

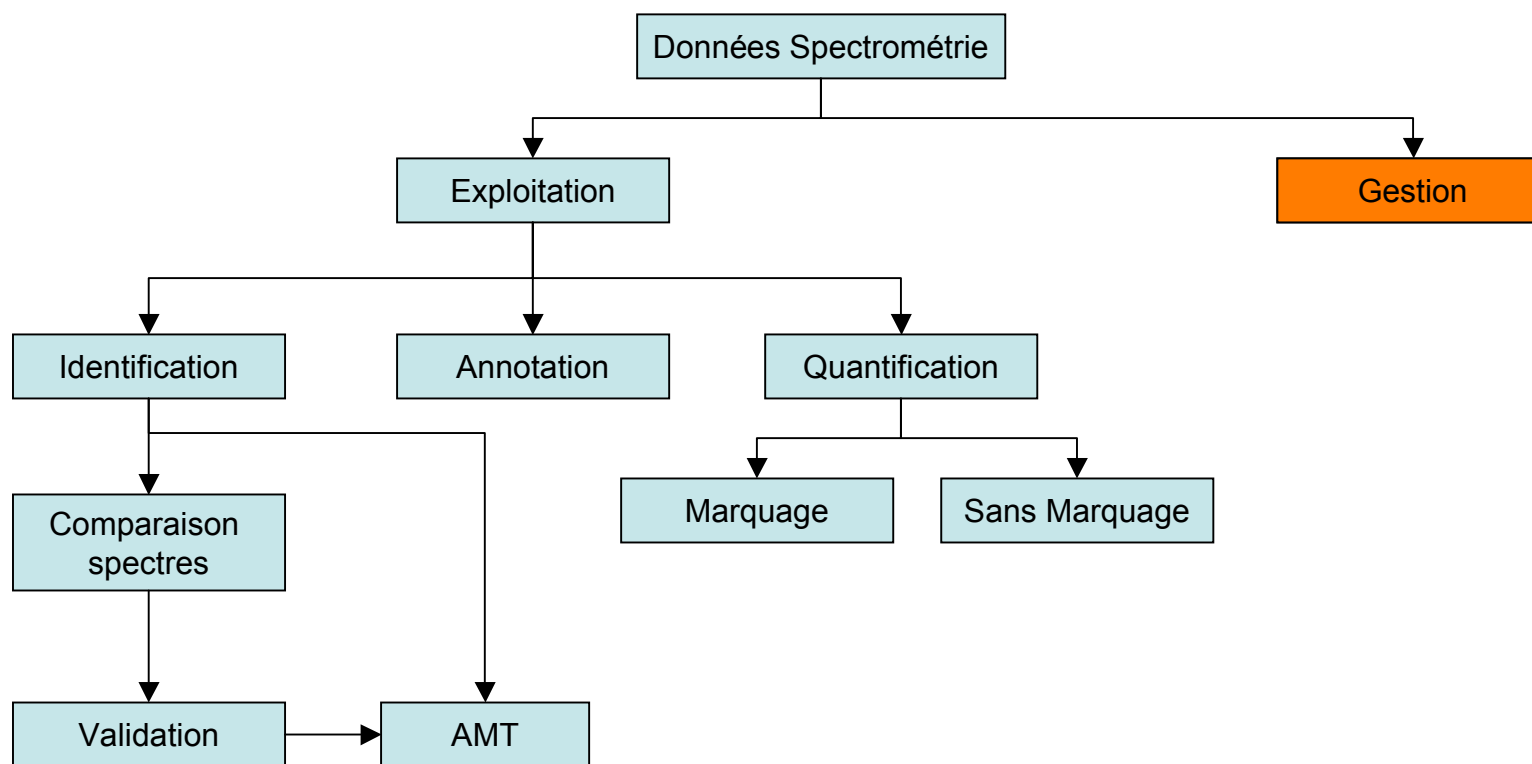


Quelle (bio)informatique pour la protéomique ?

Christophe Bruley - CEA Grenoble

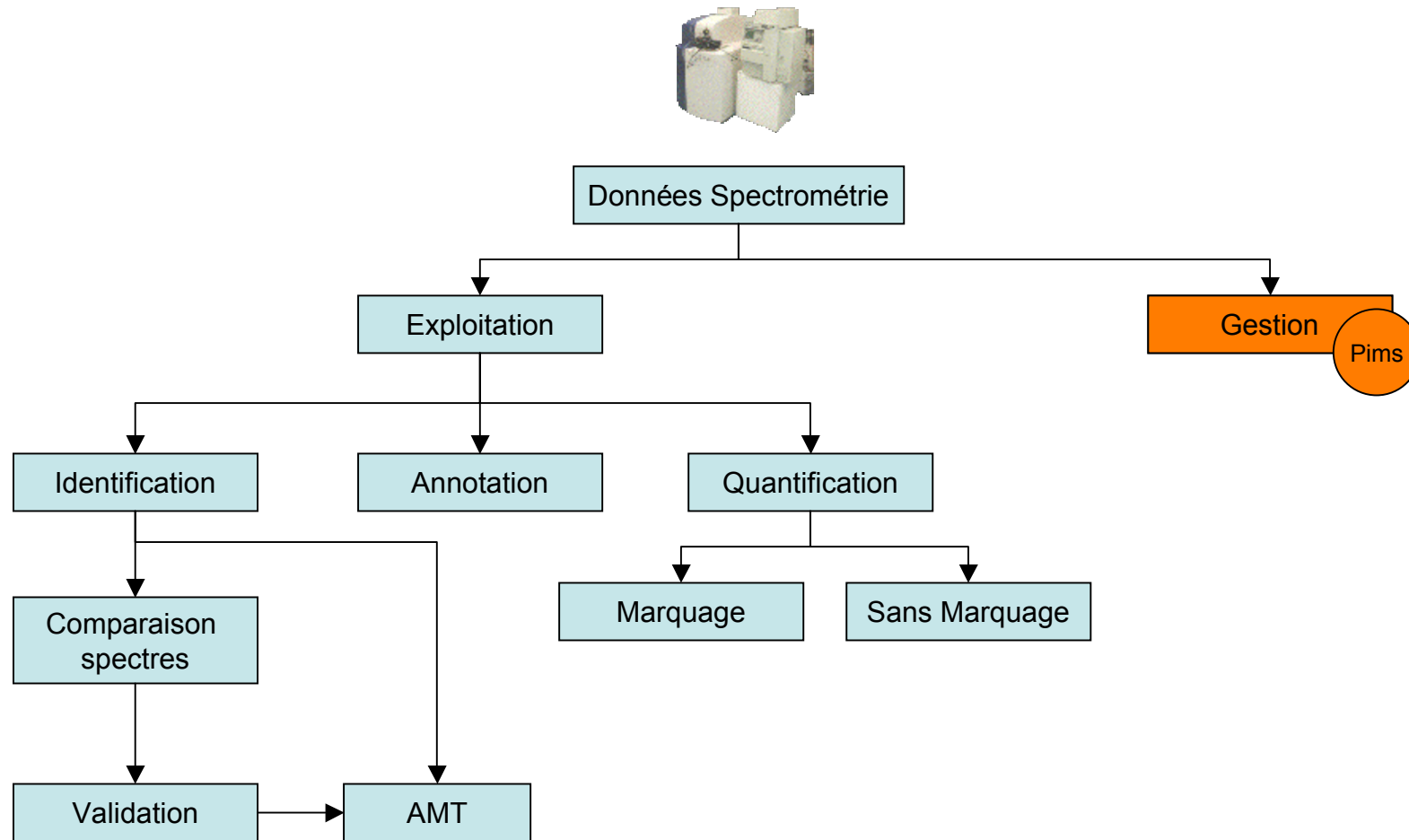
- **Positionnement et activité de l'équipe (bio)informatique**
 - Pas d'activité qui ne soit motivée par un besoin en spectrométrie de masse
 - Développement d'outils destinés à un réel usage et pas seulement des « preuves de concepts »
 - Organisation en groupes de travail thématiques
 - Suivi du développement (dicté par la demande)
 - Effort particulier d'utilisabilité
 - Mise en place et administration de l'infrastructure logicielle et matérielle

