

# MassChroQ - Mass Chromatogram Quantification

Logiciel de quantification par spectrométrie de masse

Edlira Nano

Plateforme d'Analyse Protéomique de Paris Sud-Ouest



Colloque de l'UMR de Génétique Végétale  
5 Décembre 2011

1. Mais pourquoi avoir développé notre propre logiciel de quantification par LC-MS (Chromatographie Liquide - Spectrométrie de Masse) ?

# Plan

1. Mais pourquoi avoir développé notre propre logiciel de quantification par LC-MS (Chromatographie Liquide - Spectrométrie de Masse) ?
2. Mais qu'est que c'est la quantification par LC-MS ?

# Plan

1. Mais pourquoi avoir développé notre propre logiciel de quantification par LC-MS (Chromatographie Liquide - Spectrométrie de Masse) ?
2. Mais qu'est que c'est la quantification par LC-MS ?
3. Et sinon MassChroQ, il fait quoi exactement ?

1. Mais pourquoi avoir développé notre propre logiciel de quantification par LC-MS (Chromatographie Liquide - Spectrométrie de Masse) ?

Le développement de MassChroQ est étroitement lié à celui de la technique de quantification par LC-MS en protéomique en général, et à PAPPSO en particulier.

# Histoire de MassChroQ : la préhistoire

## Avant 2006

- L'électrophorèse bidimensionnelle (gels 2D) domine la quantification.
- La LC-MS sur des spots de gels 2D est utilisée pour l'identification des protéines uniquement.

# Histoire de MassChroQ : la préhistoire

## Avant 2006

- L'électrophorèse bidimensionnelle (gels 2D) domine la quantification.
- La LC-MS sur des spots de gels 2D est utilisée pour l'identification des protéines uniquement.

## 2006

Sur des mélanges peptidiques peu complexes :

- estimation de variations quantitatives relatives : technique de spectral counting ;
- extraction manuelle des valeurs quantitatives.

# Histoire de MassChroQ : la naissance

2007 - 2008

- Nouveaux appareils de LC-MS : échantillons plus complexes. Un saut de technologie s'opère.
- Premiers scripts Perl (B. Valot) : extraction automatique des valeurs quantitatives par mesure précise de l'intensité des pics. Les tests sur un premier projet révèlent une belle cinétique et une bonne précision. Leur utilisation reste toutefois marginale.



# Histoire de MassChroQ : la naissance

## 2007 - 2008

- Nouveaux appareils de LC-MS : échantillons plus complexes. Un saut de technologie s'opère.
- Premiers scripts Perl (B. Valot) : extraction automatique des valeurs quantitatives par mesure précise de l'intensité des pics. Les tests sur un premier projet révèlent une belle cinétique et une bonne précision. Leur utilisation reste toutefois marginale.

## 2009

- Échantillons de plus en plus complexes, besoin de rapidité.
- Traduction des scripts Perl en C++ (O. Langella, QuantiMsCpp). Algorithme Zivy de détection de pics.
- Fin 2009, QuantiMsCpp est de plus en plus utilisé sur des projets de la plateforme.

# Histoire de MassChroQ : la catalyse

2010

- HeterosYeast : premier projet d'analyses LC-MS/MS à haut débit (arrivée de M. Blein en Février).

# Histoire de MassChroQ : la catalyse

2010

- HeterosYeast : premier projet d'analyses LC-MS/MS à haut débit (arrivée de M. Blein en Février).
- Dromadair : deuxième projet conséquent d'analyses LC-MS/MS complexes (L. Bonhomme).

# Histoire de MassChroQ : la catalyse

2010

- HeterosYeast : premier projet d'analyses LC-MS/MS à haut débit (arrivée de M. Blein en Février).
- Dromadair : deuxième projet conséquent d'analyses LC-MS/MS complexes (L. Bonhomme).
- Les logiciels existants se révèlent limitants car :
  - souvent non compatibles avec les spectromètres de faible résolution ;
  - spécifiques à un type d'analyse (label-free ou marquage) ;
  - difficilement adaptables (paramétrables) à des analyses diverses.

# Histoire de MassChroQ : la catalyse

2010

- HeterosYeast : premier projet d'analyses LC-MS/MS à haut débit (arrivée de M. Blein en Février).
- Dromadair : deuxième projet conséquent d'analyses LC-MS/MS complexes (L. Bonhomme).
- Les logiciels existants se révèlent limitants car :
  - souvent non compatibles avec les spectromètres de faible résolution ;
  - spécifiques à un type d'analyse (label-free ou marquage) ;
  - difficilement adaptables (paramétrables) à des analyses diverses.
- QuantiMsCpp requiert un développement à temps plein (arrivée de E. Nano en Juin).

