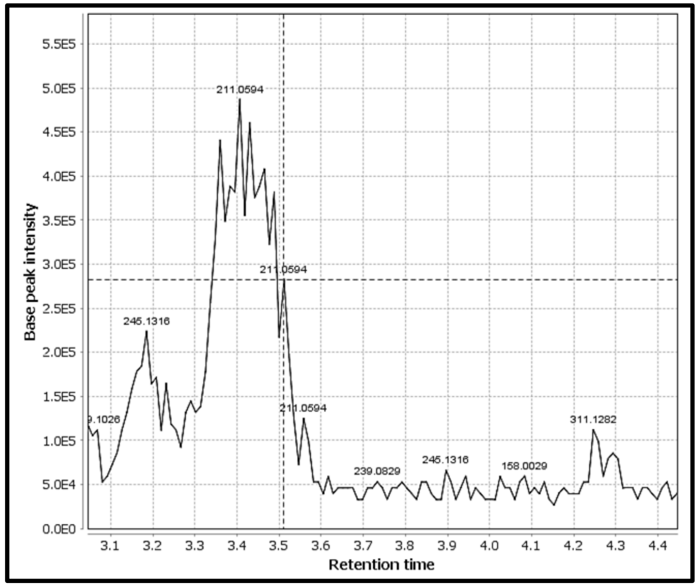


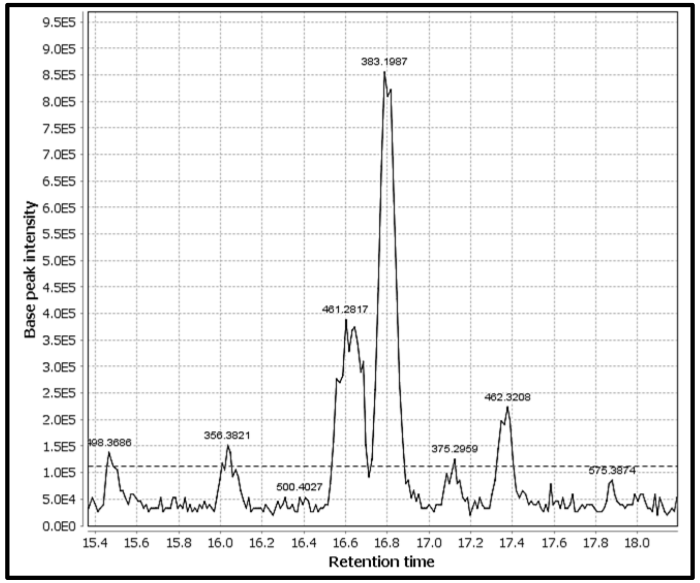
La détection automatique de pics

- **En réalité :**

Plus complexe !!!

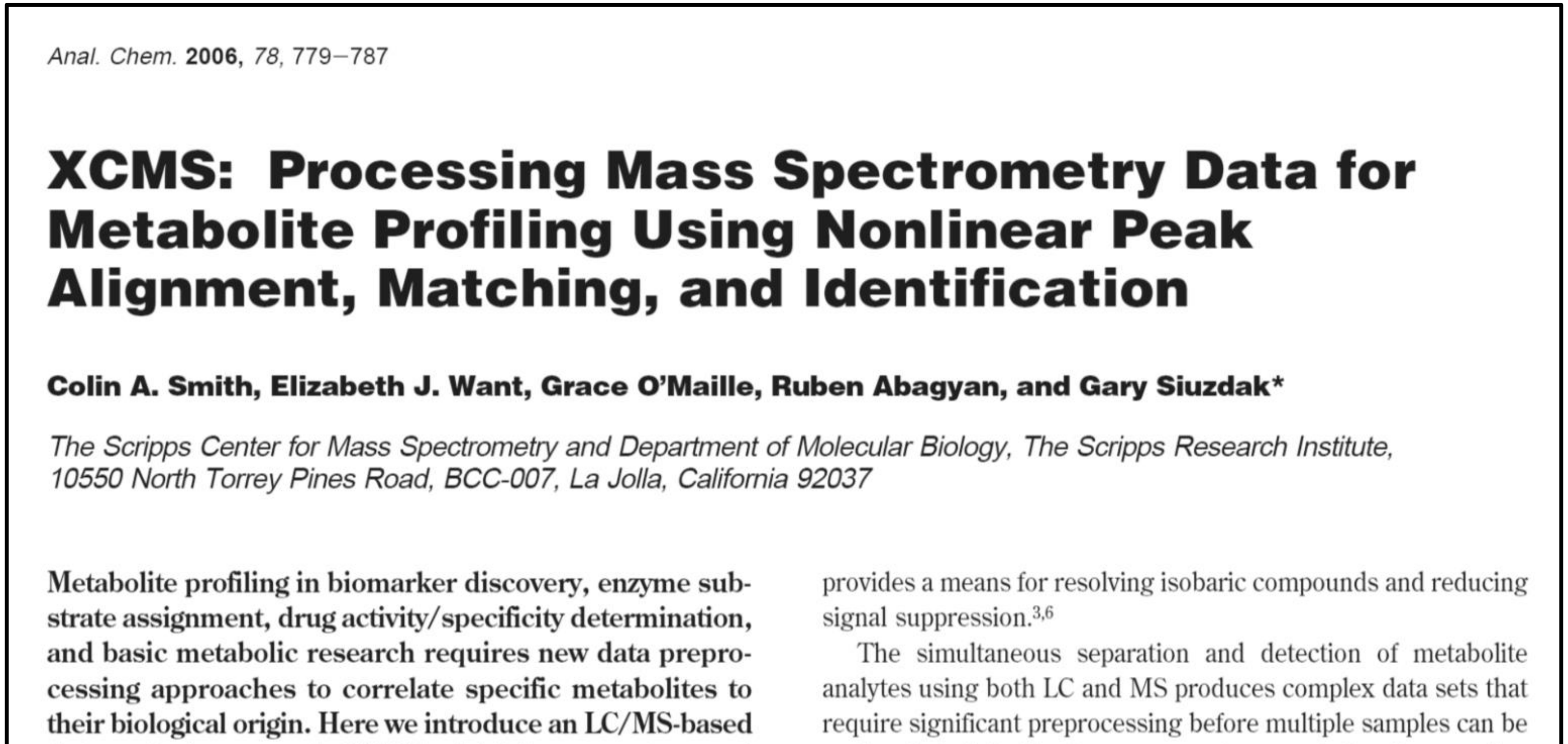






Les logiciels

- **Le « gold standard » :**



Anal. Chem. **2006**, 78, 779–787

XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification

Colin A. Smith, Elizabeth J. Want, Grace O'Maille, Ruben Abagyan, and Gary Siuzdak*

The Scripps Center for Mass Spectrometry and Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, BCC-007, La Jolla, California 92037

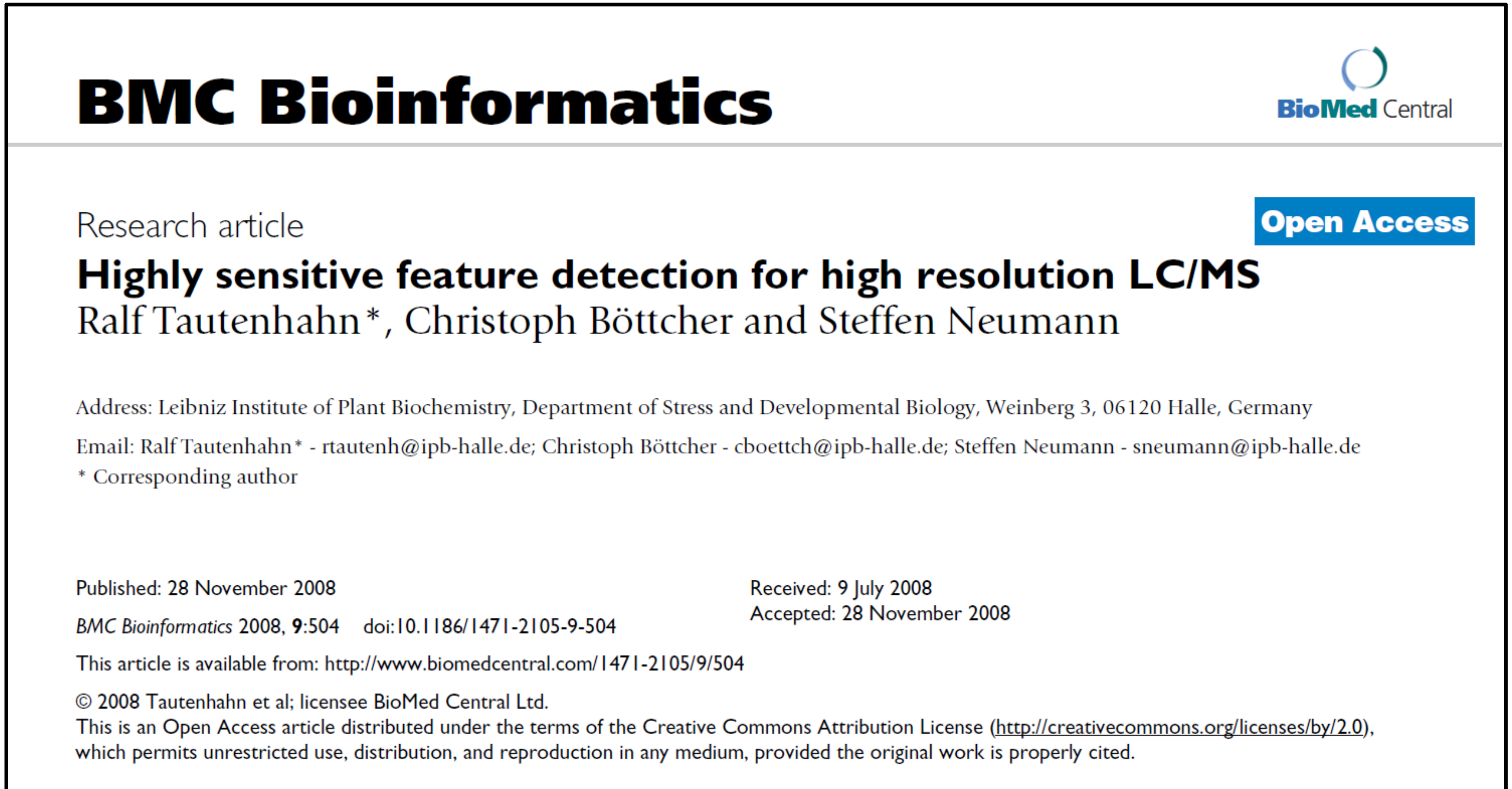
Metabolite profiling in biomarker discovery, enzyme substrate assignment, drug activity/specificity determination, and basic metabolic research requires new data preprocessing approaches to correlate specific metabolites to their biological origin. Here we introduce an LC/MS-based

provides a means for resolving isobaric compounds and reducing signal suppression.[3,6]

The simultaneous separation and detection of metabolite analytes using both LC and MS produces complex data sets that require significant preprocessing before multiple samples can be

Les logiciels

- **Le « gold standard » :**



BMC Bioinformatics

BioMed Central

Research article

Open Access

Highly sensitive feature detection for high resolution LC/MS

Ralf Tautenhahn*, Christoph Böttcher and Steffen Neumann

Address: Leibniz Institute of Plant Biochemistry, Department of Stress and Developmental Biology, Weinberg 3, 06120 Halle, Germany

Email: Ralf Tautenhahn* - rtautenh@ipb-halle.de; Christoph Böttcher - cboettch@ipb-halle.de; Steffen Neumann - sneumann@ipb-halle.de
* Corresponding author