Informatique pour la spectrométrie de masse



- **Besoin**
 - Volumétrie analyses MS et MS/MS induite par :
 - Le nombre de personnes soumettant des échantillons
 - Le nombre et la nature des instruments utilisés sur la plate-forme
 - La complexité des échantillons (durée des LC)
 - Démarche qualité
 - Traçabilité des expérimentations
 - Suivi de l'activité

Informatique pour la spectrométrie de masse



C. Pineau, R. Lavigne, D. Ousmanou – Infrastructure informatique et bioinformatique pour l'identification haut-débit.

Gestion des données de spectrométrie

• PIMS

– Fonctionnalités

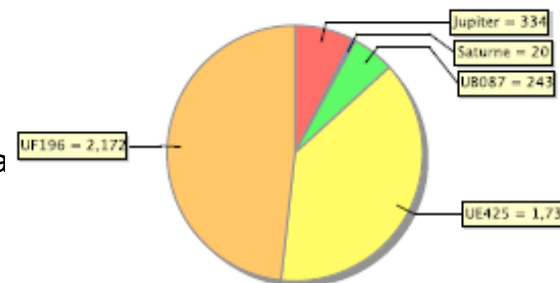
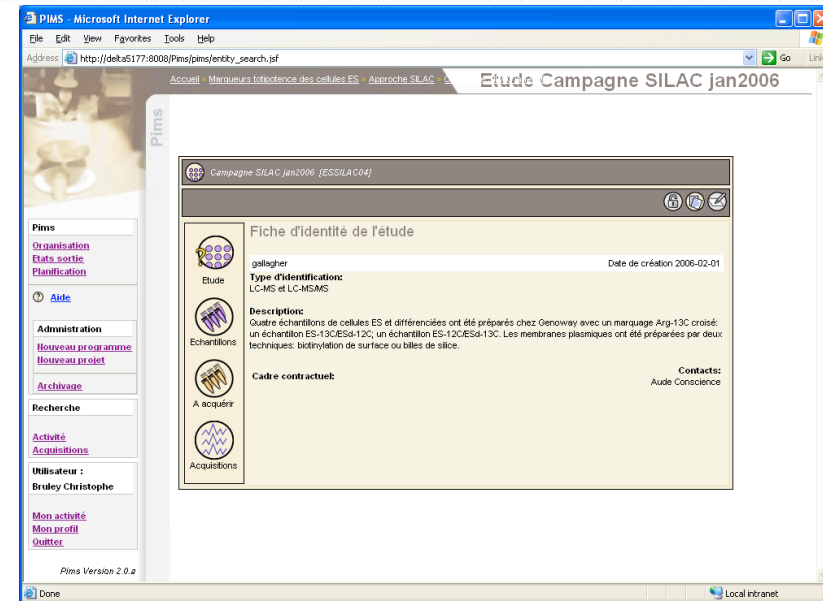
- Organisation de l'activité du laboratoire
- Recherche des études ou des acquisitions
- Accès aux données électroniques associées à une étude
- Gestion des échantillons
- Gestion des archivages sur bandes
- Génération d'états de sortie

• Statut actuel

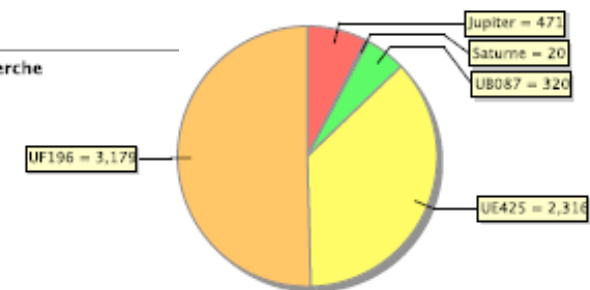
– En interne

- Mise en service de l'infrastructure logicielle et matérielle en novembre 2004
- Utilisé pour tous les instruments de la plate-forme à l'exception du MALDI
- Gère à ce jour
 - 267 études
 - 4500 échantillons analysés

- Création d'un groupe de travail « LIMS »