# Histoire de MassChroQ : le bond

2011

- QuantiMsCpp devient MassChroQ (Mass Chromatogram Quantification software).

# Histoire de MassChroQ : le bond

2011

- QuantiMsCpp devient MassChroQ (Mass Chromatogram Quantification software).
- MassChroQ : A versatile tool for mass spectrometry quantification, B. Valot, O. Langella, E. Nano, M. Zivy. Proteomics, 2011, 11 :3572–3577.

# Histoire de MassChroQ : le bond

2011

- QuantiMsCpp devient MassChroQ (Mass Chromatogram Quantification software).
- MassChroQ : A versatile tool for mass spectrometry quantification, B. Valot, O. Langella, E. Nano, M. Zivy. Proteomics, 2011, 11 :3572–3577.
- La première version publique de MassChroQ est diffusée en tant que logiciel libre à l'occasion de la sortie de la publication en Juin.



# Histoire de MassChroQ : le bond

2011

- QuantiMsCpp devient MassChroQ (Mass Chromatogram Quantification software).
- MassChroQ : A versatile tool for mass spectrometry quantification, B. Valot, O. Langella, E. Nano, M. Zivy. Proteomics, 2011, 11 :3572–3577.
- La première version publique de MassChroQ est diffusée en tant que logiciel libre à l'occasion de la sortie de la publication en Juin.
- Une fondation solide (QuantiMsCpp), puis les projets HeterosYeast et Dromadair d'une part et la préparation de la publication d'autre part, ont fait de MassChroQ un logiciel de quantification versatile et performant.

2. Mais qu'est que c'est la quantification par LC-MS/MS ?

# LC-MS/MS

- Besoin d'identifier et de quantifier les peptides et les protéines contenus dans les échantillons.

# LC-MS/MS

- Besoin d'identifier et de quantifier les peptides et les protéines contenus dans les échantillons.
- *Chromatographie Liquide (LC)* : technique chimique d'analyse permettant de séparer les molécules d'un composé au cours du temps  $\Rightarrow$  **chromatogrammes** (courbes d'intensité en fonction du temps).

# LC-MS/MS

- Besoin d'identifier et de quantifier les peptides et les protéines contenus dans les échantillons.
- *Chromatographie Liquide (LC)* : technique chimique d'analyse permettant de séparer les molécules d'un composé au cours du temps  $\Rightarrow$  **chromatogrammes** (courbes d'intensité en fonction du temps).
- *Spectrométrie de Masse (MS)* : technique physique d'analyse permettant de mesurer le rapport masse sur charge des molécules d'un composé  $\Rightarrow$  **spectres de masse** (intensité en fonction de masse/charge).

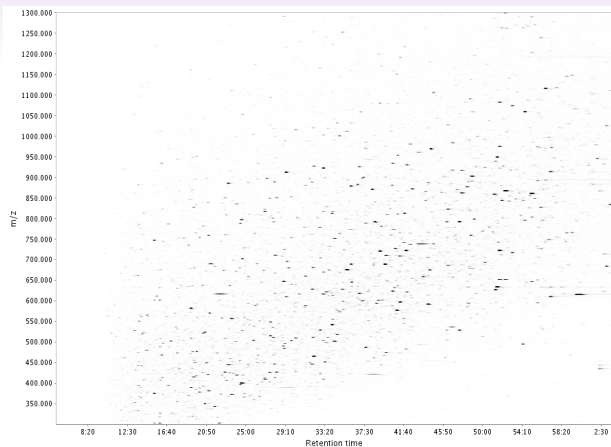
# LC-MS/MS

- Besoin d'identifier et de quantifier les peptides et les protéines contenus dans les échantillons.
- *Chromatographie Liquide (LC)* : technique chimique d'analyse permettant de séparer les molécules d'un composé au cours du temps  $\Rightarrow$  **chromatogrammes** (courbes d'intensité en fonction du temps).
- *Spectrométrie de Masse (MS)* : technique physique d'analyse permettant de mesurer le rapport masse sur charge des molécules d'un composé  $\Rightarrow$  **spectres de masse** (intensité en fonction de masse/charge).
- *MS/MS* : au cours de la MS, on sélectionne un peptide, on le fragmente et on analyse ces fragments lors d'une deuxième MS  $\Rightarrow$  permet l'identification du peptide.