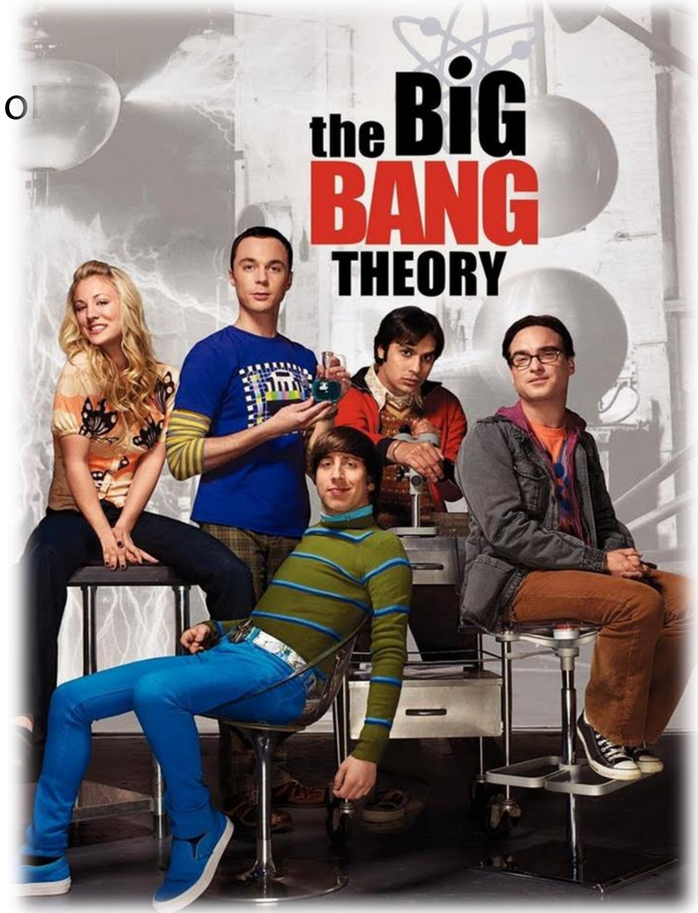
Published: 28 November 2008

BMC Bioinformatics 2008, **9**:504 doi:10.1186/1471-2105-9-504

Received: 9 July 2008
Accepted: 28 November 2008

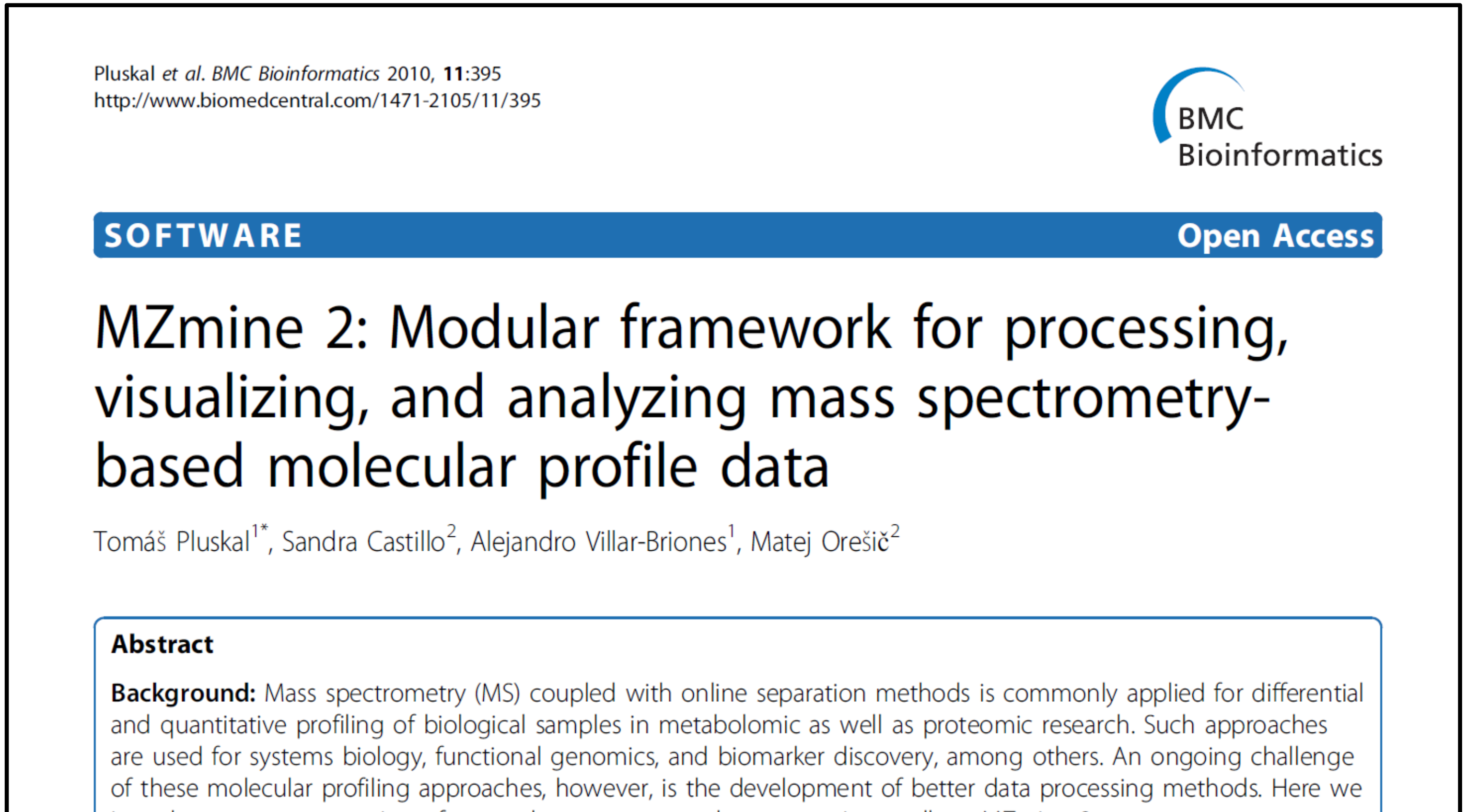
This article is available from: http://www.biomedcentral.com/1471-2105/9/504

```
R version 3.0.2 (2013-09-25) -- "Frisbee Sailing"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R est un logiciel libre livré sans AUCUNE GARANTIE. Vous pouvez le redistribuer sous certaines conditions. Tapez 'license()'
ou 'licence()' pour plus de détails.

R est un projet collaboratif avec de nombreux contributeurs. Tapez 'contributors()' pour plus d'information et 'citation()'
pour la façon de le citer dans les publications.

Tapez 'demo()' pour des démonstrations, 'help()' pour l'aide en ligne ou 'help.start()' pour o
Tapez 'q()' pour quitter R.

|
```



Les logiciels

- **Alternative :**
  - **Mzmine 2**



Pluskal *et al. BMC Bioinformatics* 2010, **11**:395
http://www.biomedcentral.com/1471-2105/11/395

BMC Bioinformatics

SOFTWARE **Open Access**

MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data

Tomáš Pluskal[1*], Sandra Castillo[2], Alejandro Villar-Briones[1], Matej Orešič[2]

Abstract

Background: Mass spectrometry (MS) coupled with online separation methods is commonly applied for differential and quantitative profiling of biological samples in metabolomic as well as proteomic research. Such approaches are used for systems biology, functional genomics, and biomarker discovery, among others. An ongoing challenge of these molecular profiling approaches, however, is the development of better data processing methods. Here we

Les logiciels

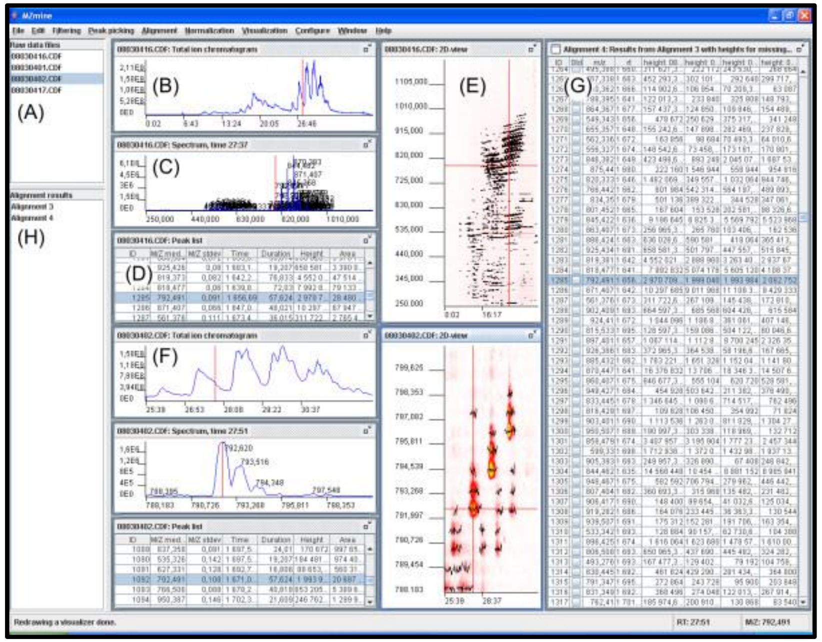
- **Alternative :**

  - **Mzmine 2**

    Avantage 1: interface graphique

    Avantage 2: compréhension des étapes

    Avantage 3: utilisable avec XCMS



La détection de pic

- **Comment ça fonctionne ???**

Détection des Ions d'intérêt

Création des chromatogrammes

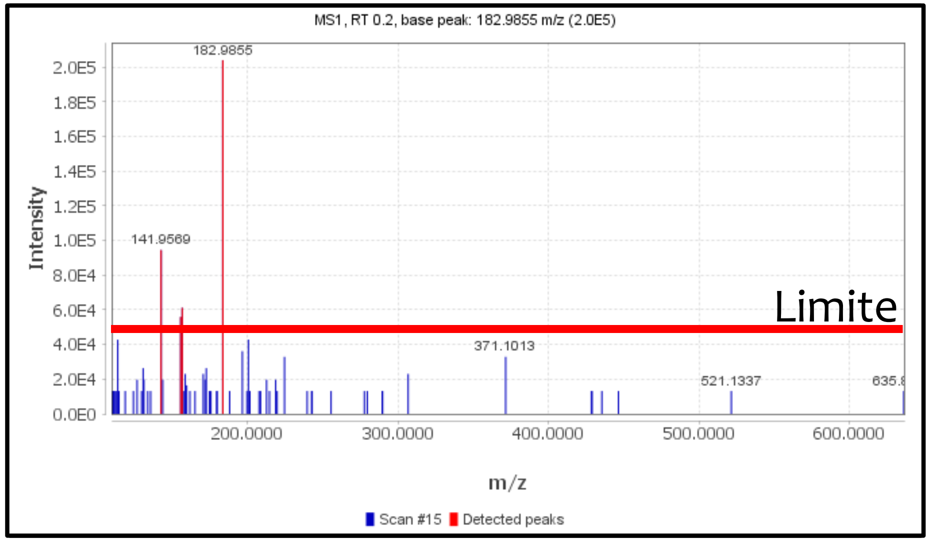
Déconvolution des chromatogrammes

Alignement des pics

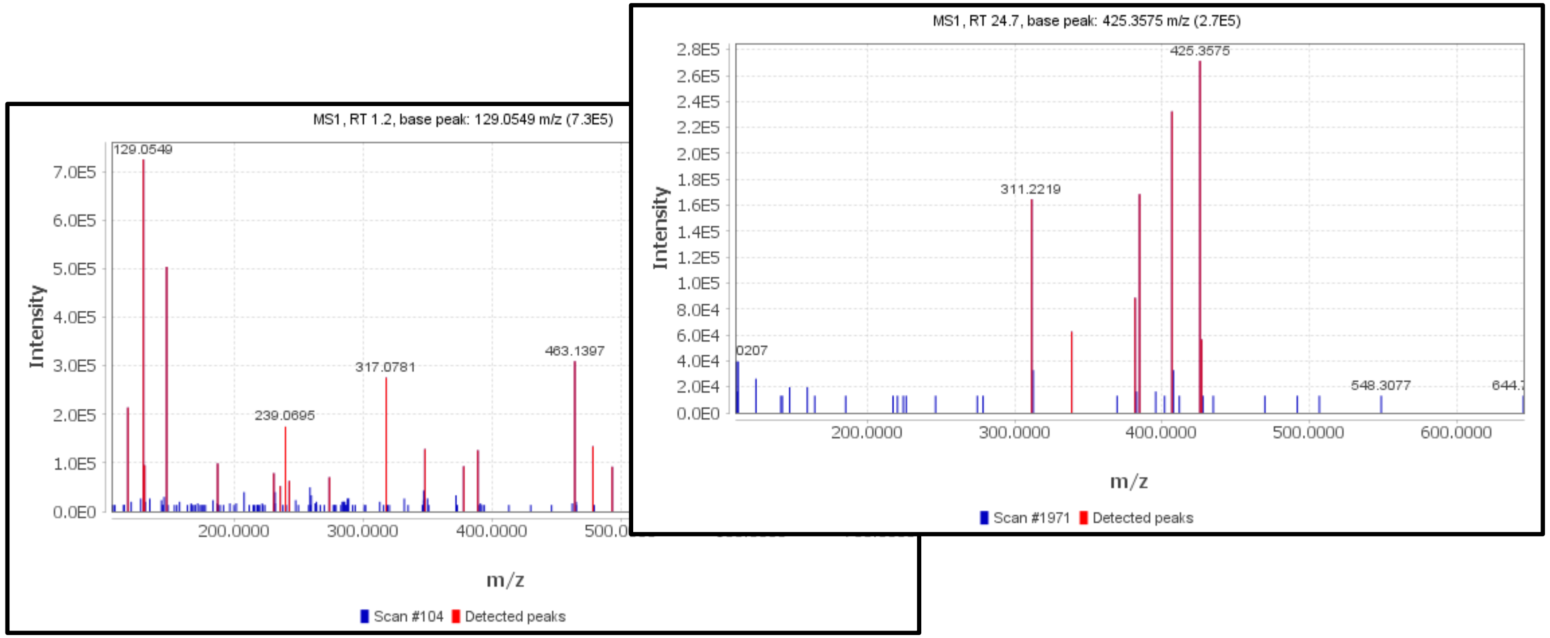
Globalement même fonctionnement pour toutes les approches

La détection de pic

ETAPE I : Détection des ions d’intérêt

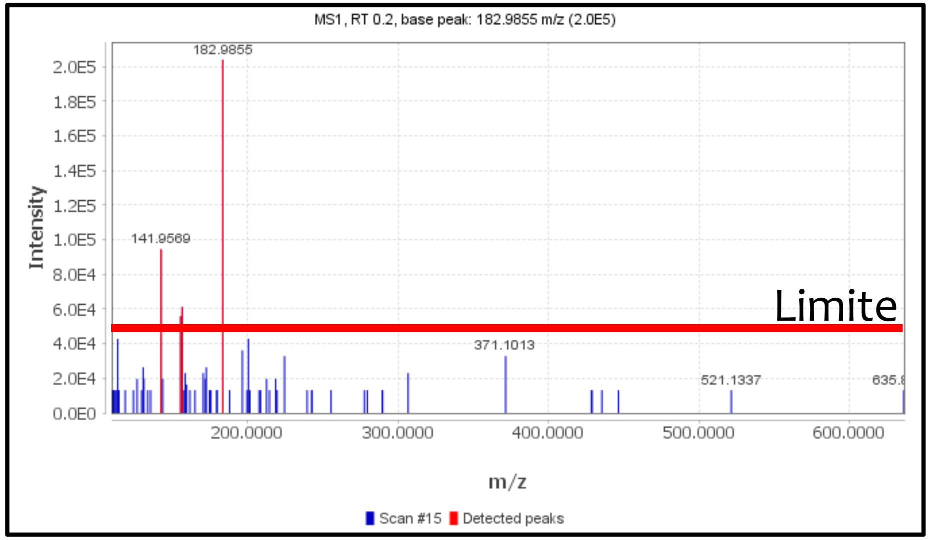


Recherche spectre après spectres de l’ensemble des ions ayant une intensité supérieur à la limite.



La détection de pic

ETAPE I : Détection des ions d'intérêt

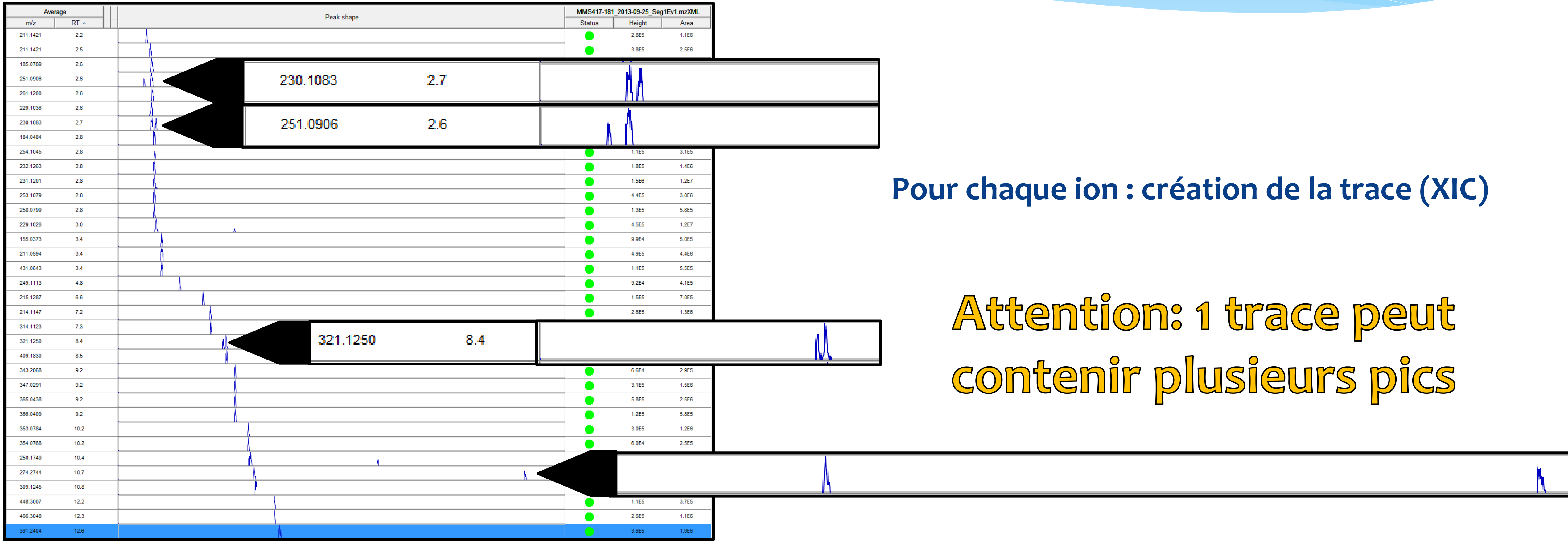


Recherche spectre après spectres de l'ensemble des ions ayant une intensité supérieur à la limite.