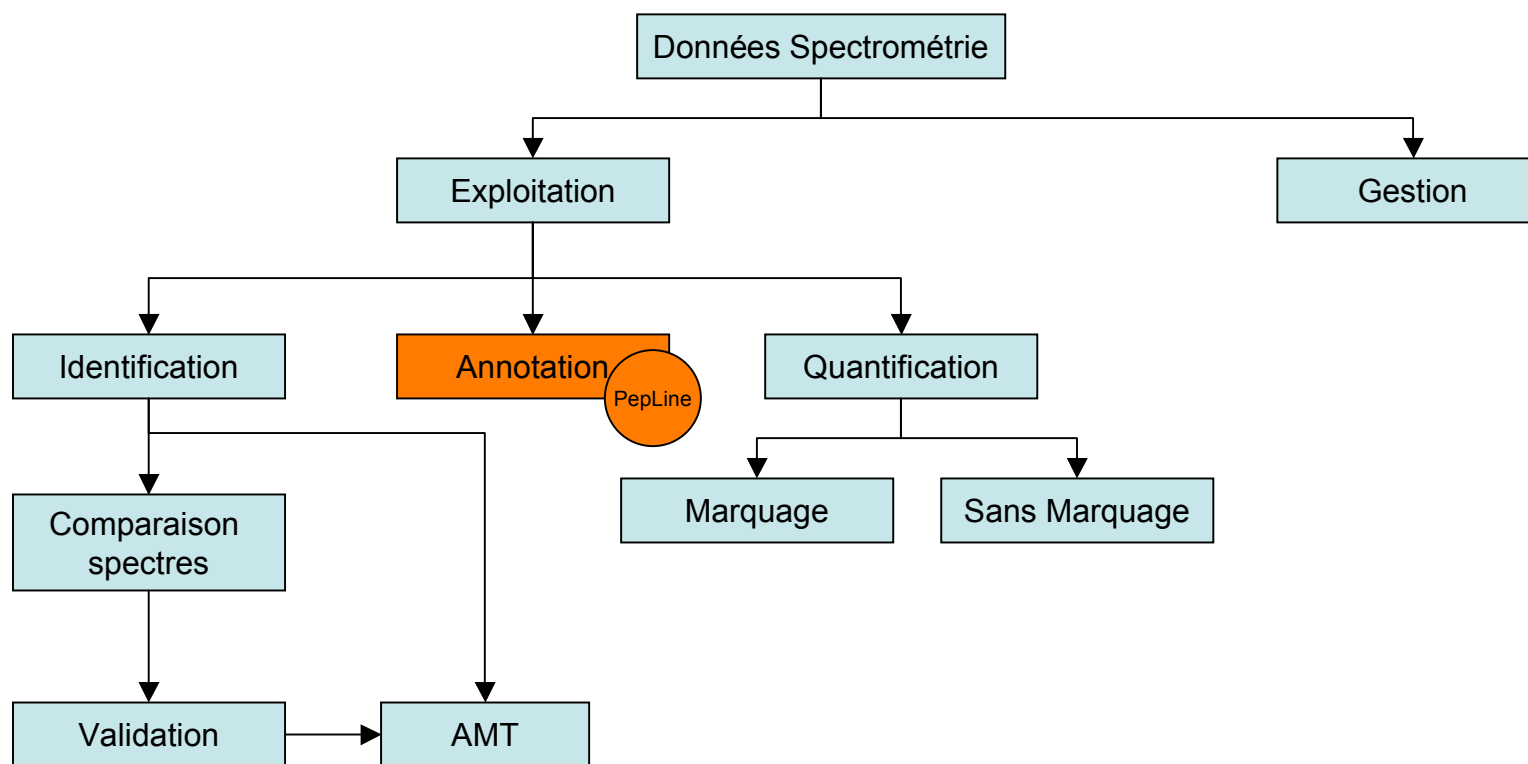


Number of Acquisitions - Recherche



Laboratory Activity (Nbr Acq.)

Informatique pour la spectrométrie de masse



Protéomique et annotation des génomes

- **Contexte**

- Annotation du génome de *Chlamydomonas reinhardtii*

- **Principe**

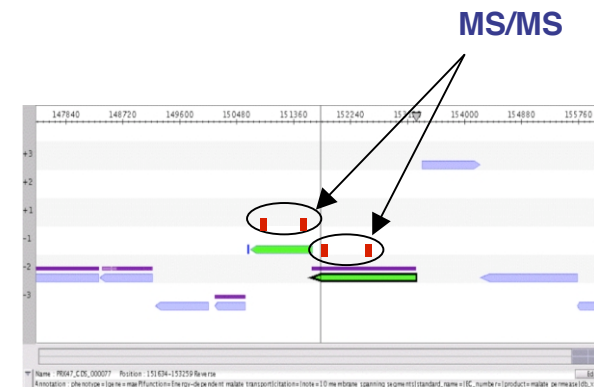
- Interprétation partielle des spectres MSMS en PST
- Localisation des PST sur le génome
- Regroupement des PST pour localiser les gènes

- **Outil**

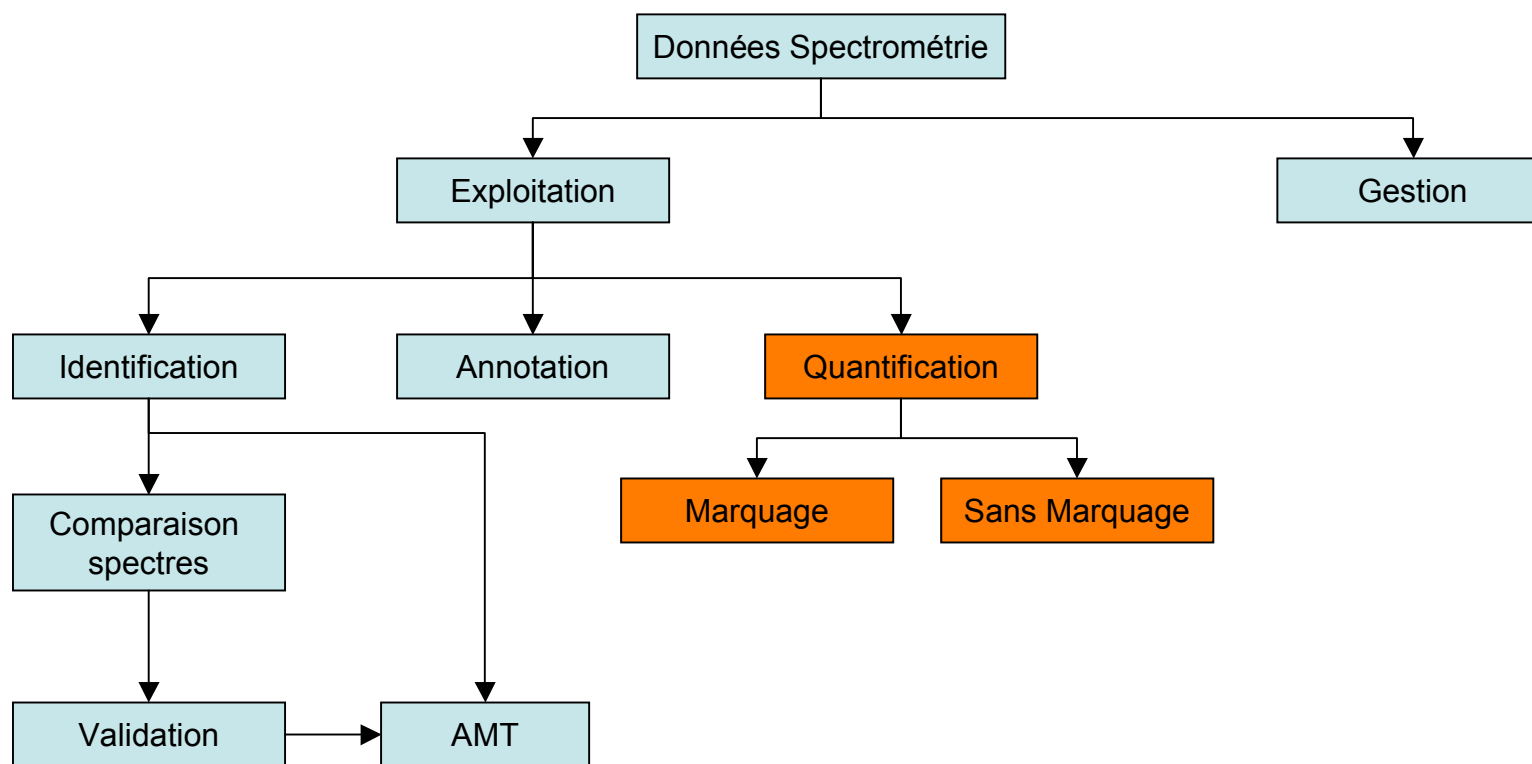
- PepLine
 - Collaboration CEA – INRIA Rhône-Alpes – Genome Express

- **Futur**

- GénoProtéo
 - Intégration de PepLine dans la plate-forme Genostar (iogma)
 - *Confrontation des annotations issues de données expérimentales avec les prédictions obtenues in-silico*
 - Encadrement d'un ingénieur informaticien (12 mois de CDD financé par ACI IMPBio)
 - *Améliorations PepLine*
 - *Développement du module de Genostar*