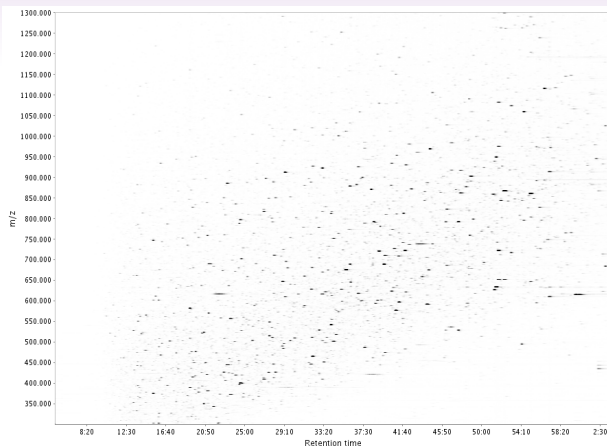
# LC-MS/MS

- Besoin d'identifier et de quantifier les peptides et les protéines contenus dans les échantillons.
- *Chromatographie Liquide (LC)* : technique chimique d'analyse permettant de séparer les molécules d'un composé au cours du temps  $\Rightarrow$  **chromatogrammes** (courbes d'intensité en fonction du temps).
- *Spectrométrie de Masse (MS)* : technique physique d'analyse permettant de mesurer le rapport masse sur charge des molécules d'un composé  $\Rightarrow$  **spectres de masse** (intensité en fonction de masse/charge).
- *MS/MS* : au cours de la MS, on sélectionne un peptide, on le fragmente et on analyse ces fragments lors d'une deuxième MS  $\Rightarrow$  permet l'identification du peptide.
- *LC-MS/MS* : En protéomique, la MS couplée à une phase préalable de LC, permet de séparer, d'identifier et de quantifier les peptides et les protéines d'un composé.

# LC-MS/MS

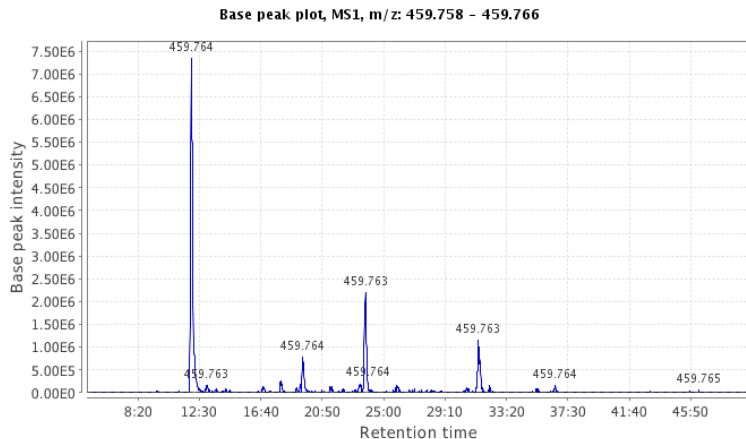






## Extraction des chromatogrammes (XICs)

Pour chaque peptide ( $m/z$ ) à quantifier, on va extraire à partir des données MS, son intensité au cours du temps chromatographique.

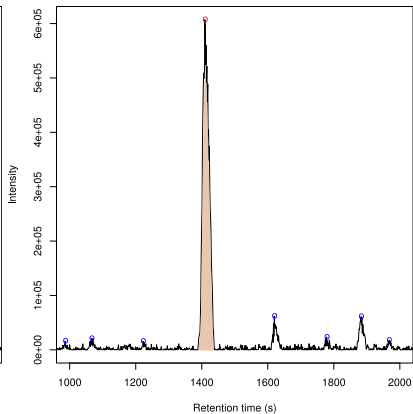
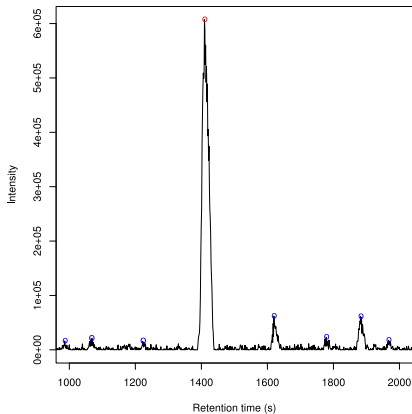


## Extraction des chromatogrammes (XICs)

Pour chaque peptide (m/z) à quantifier, on va extraire à partir des données MS, son intensité au cours du temps chromatographique.