Liste d'ions d'intérêt (n° spectre × *m/z*)

La détection de pic

ETAPE II : création des chromatogrammes



Average			MMS417-181_2013-09-25_Seg1Ev1.mzXML		
m/z	RT	Peak shape	Status	Height	Area
211.1421	2.2		●	2.8E5	1.1E6
211.1421	2.5		●	3.8E5	2.5E6
185.0789	2.6				
251.0906	2.6				
261.1200	2.6				
229.1036	2.6				
230.1083	2.7				
184.0484	2.8				
254.1045	2.8		●	1.1E5	3.1E5
232.1263	2.8		●	1.8E5	1.4E6
231.1201	2.8		●	1.5E6	1.2E7
253.1079	2.8		●	4.4E5	3.0E6
258.0799	2.8		●	1.3E5	5.8E5
229.1026	3.0		●	4.5E5	1.2E7
155.0373	3.4		●	9.9E4	5.0E5
211.0594	3.4		●	4.9E5	4.4E6
431.0643	3.4		●	1.1E5	5.5E5
249.1113	4.8		●	9.2E4	4.1E5
215.1287	6.6		●	1.5E5	7.0E5
214.1147	7.2		●	2.6E5	1.3E6
314.1123	7.3				
321.1250	8.4				
409.1830	8.5				
343.2068	9.2		●	6.6E4	2.9E5
347.0291	9.2		●	3.1E5	1.5E6
365.0438	9.2		●	5.8E5	2.5E6
366.0409	9.2		●	1.2E5	5.8E5
353.0784	10.2		●	3.0E5	1.2E6
354.0768	10.2		●	6.0E4	2.5E5
250.1749	10.4				
274.2744	10.7				
309.1245	10.8				
448.3007	12.2		●	1.1E5	3.7E5
466.3048	12.3		●	2.6E5	1.1E6
391.2404	12.6		●	3.6E5	1.9E6

230.1083 2.7

251.0906 2.6

321.1250 8.4

Pour chaque ion : création de la trace (XIC)

Attention: 1 trace peut contenir plusieurs pics

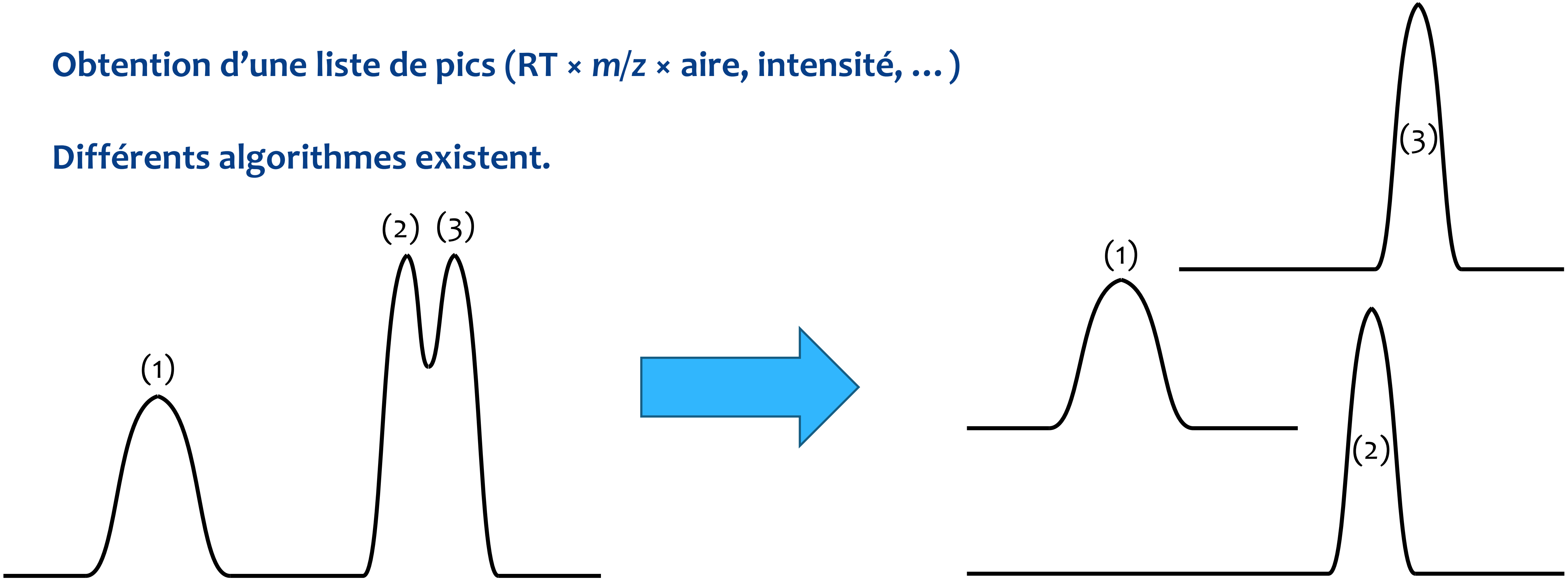
Détection des Ions d'intérêt | **Création des chromatogrammes** | Déconvolution des chromatogrammes | Alignement des pics

La détection de pic

ETAPE III : Déconvolution des chromatogrammes

Obtention d'une liste de pics (RT × *m/z* × aire, intensité, …)

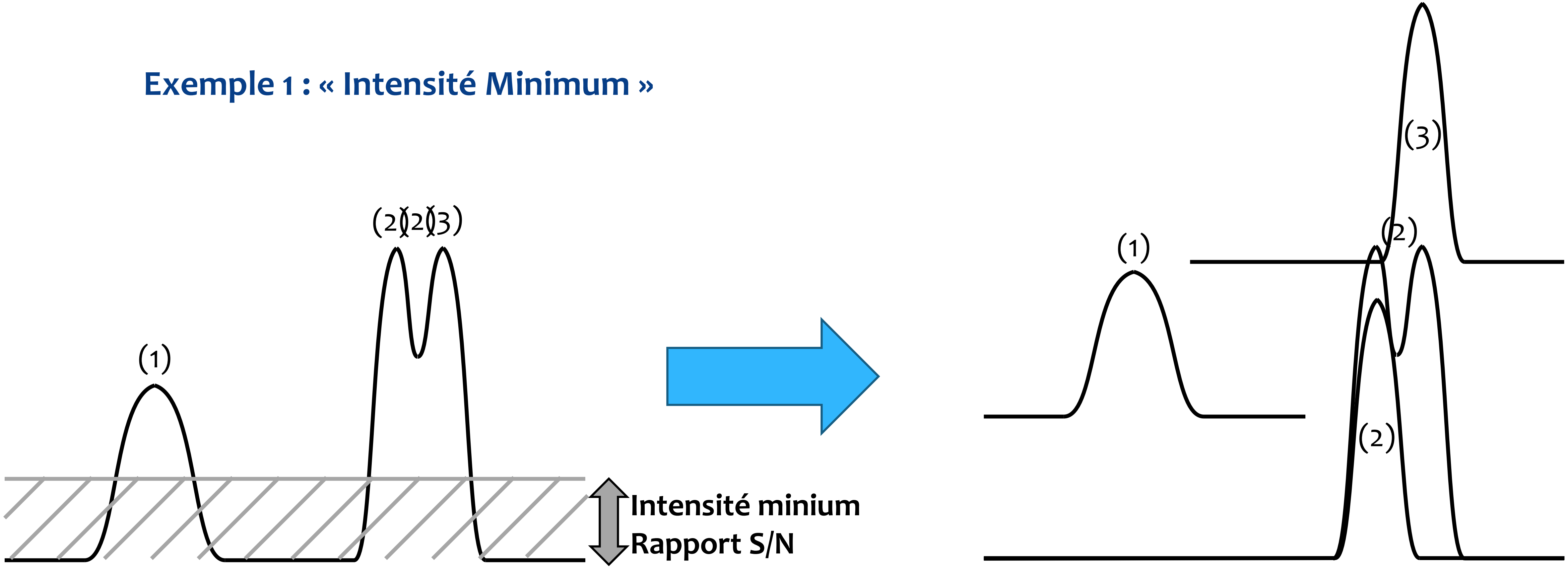
Différents algorithmes existent.



La détection de pic

ETAPE III : Déconvolution des chromatogrammes

Exemple 1 : « Intensité Minimum »

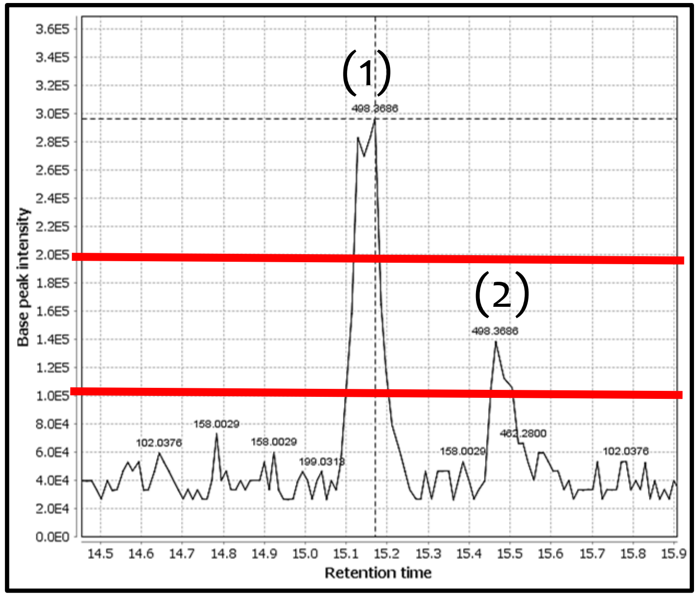


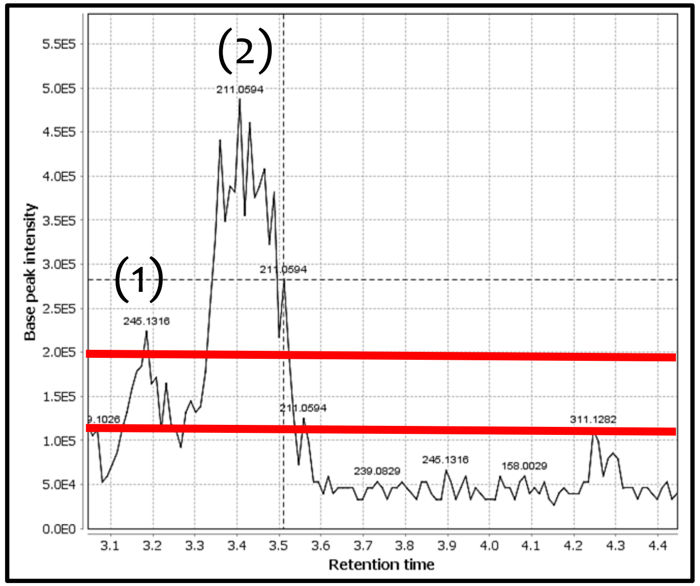
Intensité minium
Rapport S/N

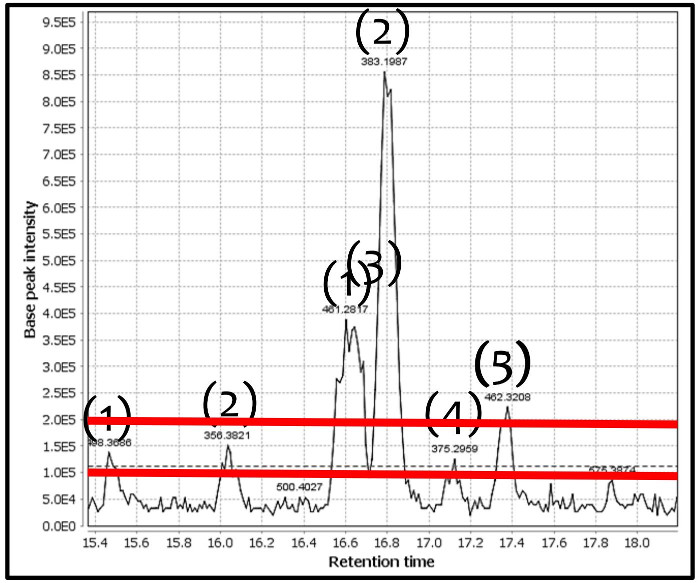
La détection de pic

ETAPE III : Déconvolution des chromatogrammes

Exemple 1 : « Intensité Minimum »



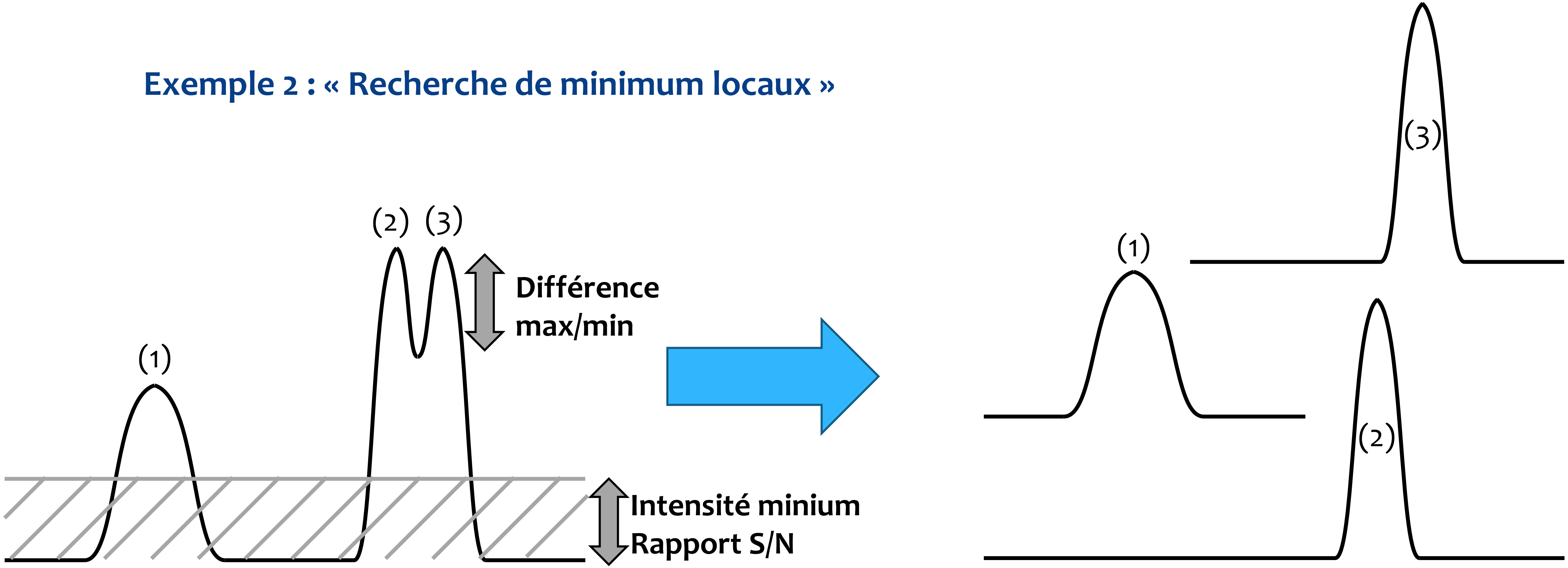




La détection de pic

ETAPE III : Déconvolution des chromatogrammes

Exemple 2 : « Recherche de minimum locaux »