Informatique pour la spectrométrie de masse

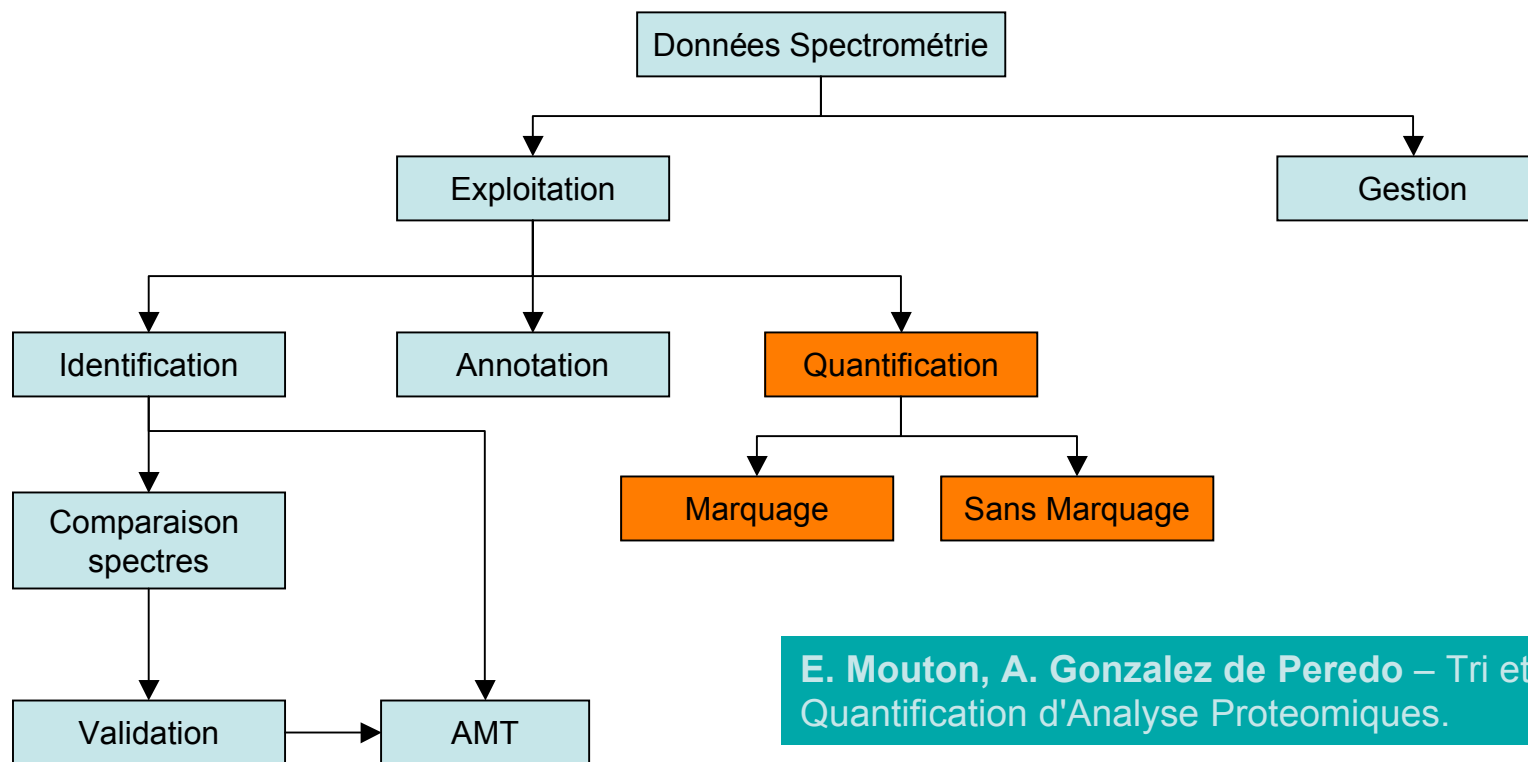


- **Analyses différentielles ou quantitatives**
 - Analyse différentielle
 - Comparer le contenu (peptides) de 2 analyses MS
 - *Comparaison de cartes peptidiques*
 - *Comparaison de liste de peptides*
 - Marquage isotopique

Principe : une unique analyse en masse d'un mélange de deux échantillons différemment marqués

 - Co-élution des peptides marqués et non-marqués en nanoLC
 - *Identification puis quantification*
 - » Nombre de peptides fragmentés
 - » Signal MS
 - *Quantification puis identification*
 - » Signal MS
 - Co-fragmentation des peptides marqués et non-marqués en MS/MS
 - » Pics MS/MS
 - Peptides marqueurs

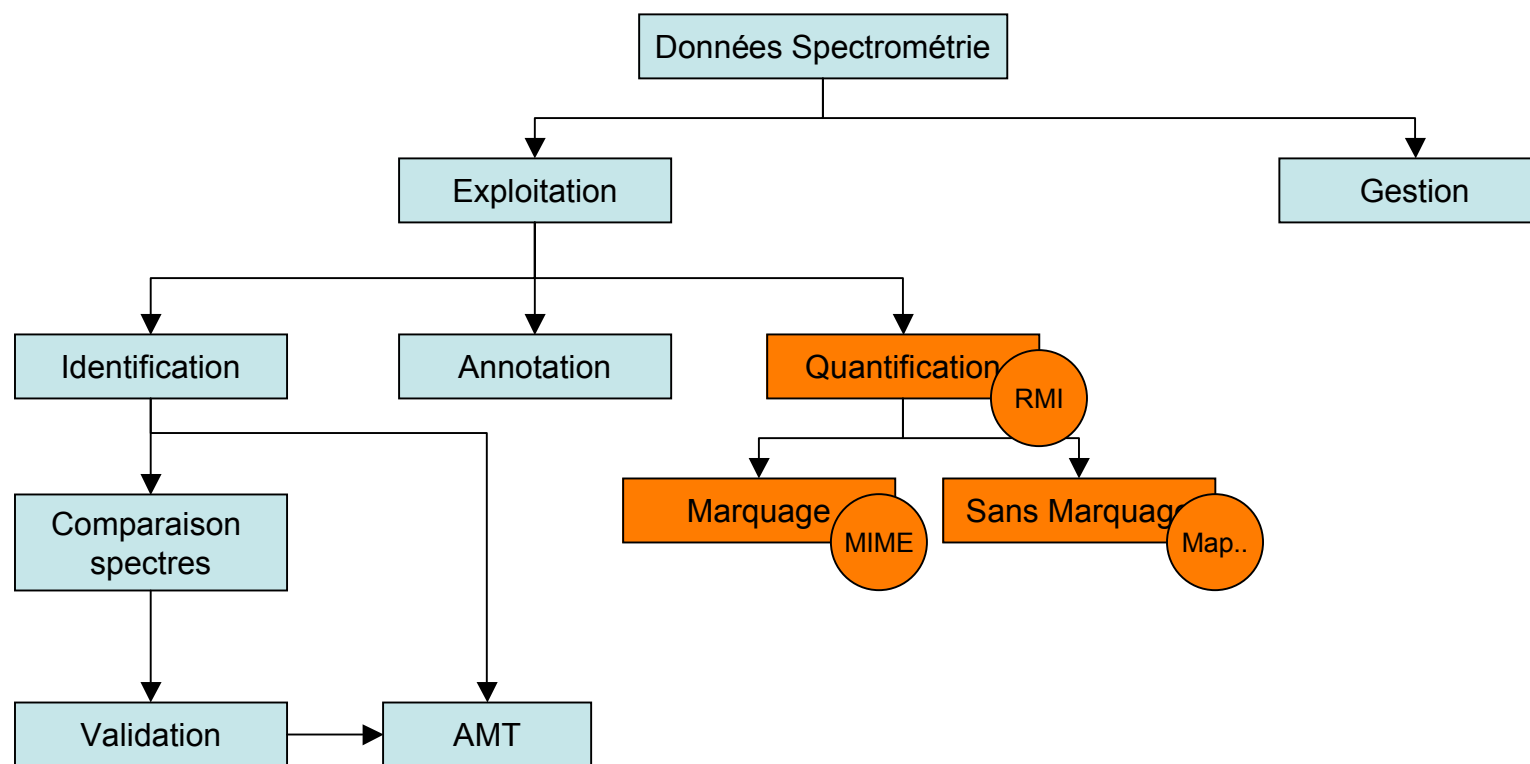
Informatique pour la spectrométrie de masse