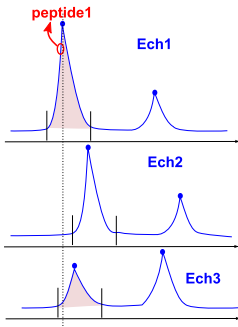
## Détection de pics et quantification

Après traitement du signal de ces XICs, nous allons y détecter les pics, puis mesurer leur aire (valeur quantitative).



## Alignement des LC

La LC génère inévitablement des décalages de temps de rétention : besoin d'alignement des échantillons LC-MS semblables afin d'éviter les biais expérimentaux.



- 
- 
3. Et sinon MassChroQ, il fait quoi exactement ?

# Illustration avec les données de la publication

## Protéome complexe

Digestat d'un extrait protéique de levure :

- 6 répétitions techniques
- 700 ng injectés

# Illustration avec les données de la publication

## Protéome complexe

Digestat d'un extrait protéique de levure :

- 6 répétitions techniques
- 700 ng injectés

## Protéine standard

Injection de gamme croissante de BSA : 4,5, 15, 45, 105, 450 et 1500 fmol.

# Illustration avec les données de la publication

## Protéome complexe

Digestat d'un extrait protéique de levure :

- 6 répétitions techniques
- 700 ng injectés

## Protéine standard

Injection de gamme croissante de BSA : 4,5, 15, 45, 105, 450 et 1500 fmol.

## Analyse en LC-MS/MS

HR High Resolution (Orbitrap - Jouy)

LR Low Resolution (LTQ-XL - Moulon)

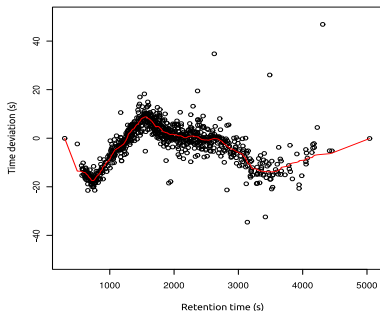


# Alignement dans MassChroQ

## Deux méthodes d'alignement implémentées

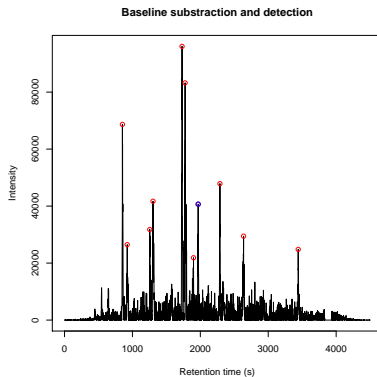
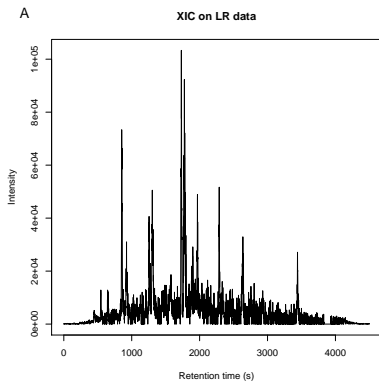
- Une méthode basée sur les données MS : intégration du logiciel libre *ObiWarp*
- Une méthode utilisant les résultats de l'identification (MS/MS) : algorithme développé à Pappso

Exemple d'alignement  
MS/MS



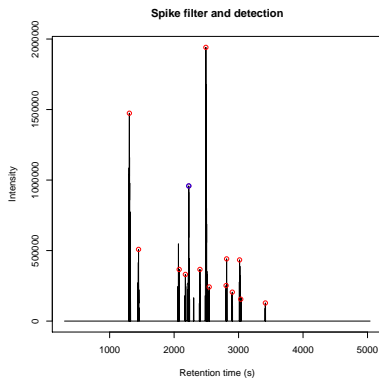
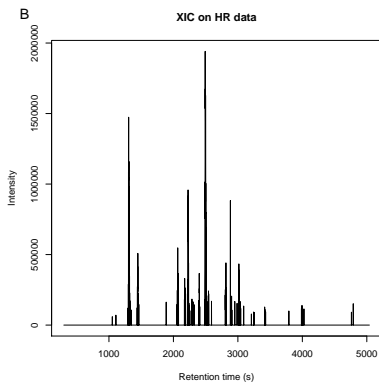
# Extraction des valeurs quantitatives

MassChroQ extrait les XICs, filtre les différents bruits du signal, détecte les pics et calcule leurs aires.