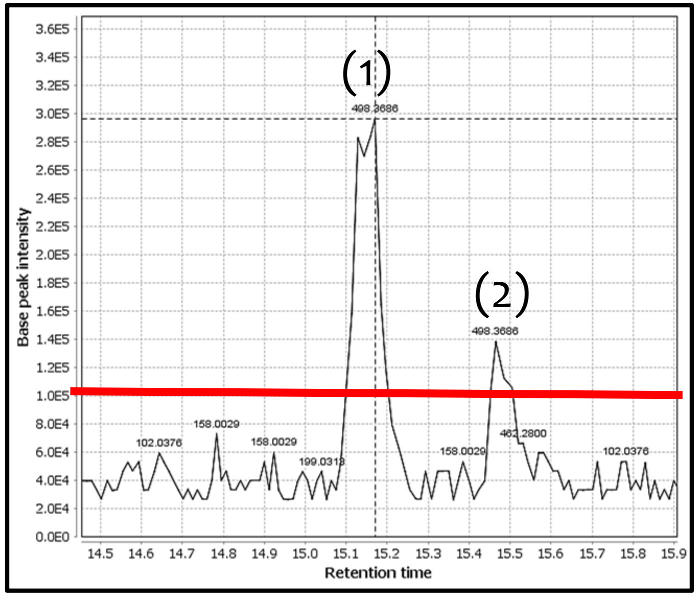


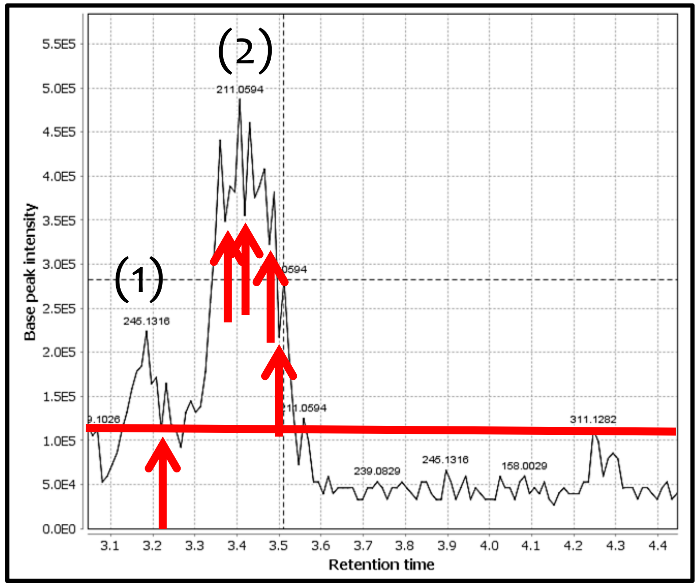
La détection de pic

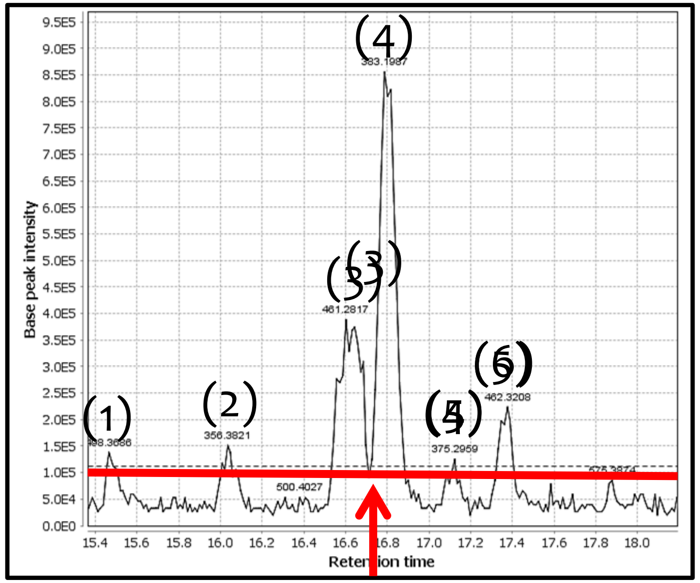
ETAPE III : Déconvolution des chromatogrammes

Exemple 2 : « Recherche de minimum locaux »

Les paramètres sélectionnés sont cruciaux !



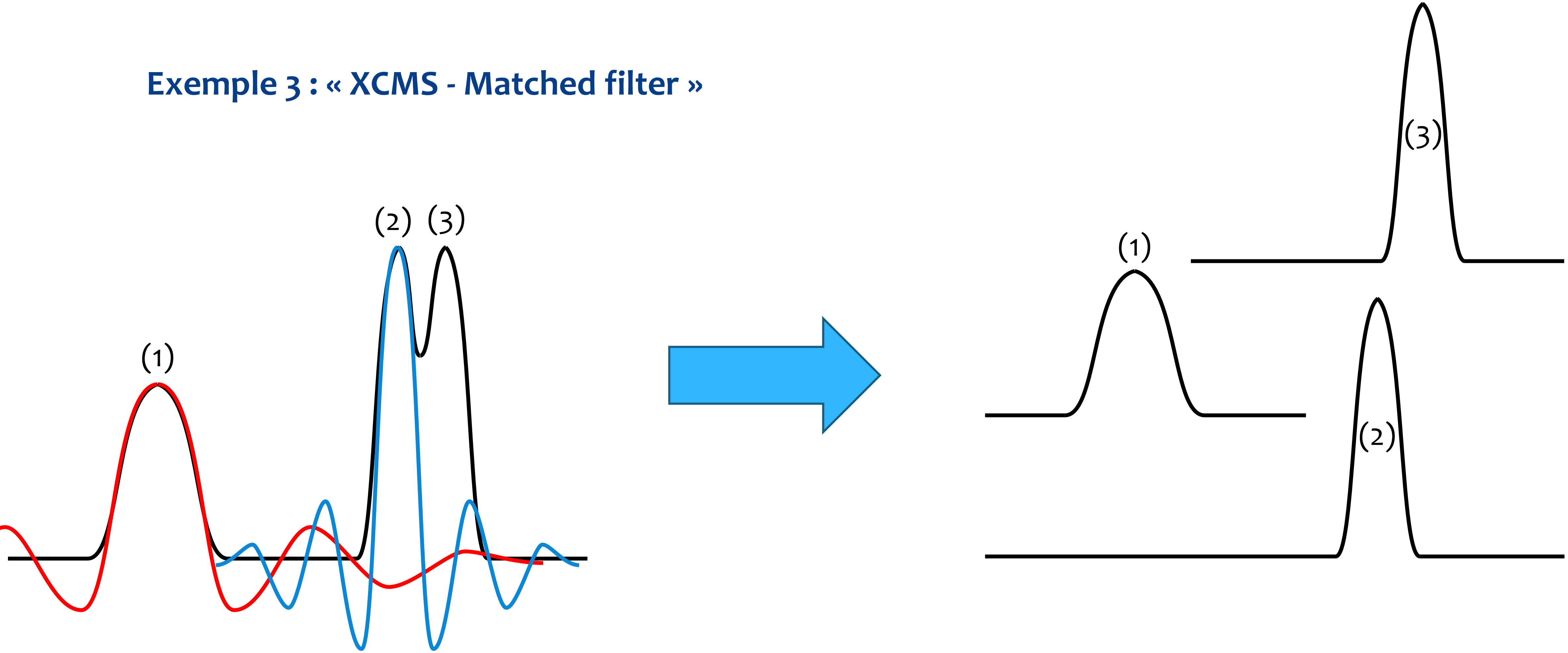




La détection de pic

ETAPE III : Déconvolution des chromatogrammes

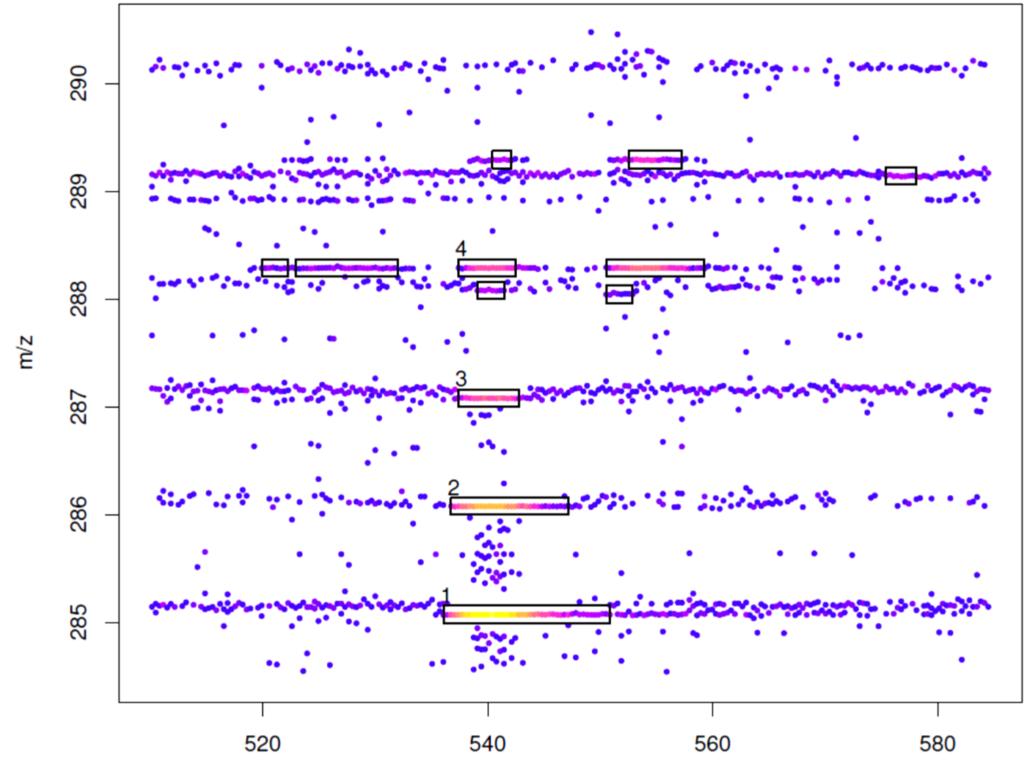
Exemple 3 : « XCMS - Matched filter »



La détection de pic

ETAPE III : Déconvolution des chromatogrammes

Exemple 4 : « XCMS - CentWave »



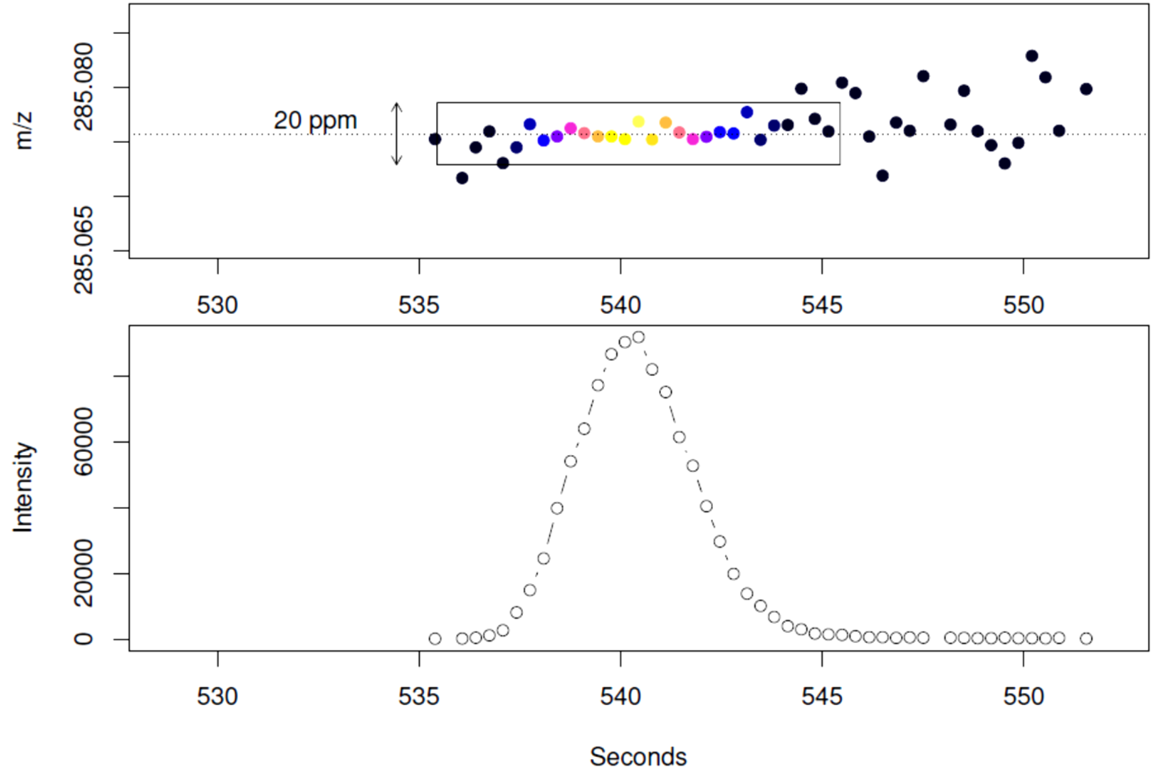
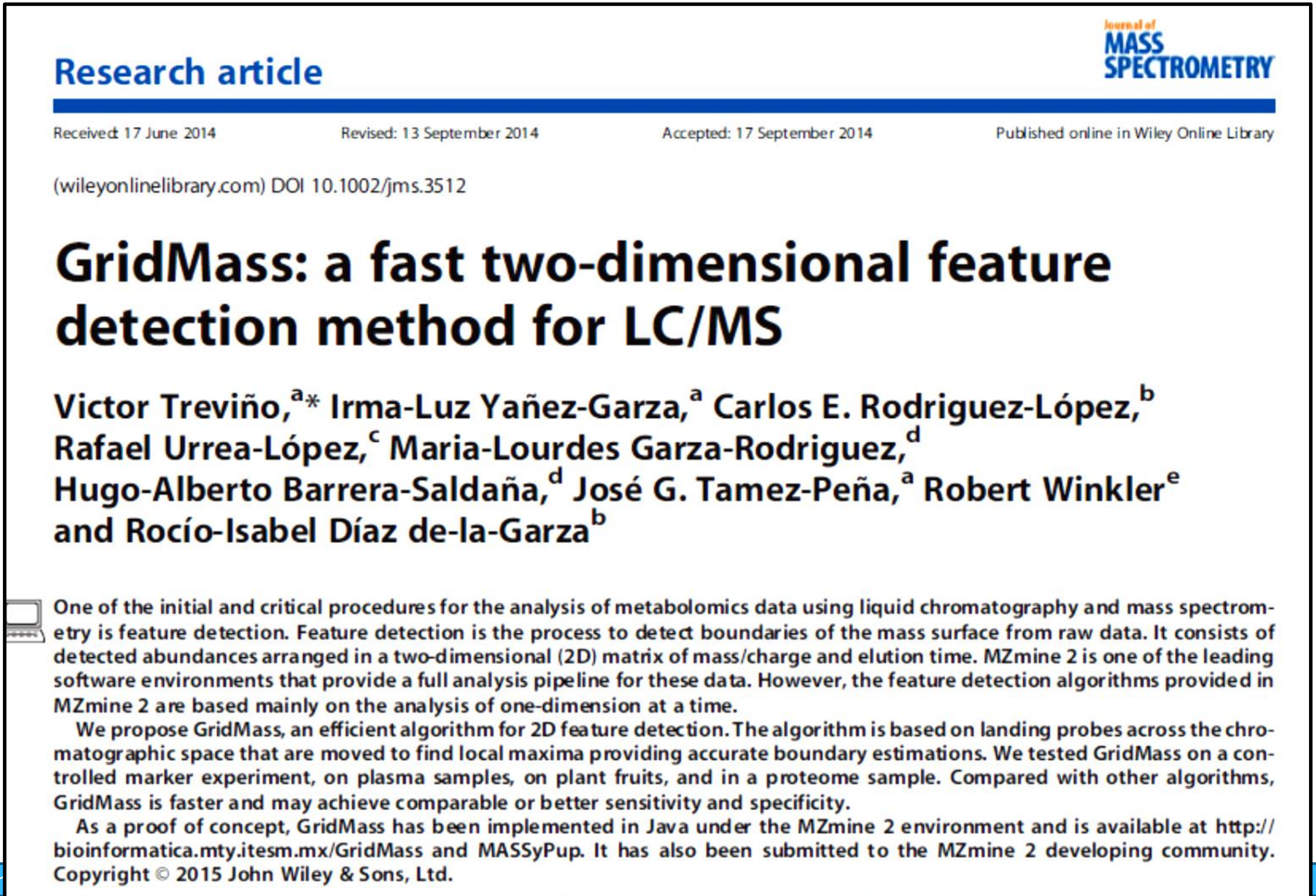


Figure 1
Mass trace and chromatographic peak of Biochanin A $[M + H]^+$ mass signal. The upper panel shows the mass trace of the biochanin A $[M + H]^+$ mass signal across 10 seconds with colour-coded intensities. The corresponding chromatographic peak is shown below.