E. Mouton, A. Gonzalez de Peredo – Tri et Quantification d'Analyse Proteomiques.

B. Schwikowski – Comparative LC Proteomics.

Informatique pour la spectrométrie de masse

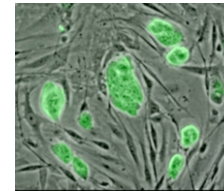


Quantification par marquage isotopique

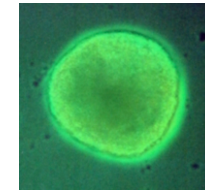
- **Contexte**

- Application à l'étude de la différenciation des cellules souches totipotentes embryonnaires murines.

- Analyse par marquage SILAC (sept 2004)



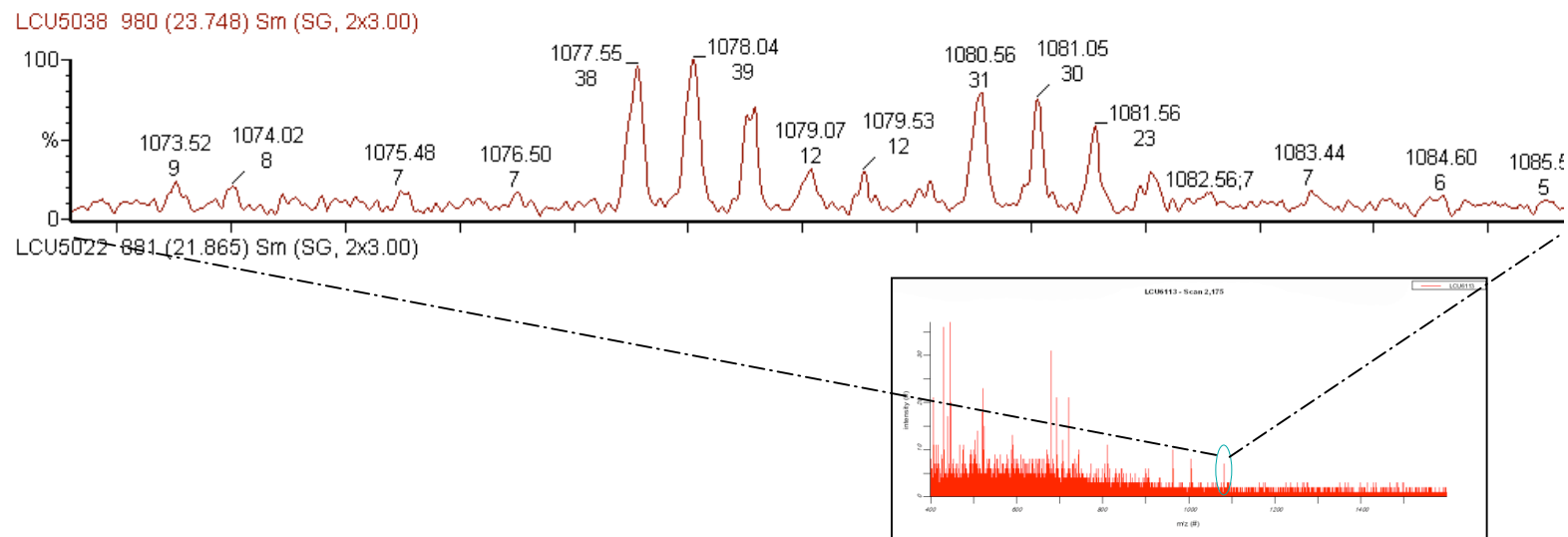
Cellules ES totipotentes



Corps embryoïde

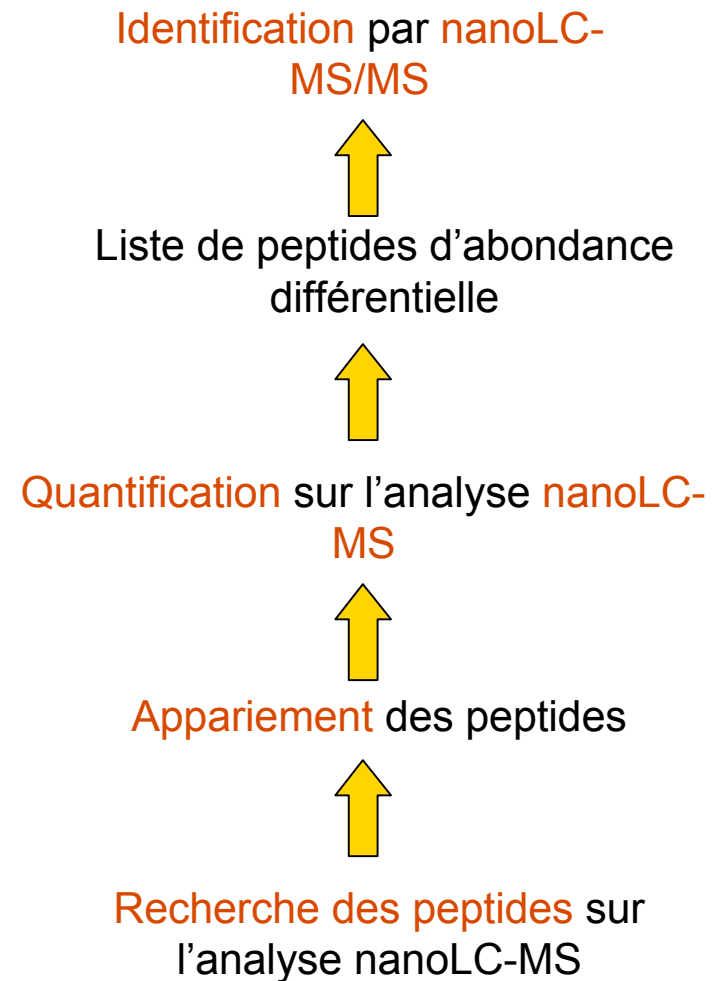
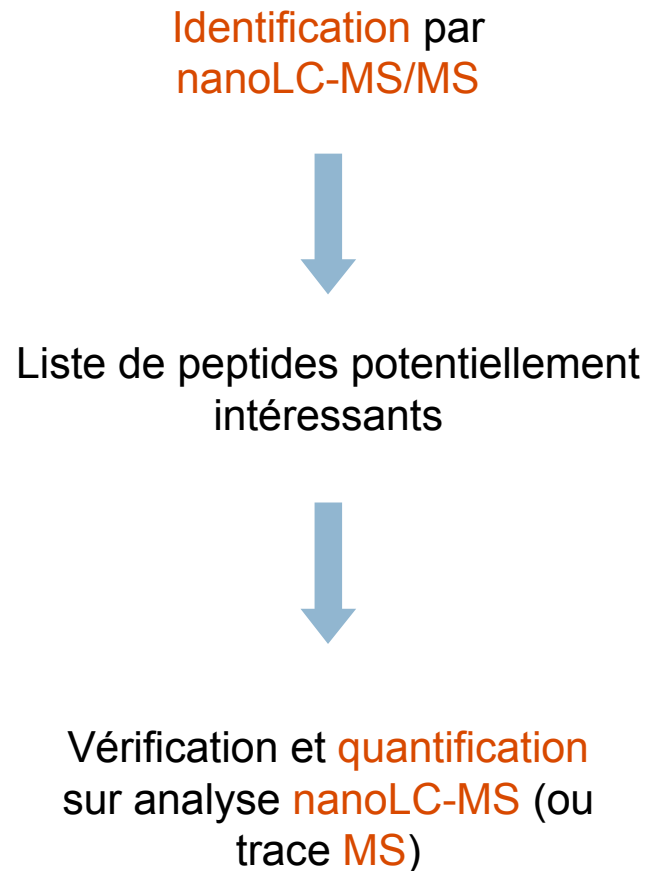
- **Principe**