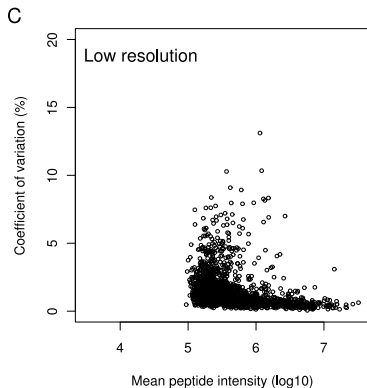
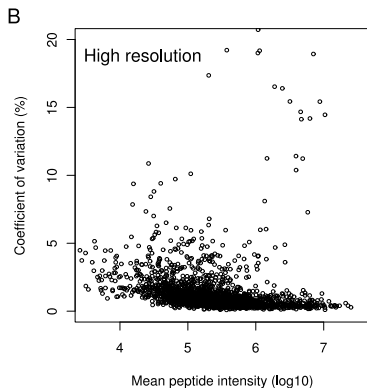


# Extraction des valeurs quantitatives

MassChroQ extrait les XICs, filtre les différents bruits du signal, détecte les pics et calcule leurs aires.

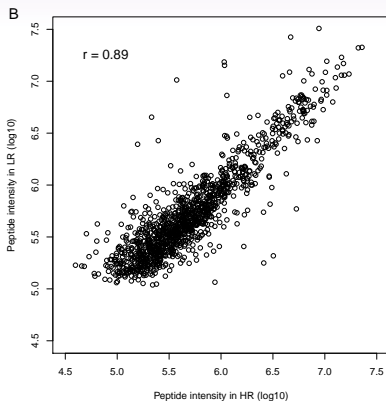


# Reproductibilité de la quantification



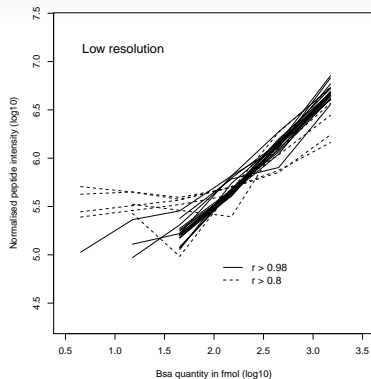
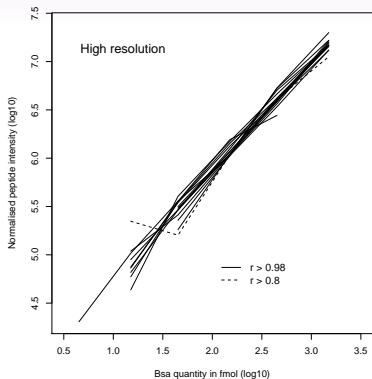
Les données montrent des CV inférieurs à 5%.

# Comparaison entre LR et HR



Les données quantitatives obtenues à la fois en LR et HR sont comparables (1179 peptides).

# Estimation d'une variation quantitative



La corrélation entre l'intensité des peptides et la quantité de BSA injectée, est meilleure en HR qu'en LR.

## En résumé, MassChroQ peut :

- quantifier des données HR, aussi bien que LR (publication) ;
- quantifier des données avec marquage isotopique (projet Dromadair), aussi bien que sans marquage (projet HeterosYeast) ;
- s'adapter à divers traitements (par exemple aux analyses avec préfractionnement de Dromadair) ;
- traiter une grande quantité de données en une seule fois, rapidement, sur un PC de bureau qui reste opérationnel (voir le poster).

## En résumé, MassChroQ peut :

- quantifier des données HR, aussi bien que LR (publication) ;
- quantifier des données avec marquage isotopique (projet Dromadair), aussi bien que sans marquage (projet HeterosYeast) ;
- s'adapter à divers traitements (par exemple aux analyses avec préfractionnement de Dromadair) ;
- traiter une grande quantité de données en une seule fois, rapidement, sur un PC de bureau qui reste opérationnel (voir le poster).

## De plus :

- MassChroQ est un logiciel libre : son code source est ouvert, il est librement téléchargeable et son utilisation n'est pas limitée (et ça, mine de rien, c'est beau).
- Toutes les données qu'il utilise ou qu'il produit sont sous des formats standards et ouverts (et ça, mine de rien, c'est important).



## Pour conclure

- MassChroQ est aujourd'hui un projet libre, ouvert au public.  
Site internet (code source, documentation, publication, ...) :

<http://pappso.inra.fr/bioinfo/masschroq/>

- La troisième version de release est actuellement en production.  
La quatrième version avec une interface graphique est en cours de développement.
- Il est utilisé pour quasiment toutes les analyses de quantification LC-MS de la plateforme (et il y en a beaucoup).
- Des laboratoires externes à l'UMR l'ont déjà adopté et d'autres rêvent d'y parvenir un jour ...