La détection de pic

ETAPE III : Déconvolution des chromatogrammes

Exemple 5 : « GridMass »



Research article

Journal of MASS SPECTROMETRY

Received: 17 June 2014 Revised: 13 September 2014 Accepted: 17 September 2014

(wileyonlinelibrary.com) DOI 10.1002/jms.3512

GridMass: a fast two-dimensional feature detection method for LC/MS

Victor Treviño,[a]* Irma-Luz Yañez-Garza,[a] Carlos E. Rodriguez-López,[b] Rafael Urrea-López,[c] Maria-Lourdes Garza-Rodriguez,[d] Hugo-Alberto Barrera-Saldaña,[d] José G. Tamez-Peña,[a] Robert Winkler[e] and Rocío-Isabel Díaz de-la-Garza[b]

One of the initial and critical procedures for the analysis of metabolomics data using liquid chromatography and mass spectrometry is feature detection. Feature detection is the process to detect boundaries of the mass surface from raw data. It consists of detected abundances arranged in a two-dimensional (2D) matrix of mass/charge and elution time. MZmine 2 is one of the leading software environments that provide a full analysis pipeline for these data. However, the feature detection algorithms provided in MZmine 2 are based mainly on the analysis of one-dimension at a time.

We propose GridMass, an efficient algorithm for 2D feature detection. The algorithm is based on landing probes across the chromatographic space that are moved to find local maxima providing accurate boundary estimations. We tested GridMass on a controlled marker experiment, on plasma samples, on plant fruits, and in a proteome sample. Compared with other algorithms, GridMass is faster and may achieve comparable or better sensitivity and specificity.

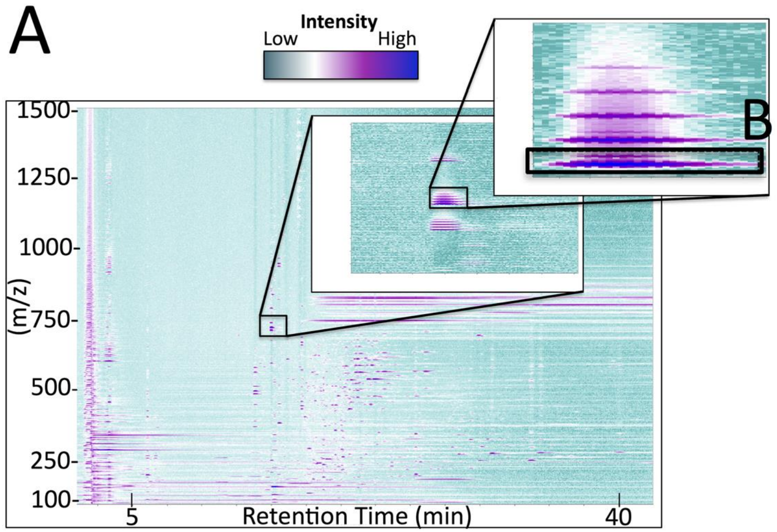
As a proof of concept, GridMass has been implemented in Java under the MZmine 2 environment and is available at http://bioinformatica.mty.itesm.mx/GridMass and MASSyPup. It has also been submitted to the MZmine 2 developing community.

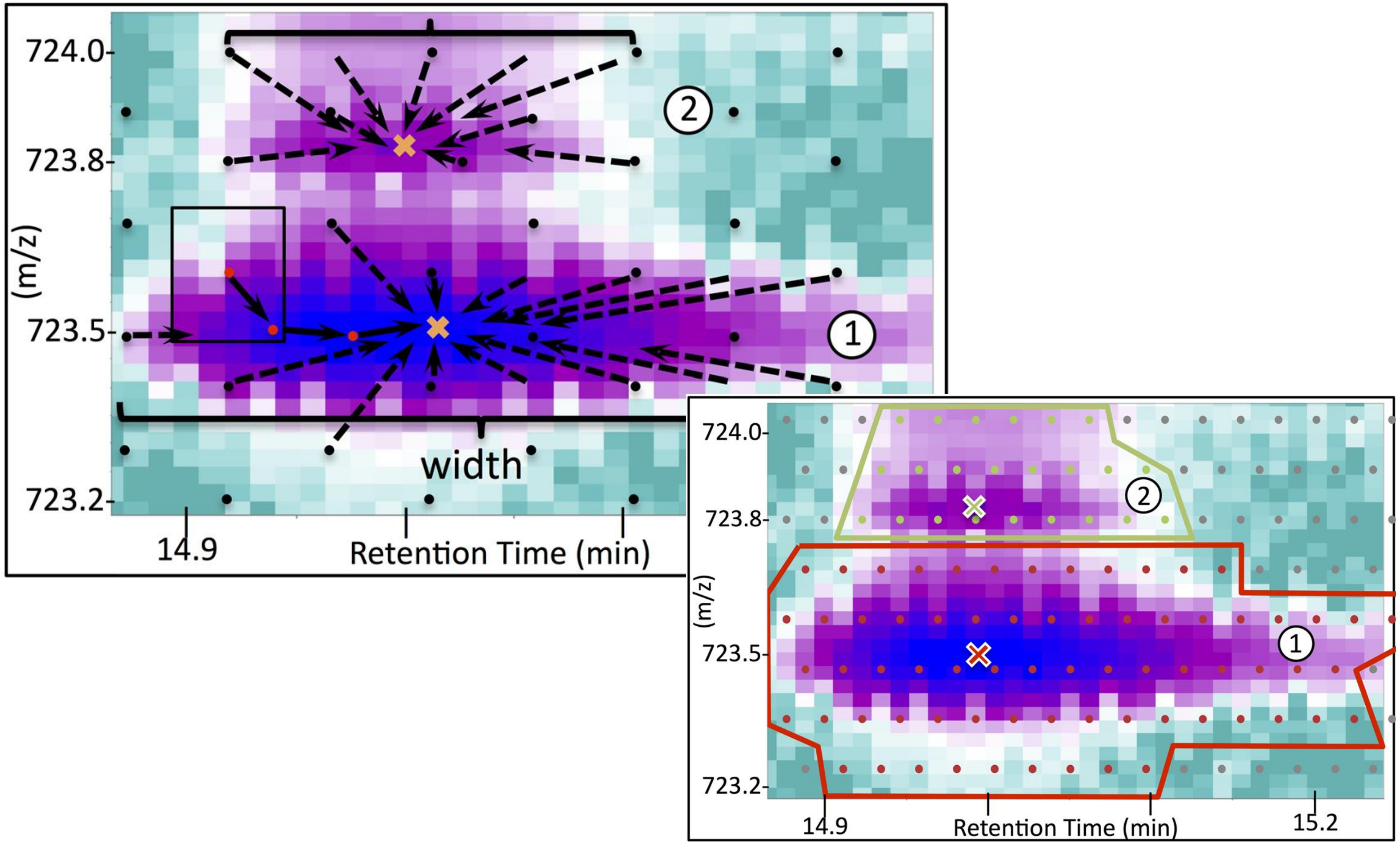
Treviño et al. J. Mass Spectrom. 2015 50 165-174

La détection de pic

ETAPE III : Déconvolution des chromatogrammes

Exemple 5 : « GridMass »



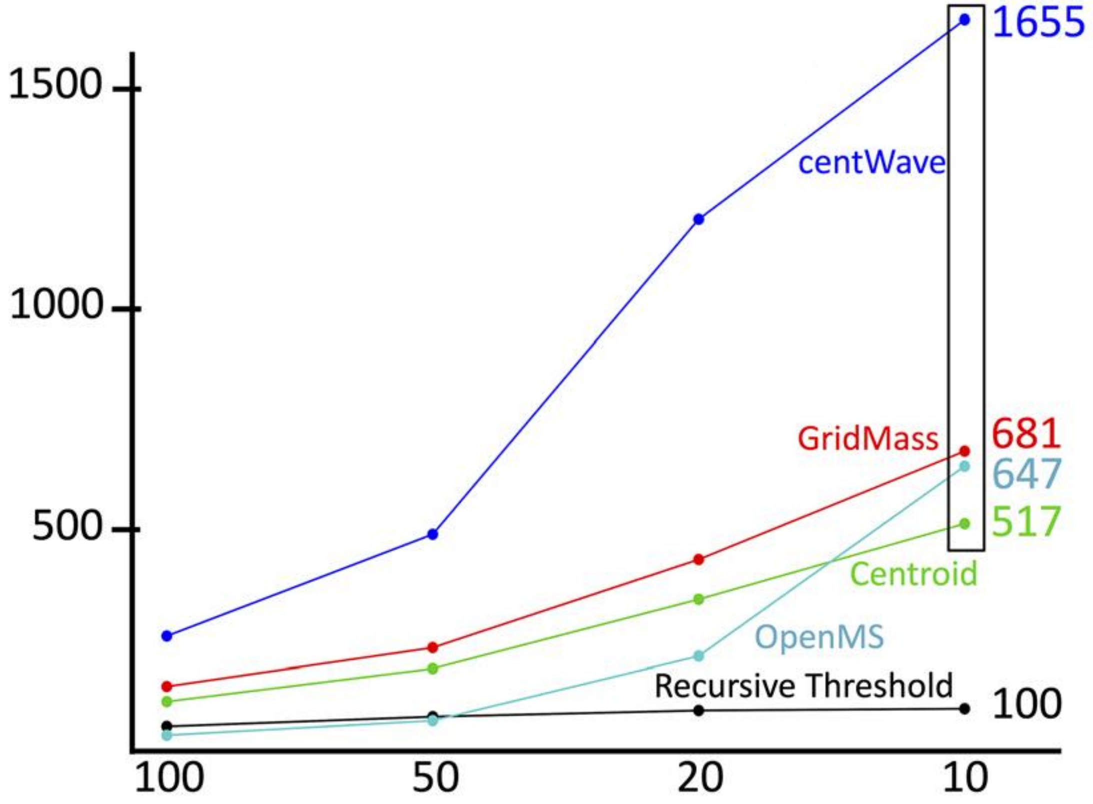


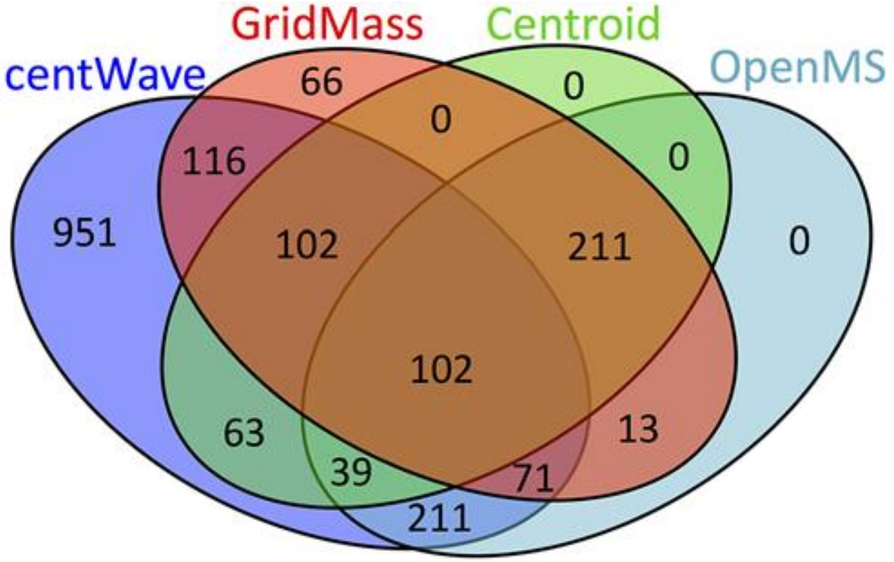
Assez rapide !

La détection de pic

ETAPE III : Déconvolution des chromatogrammes

Quel algorithme ?



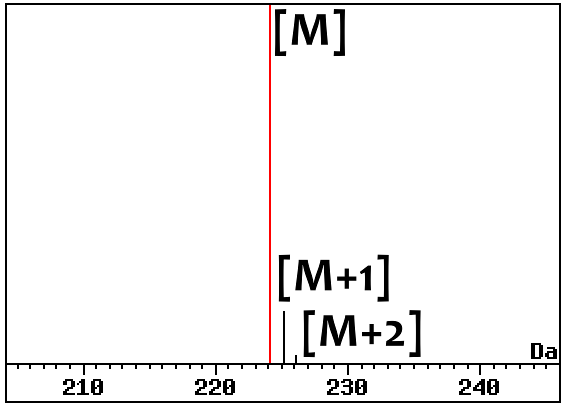


La détection de pic

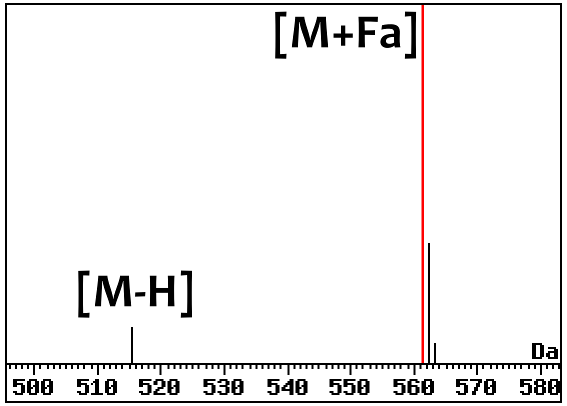
ETAPE IV : Alignement des pics

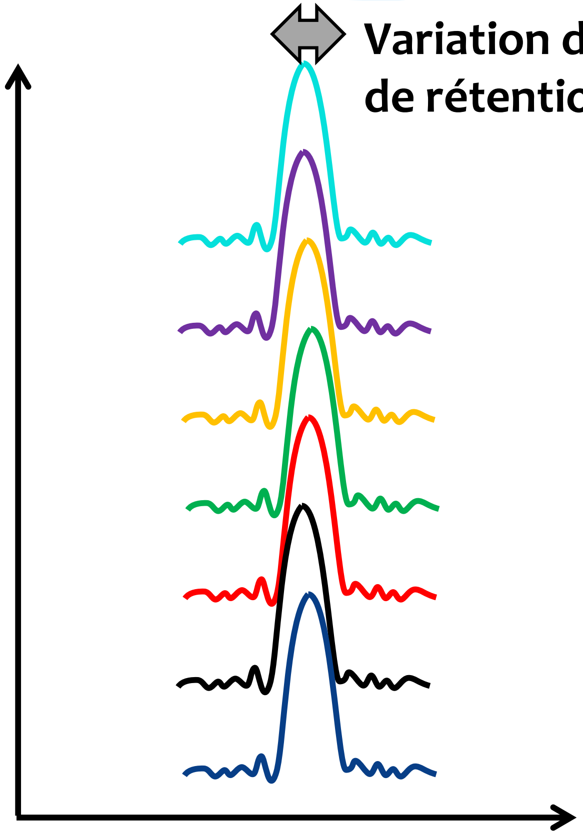
Plusieurs pics (*m*/*z* × RT) peuvent avoir le « même » temps de rétention.

isotopes



adduits





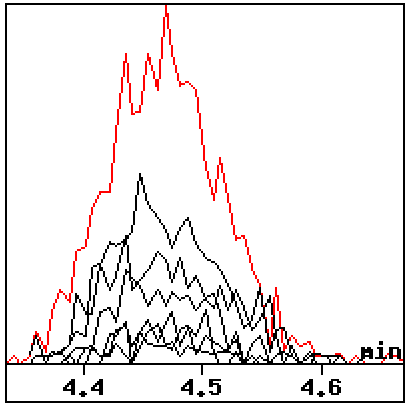
Variation des temps de rétention

Il peut y avoir quelques variations minimes.