- Co-élution des peptides marqués et non-marqués
- Différence de masse constante

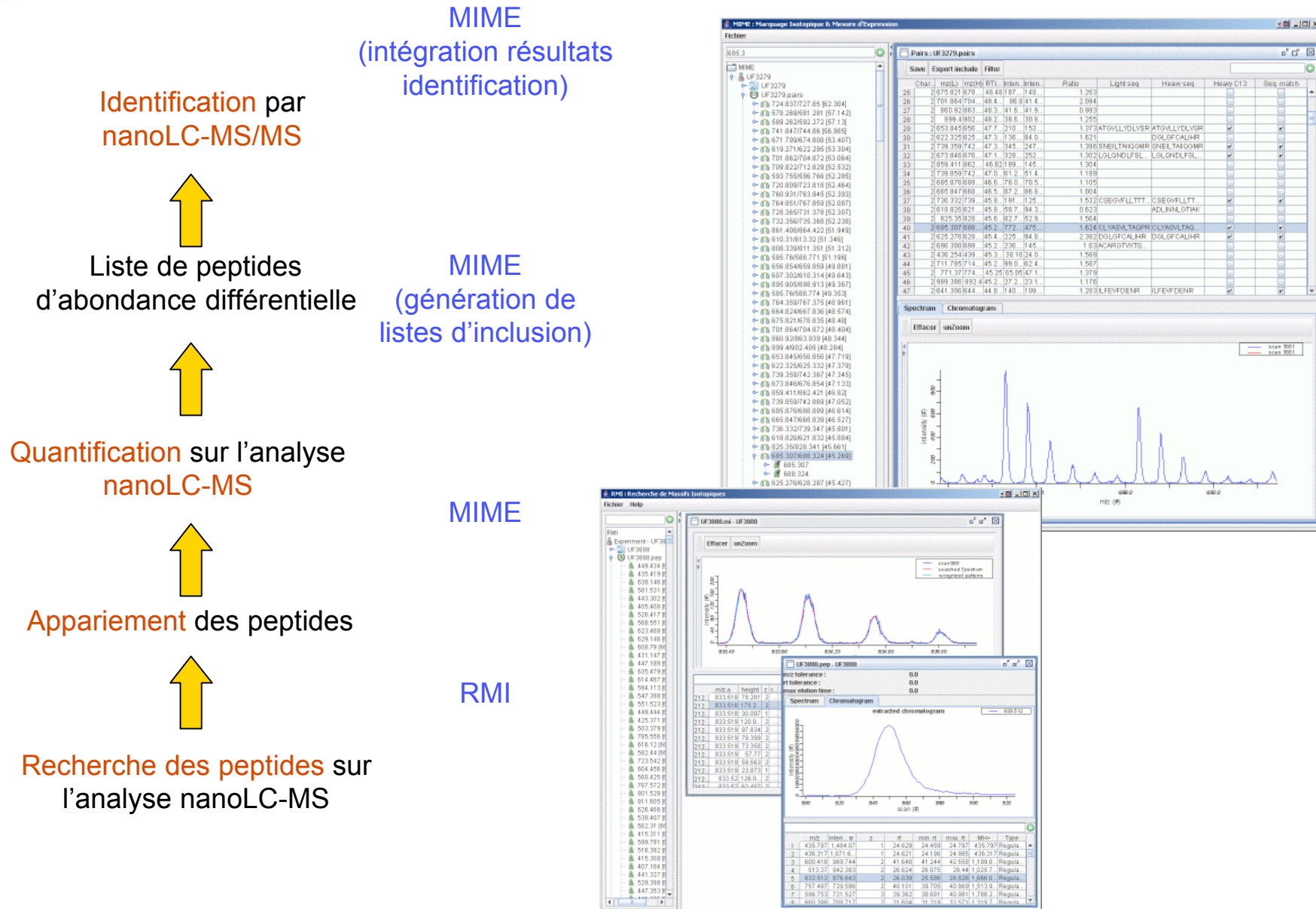


Quantification par marquage isotopique

- **Stratégies d'analyses**



Quantification par marquage isotopique



Quantification par marquage isotopique

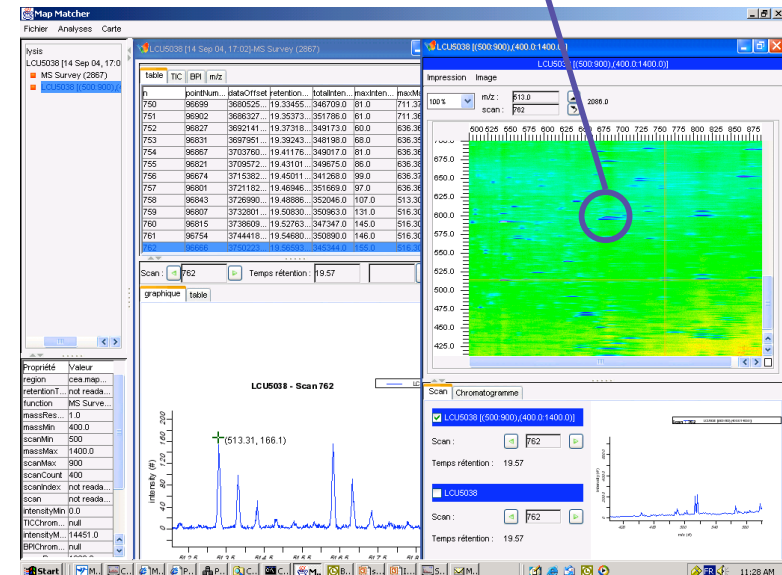
- **Statut actuel**

- Détection des peptides (**RMI**)
 - Utilisé pour l'analyse des échantillons SILAC
 - Tests préliminaires pour les analyses avec marquage N15
- Détection des paires (**MIME**)
 - Utilisé pour l'analyse des échantillons SILAC