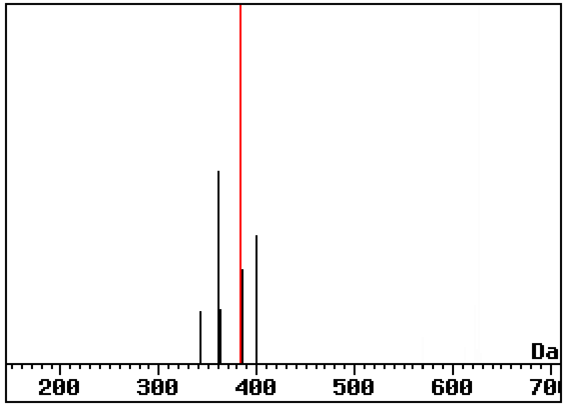
La détection de pic

ETAPE IV : Alignement des pics

Regroupement pour former des pseudo-spectres.



Pics regroupés



Pseudo spectre associé

Analyse automatique des pseudo-spectres

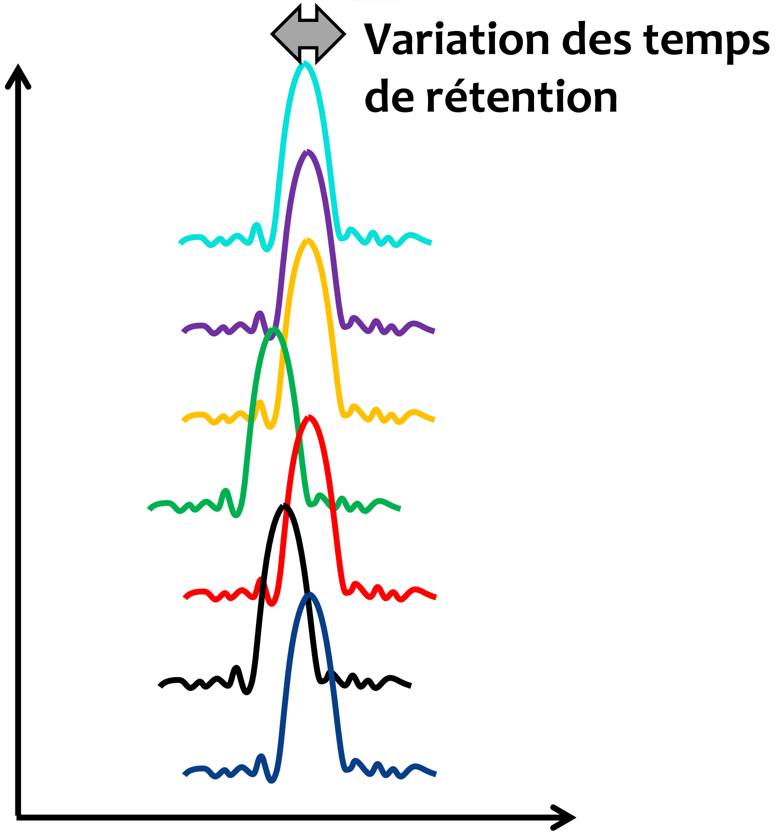
La détection de pic

ETAPE IV : Alignement des pics

Nécessaire lorsque l'on compare des chromatogrammes.

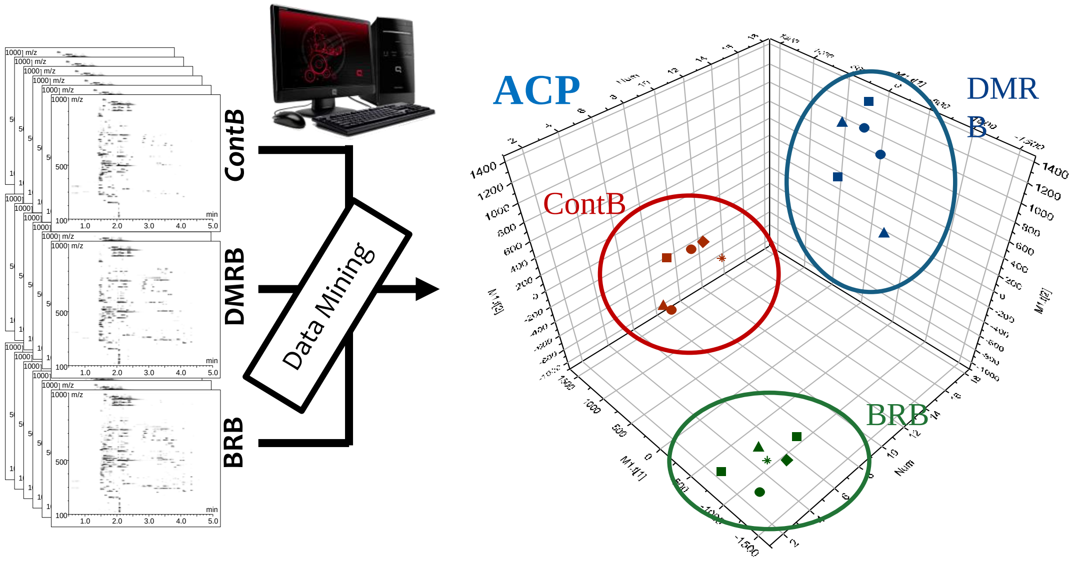
Variation entre les chromatogrammes : faible mais bien réelle

Différents algorithmes existent en fonction des logiciels.



Le cas de la métabolomique différentielle

ETAPE V :



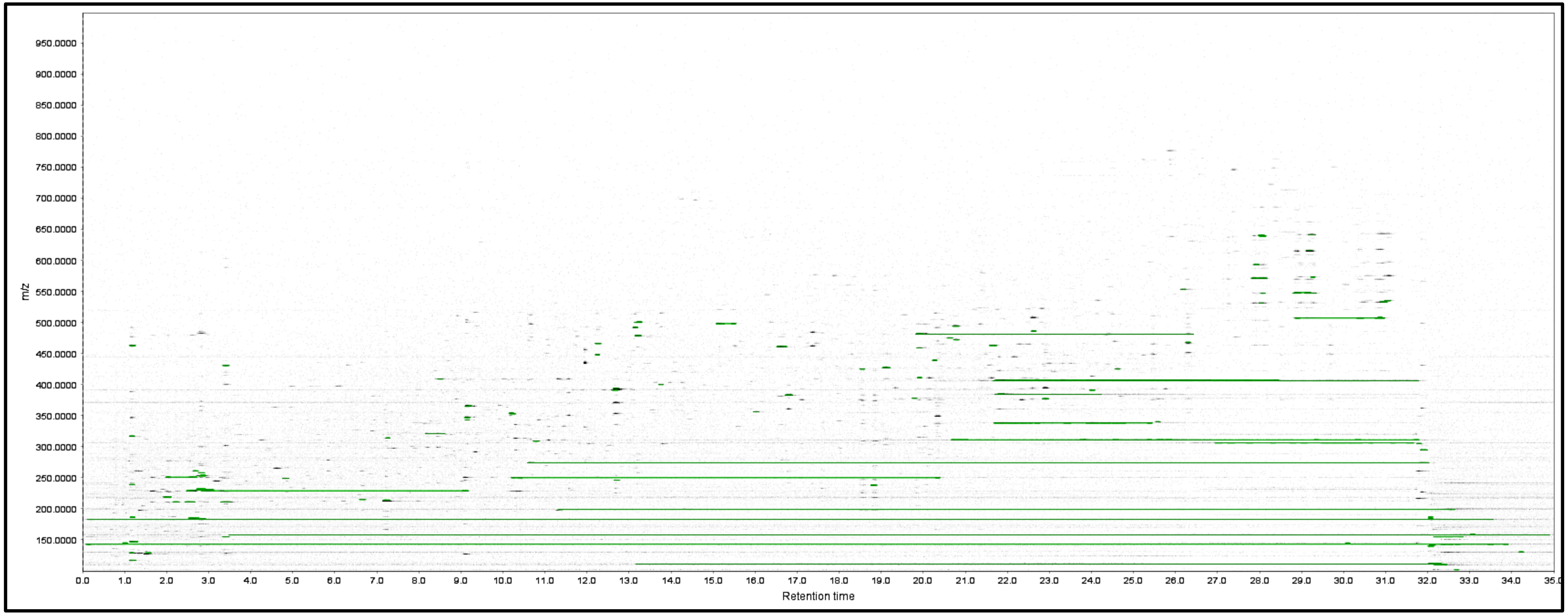
Les statistiques n'aime pas les ZERO !

Il faut impérativement faire une étape de remplissage des zéros de la matrice.

MZmine: Gapfilling **XCMS: Fillpeaks**

La détection de pic

Comment faire une bonne détection de pics ?



Comment faire une bonne détection de pics ?

Comment faire une bonne détection de pics ?

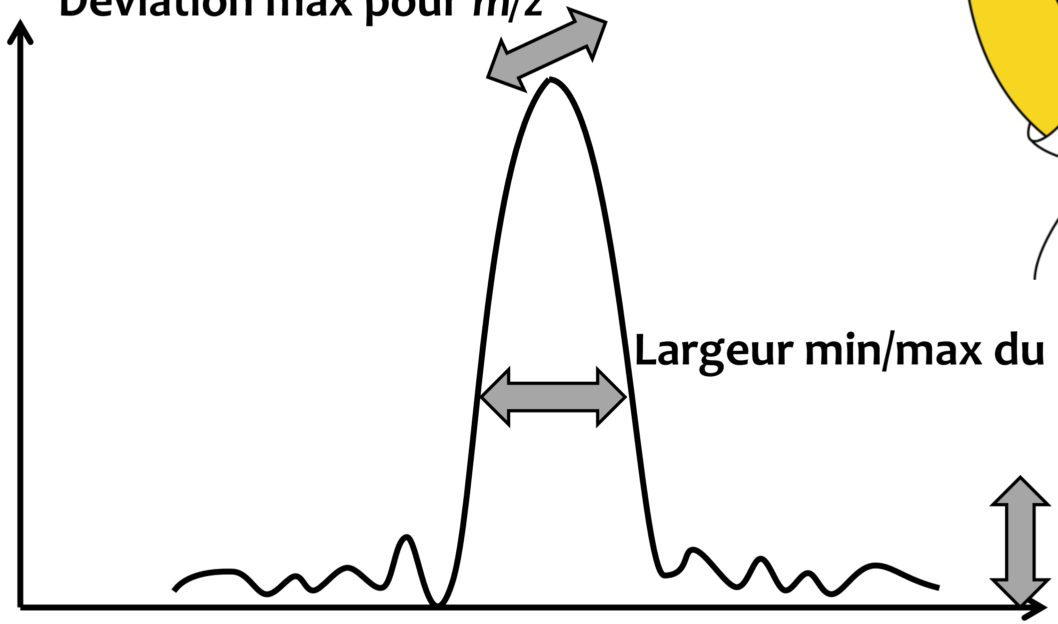
Comment faire une bonne détection de pics ?

Quel algorithme ? **Testez et faite vous votre idée !**

Quel paramètres ?

Testez et faite vous votre idée !

MAIS restez logiques !

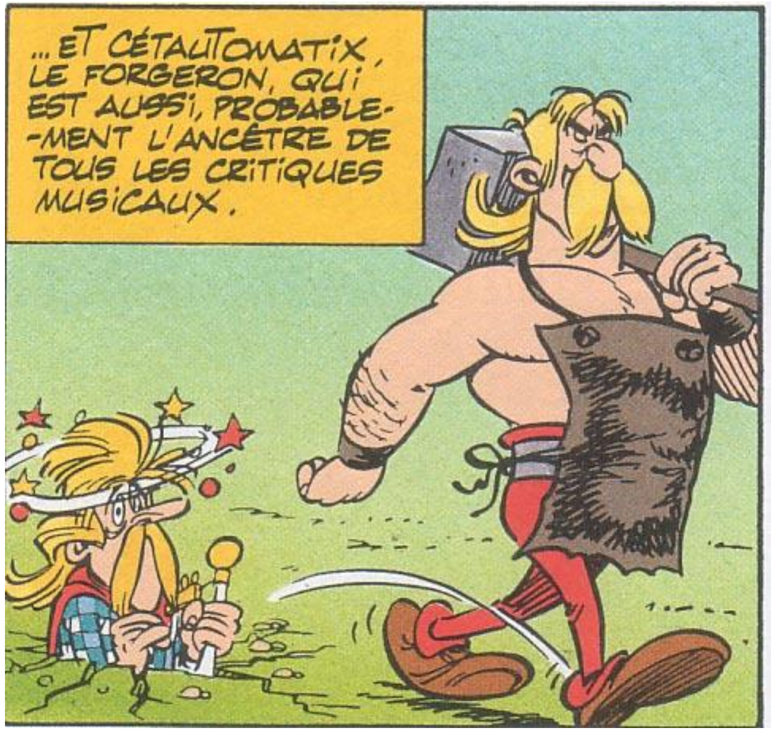


Déviation max pour *m/z*

Largeur min/max du pic

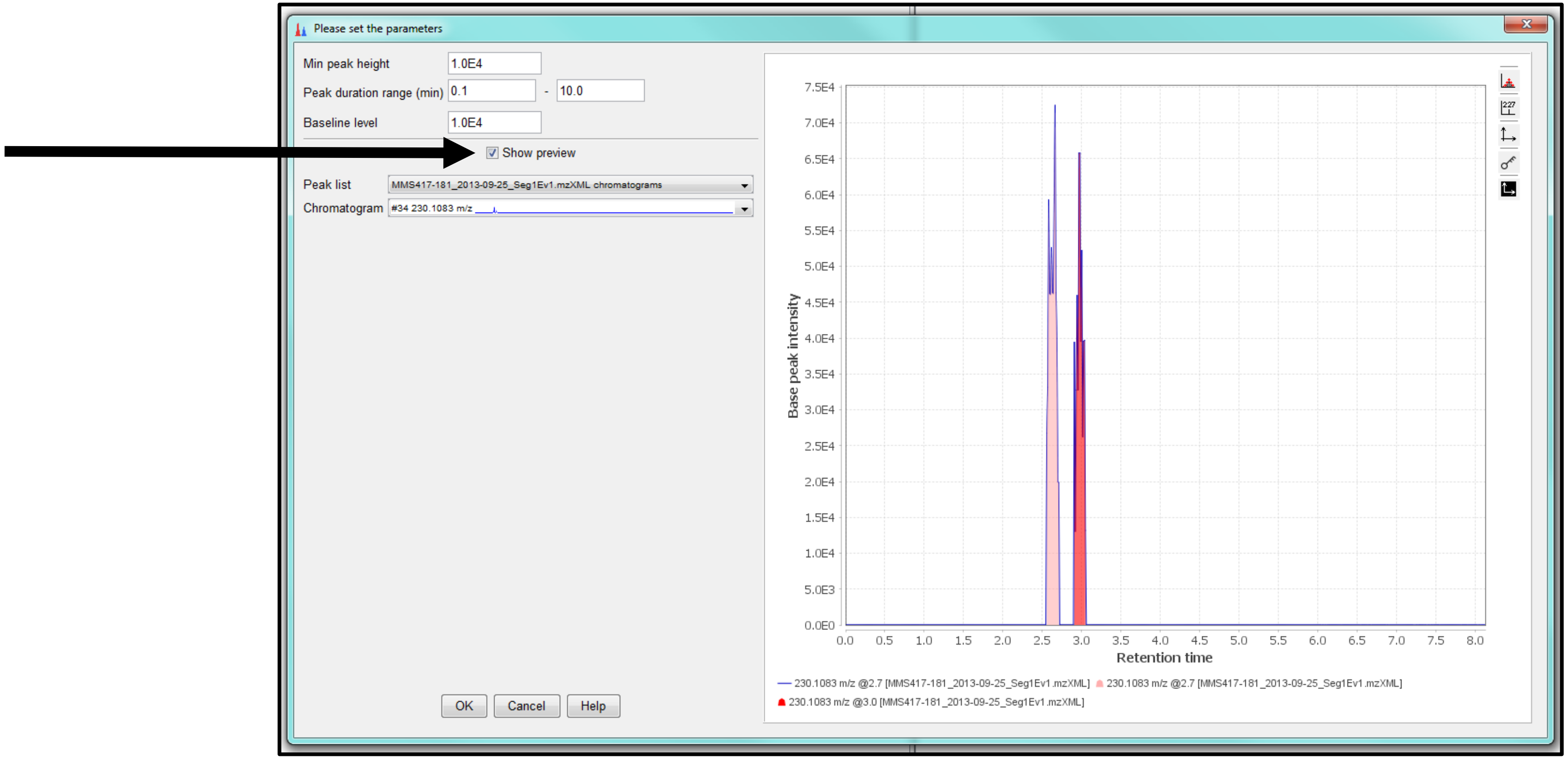
Intensité minium
Rapport S/N

C'est en forgeant que l'on devient forgeron !



Comment faire une bonne détection de pics avec Mzmine 2 ?

Avantages de MZmine 2



Please set the parameters

Min peak height: 1.0E4

Peak duration range (min): 0.1 - 10.0

Baseline level: 1.0E4

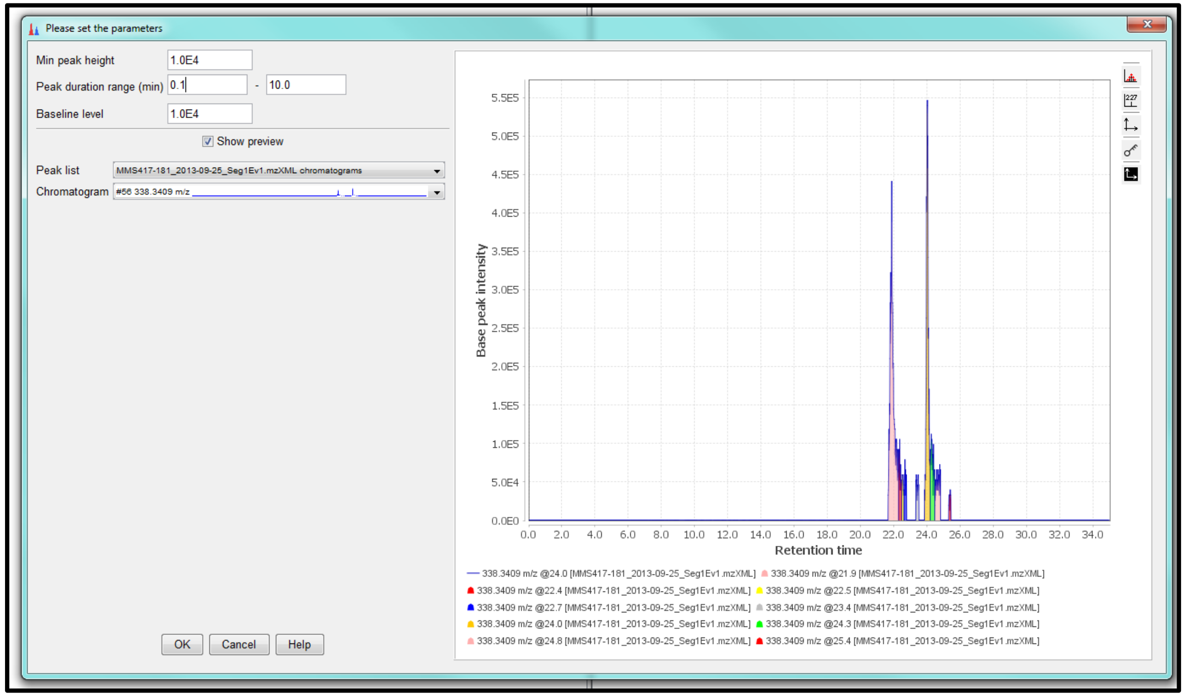
☑ Show preview

Peak list: MMS417-181_2013-09-25_Seg1Ev1.mzXML chromatograms

Chromatogram: #34 230.1083 m/z

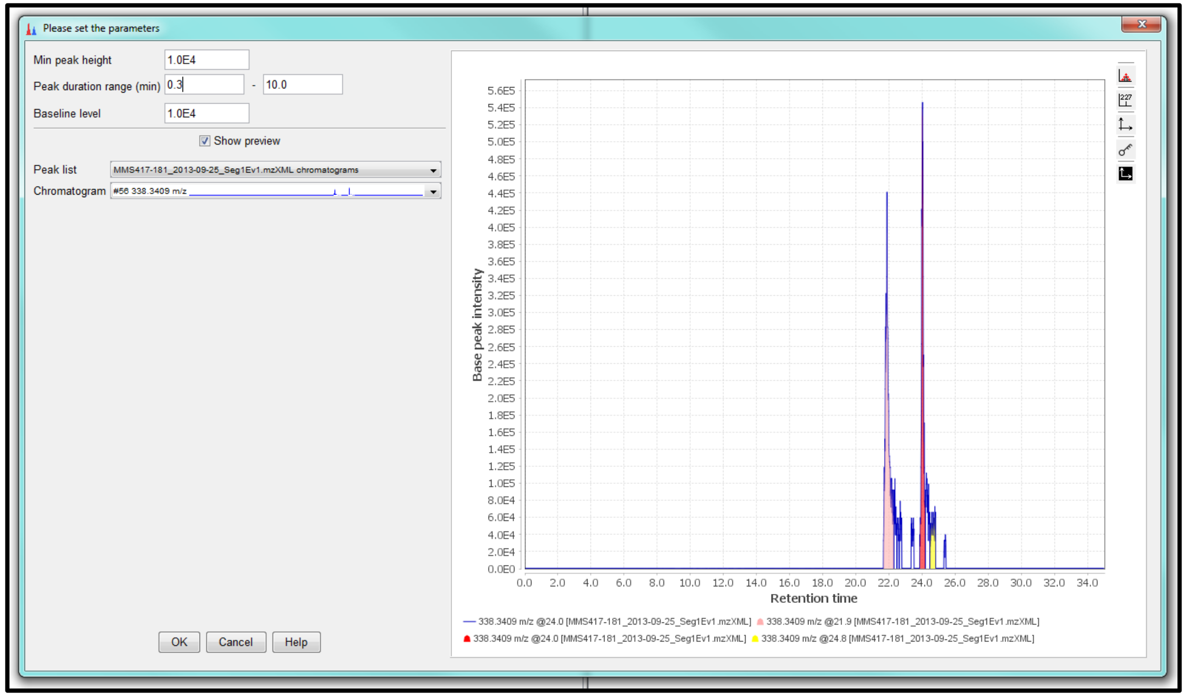
Comment faire une bonne détection de pics avec Mzmine 2 ?

Possibilité de voir l'impact des paramètres sur le résultat à chaque étape.



Please set the parameters

Min peak height: 1.0E4
Peak duration range (min): 0.1 - 10.0
Baseline level: 1.0E4
☑ Show preview
Peak list: MMS417-181_2013-09-25_Seg1Ev1.mzXML chromatograms
Chromatogram: #56 338.3409 m/z



Please set the parameters

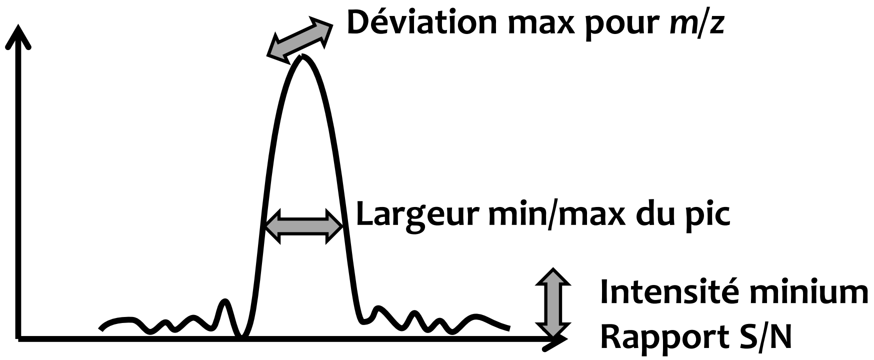
Min peak height: 1.0E4
Peak duration range (min): 0.3 - 10.0
Baseline level: 1.0E4
☑ Show preview
Peak list: MMS417-181_2013-09-25_Seg1Ev1.mzXML chromatograms
Chromatogram: #56 338.3409 m/z

Comment faire une bonne détection de pics avec XCMS ?

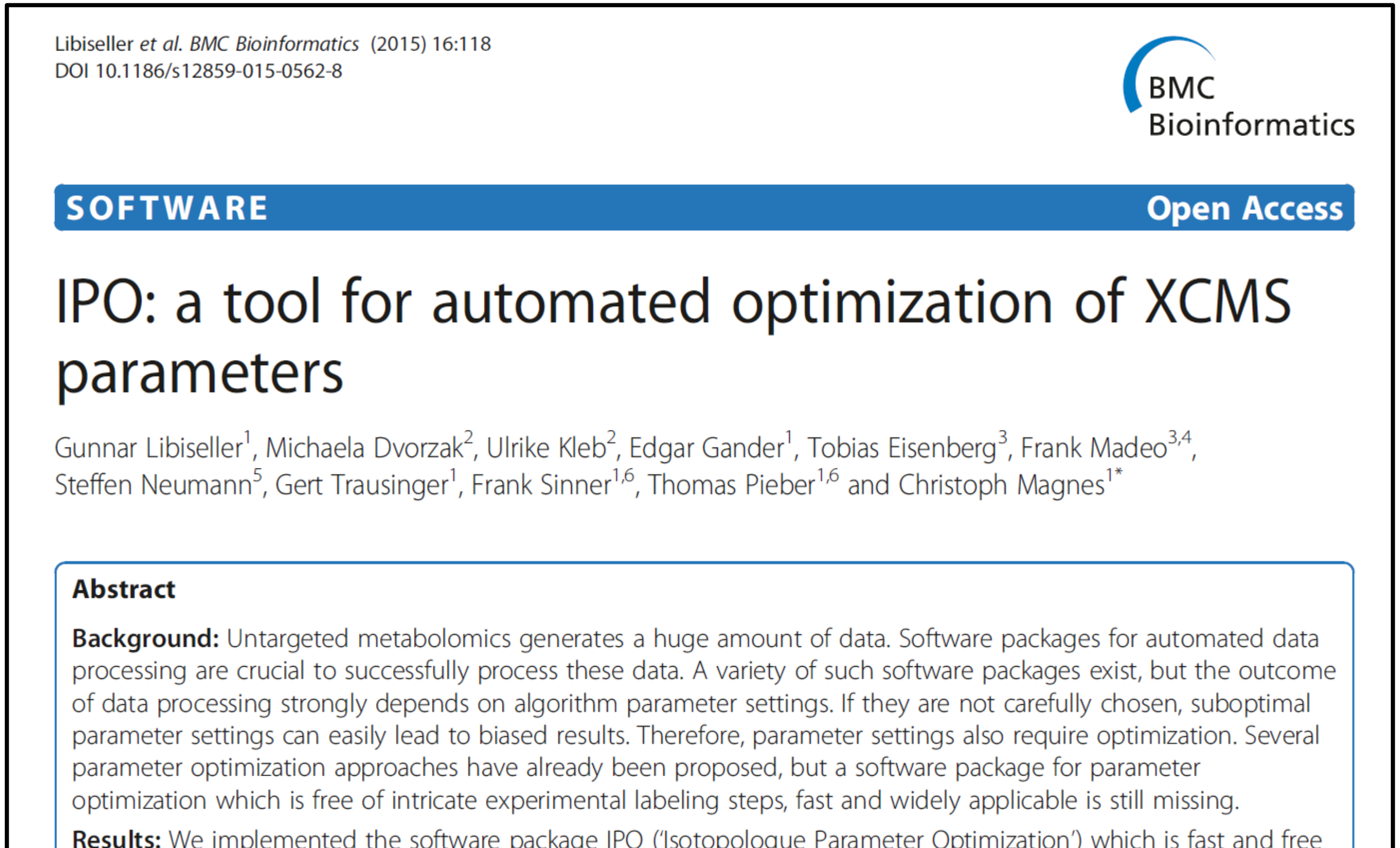
Comment faire une bonne détection de pics ?

Une solution pour XCMS : IPO

Optimisation des paramètres de pics



Optimisation des paramètres d’alignement


Libiseller *et al. BMC Bioinformatics* (2015) 16:118
DOI 10.1186/s12859-015-0562-8

BMC Bioinformatics

SOFTWARE **Open Access**

IPO: a tool for automated optimization of XCMS parameters

Gunnar Libiseller[1], Michaela Dvorzak[2], Ulrike Kleb[2], Edgar Gander[1], Tobias Eisenberg[3], Frank Madeo[3,4], Steffen Neumann[5], Gert Trausinger[1], Frank Sinner[1,6], Thomas Pieber[1,6] and Christoph Magnes[1*]

Abstract

Background: Untargeted metabolomics generates a huge amount of data. Software packages for automated data processing are crucial to successfully process these data. A variety of such software packages exist, but the outcome of data processing strongly depends on algorithm parameter settings. If they are not carefully chosen, suboptimal parameter settings can easily lead to biased results. Therefore, parameter settings also require optimization. Several parameter optimization approaches have already been proposed, but a software package for parameter optimization which is free of intricate experimental labeling steps, fast and widely applicable is still missing.