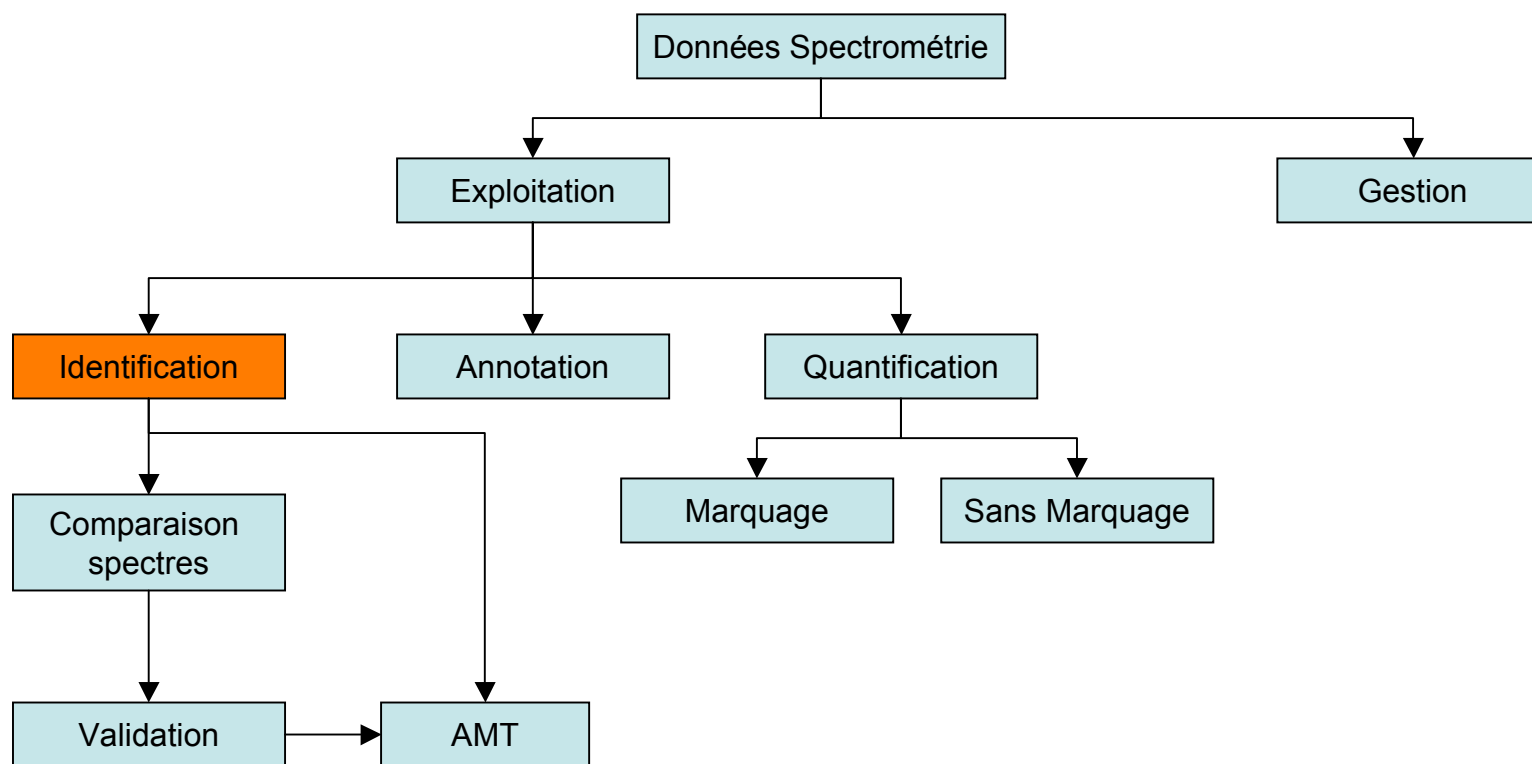
- **Futur**

- Comparaison d'analyses MS sans marquage
 - Chaque point de la carte peptidique est un peptide dont on connaît désormais les caractéristiques (**MapMatcher**)
 - Méthodes d'alignement de listes de peptides
- Détection de peptides (**RMI**)
 - Extension du modèle pour prendre en compte le temps de rétention (détection en 2D)

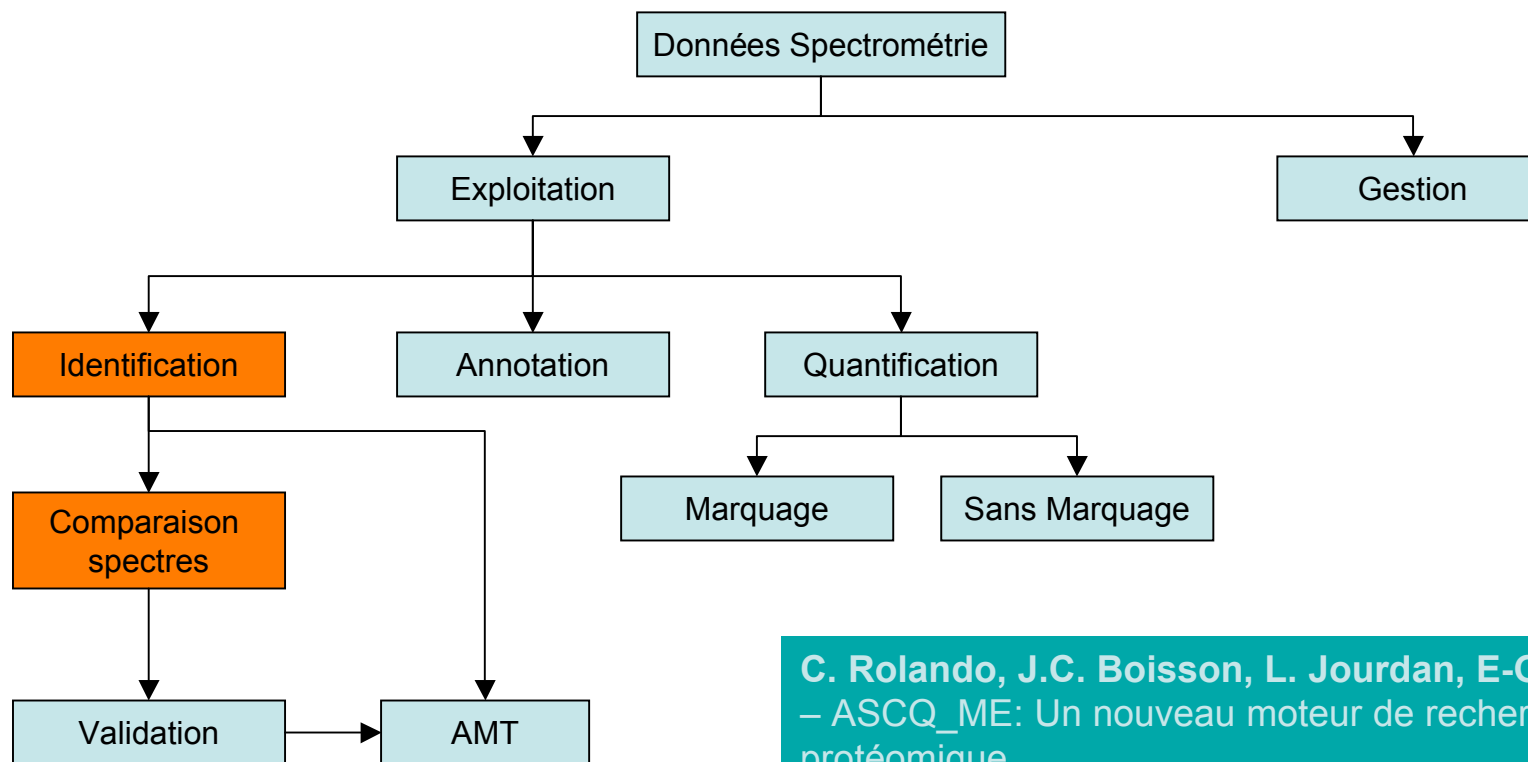
Masse monoisotopique
Charge
Temps de rétention