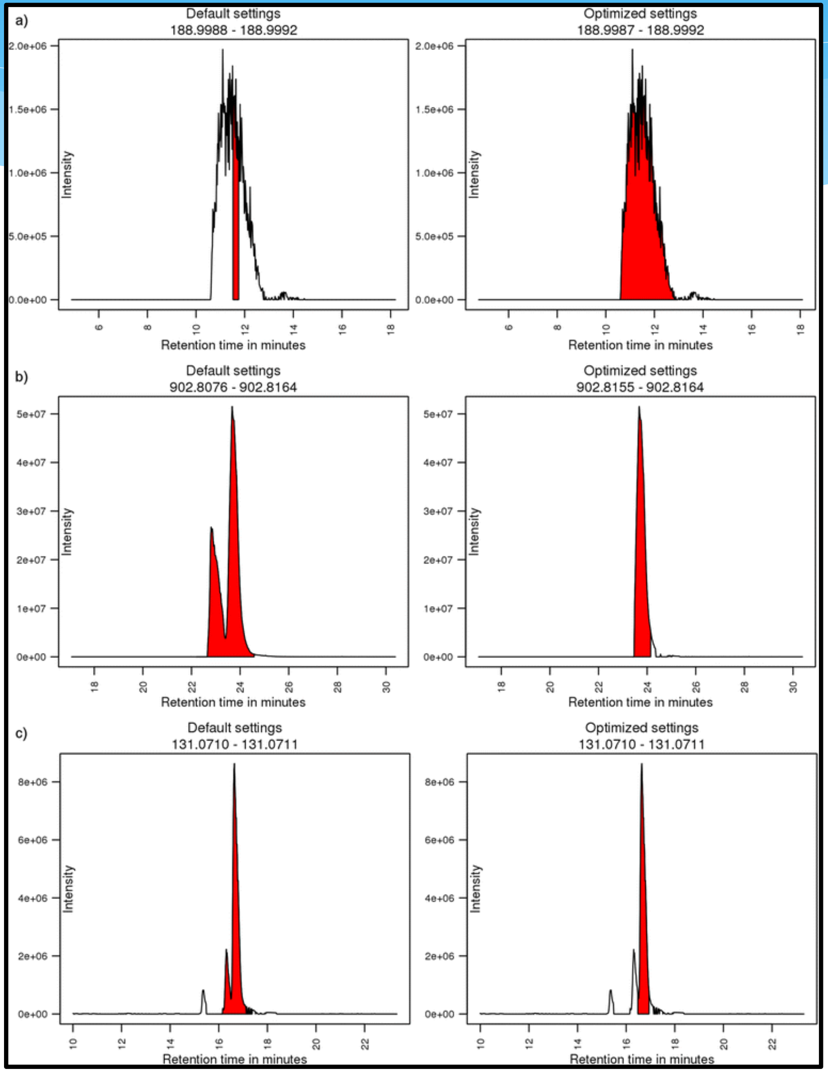
Results: We implemented the software package IPO (‘Isotopologue Parameter Optimization’) which is fast and free

La détection de pic

Comment faire une bonne détection de pics ?

Une solution pour XCMS : IPO

Un bon pic doit avoir des isotopes bien alignés !



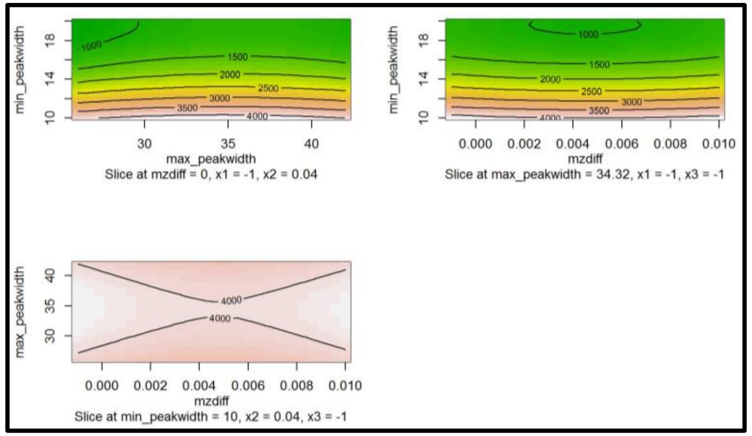
La détection de pic

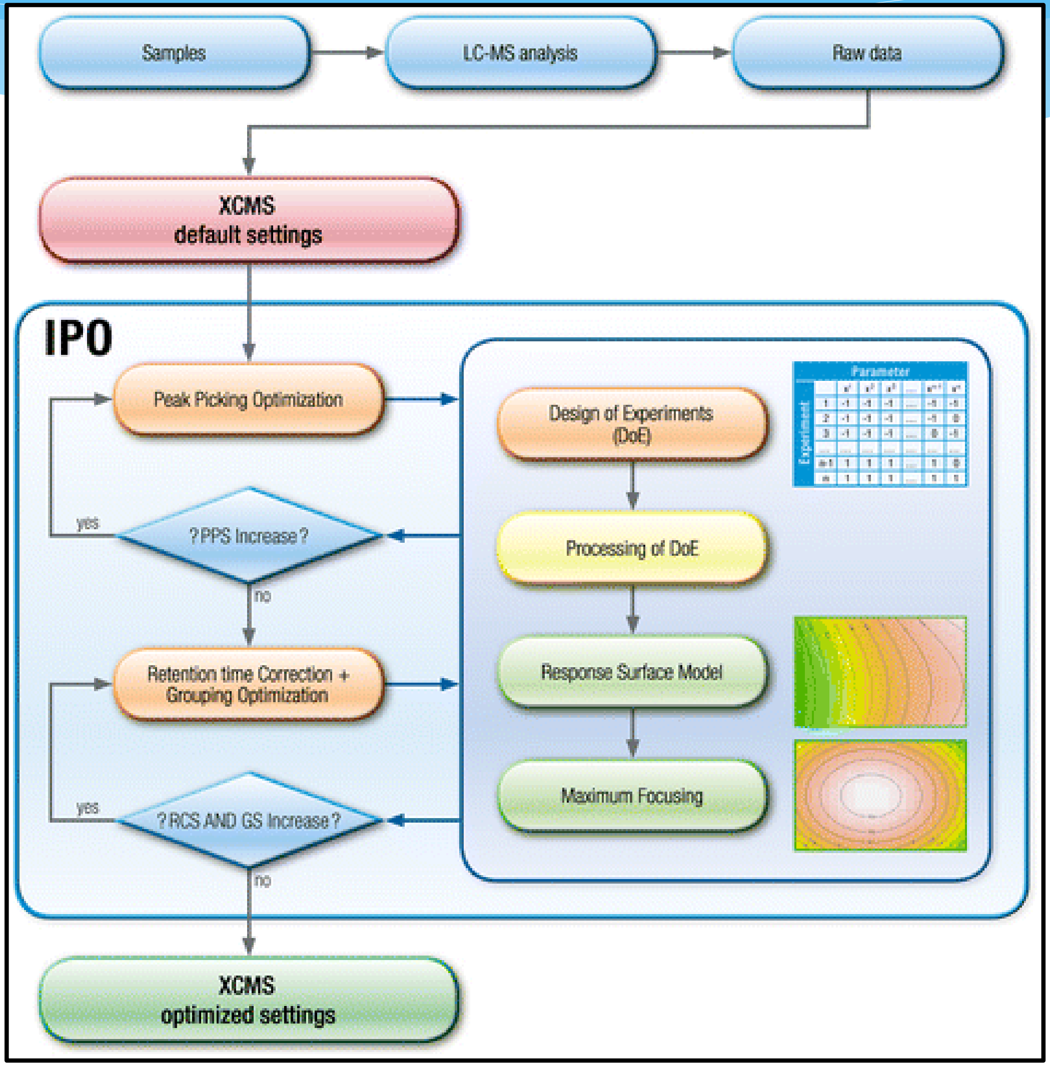
Comment faire une bonne détection de pics ?

Une solution pour XCMS : IPO

Utilisation des plans d'expérience

- optimisation de la déconvolution
- optimisation de la alignement





Samples

LC-MS analysis

Raw data

XCMS
default settings

IPO

Peak Picking Optimization

? PPS Increase ?

yes

no

Retention time Correction +
Grouping Optimization

? RCS AND GS Increase ?

yes

no

Design of Experiments
(DoE)

Processing of DoE

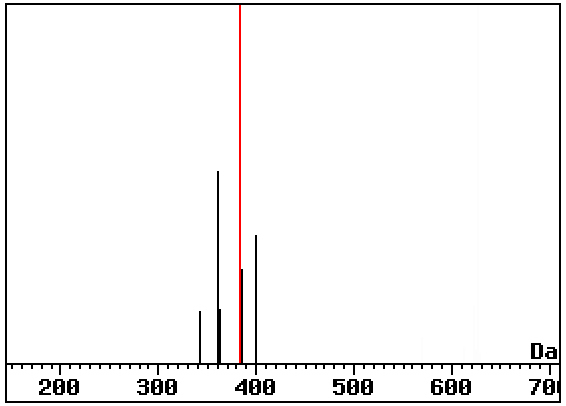
Response Surface Model

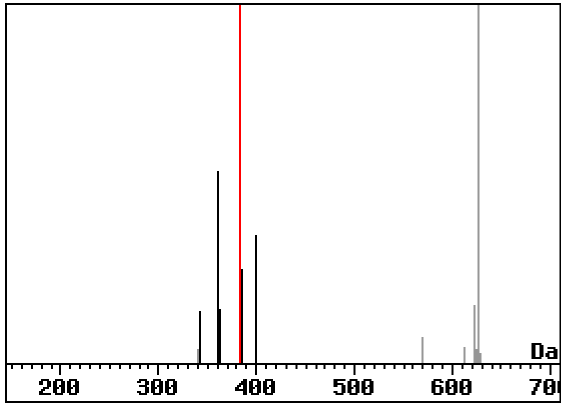
Maximum Focusing

XCMS
optimized settings

Avantages de la détection automatisé

1- faciliter l'analyse des donner



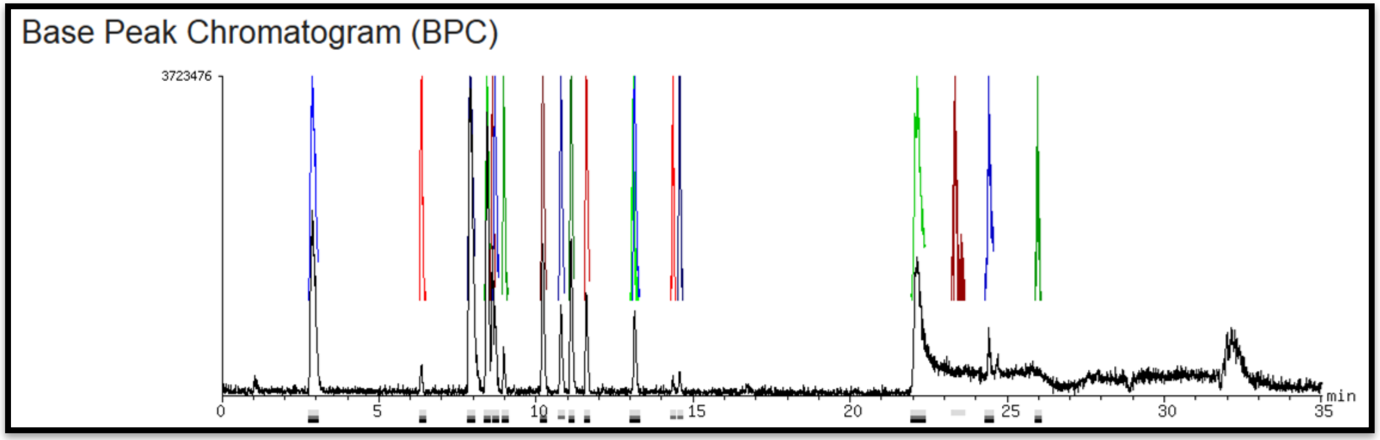


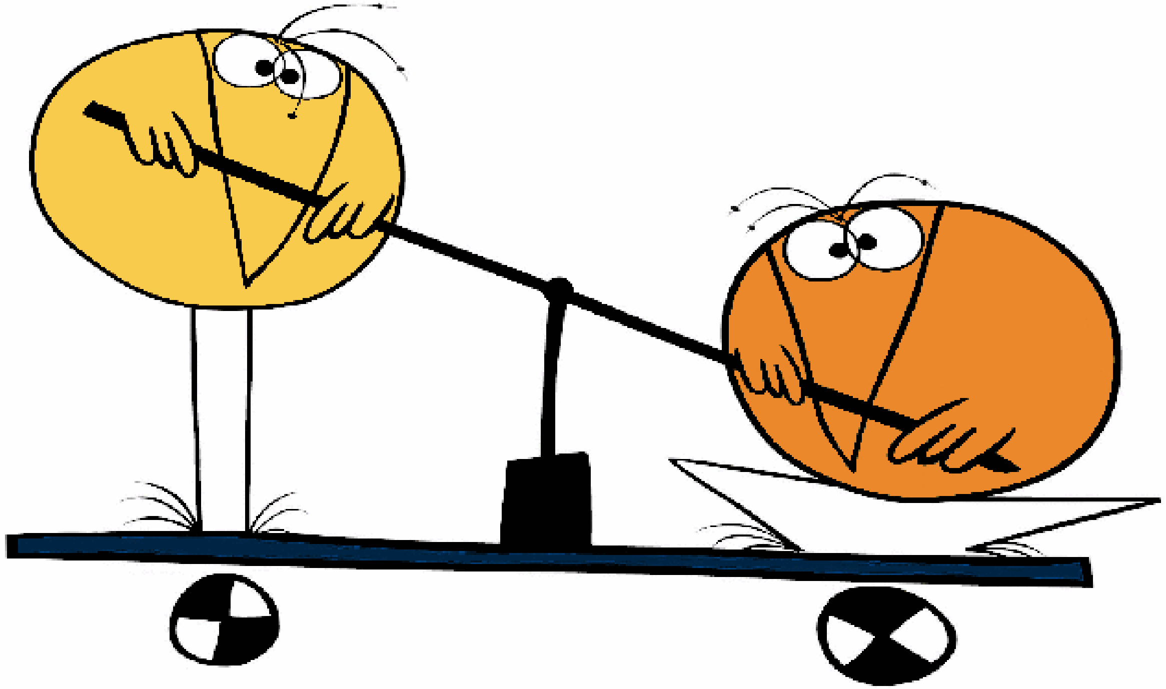
2- traitement homogène des chromatogrammes

(pas d'effet opérateur)

3- plus rapide si beaucoup de chromatogrammes à analyser

Bilan





N’ayez pas peur ca va bien se passer !

Merci

mms
mer molécules santé

UNIVERSITÉ DE NANTES

institut universitaire
mer et littoral
FR CNRS 3473

ThalassOMICS

Biogenouest

Analyse structurale et métabolomique

Plate-Forme
Corsaire