Informatique pour la spectrométrie de masse

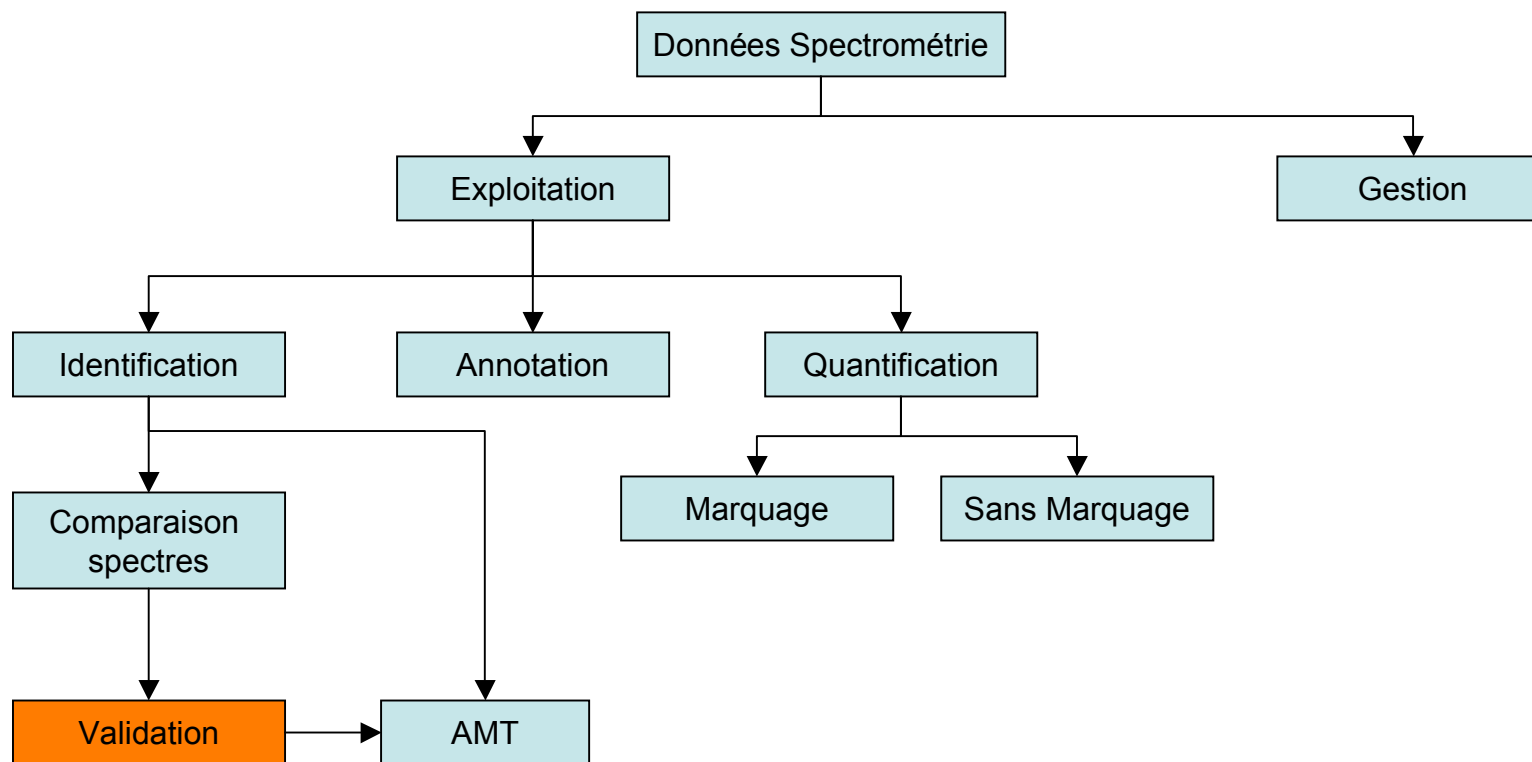


Informatique pour la spectrométrie de masse

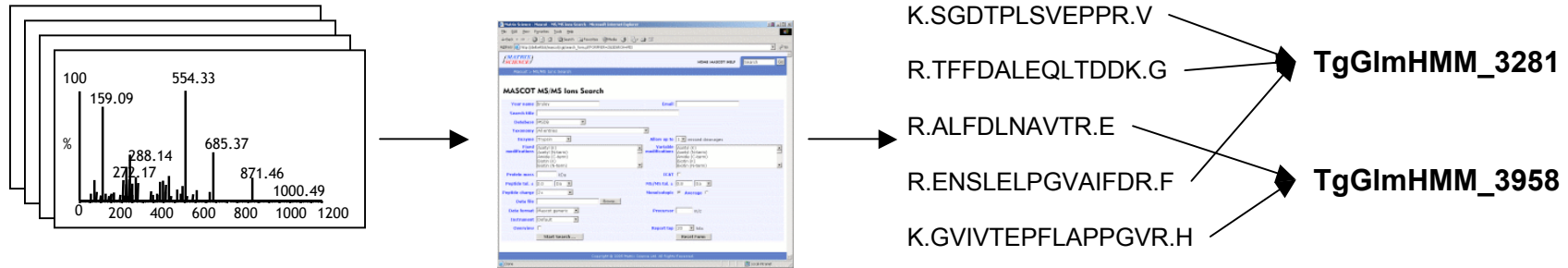


C. Rolando, J.C. Boisson, L. Jourdan, E-G. Talbi
– ASCQ_ME: Un nouveau moteur de recherche en protéomique.

Informatique pour la spectrométrie de masse



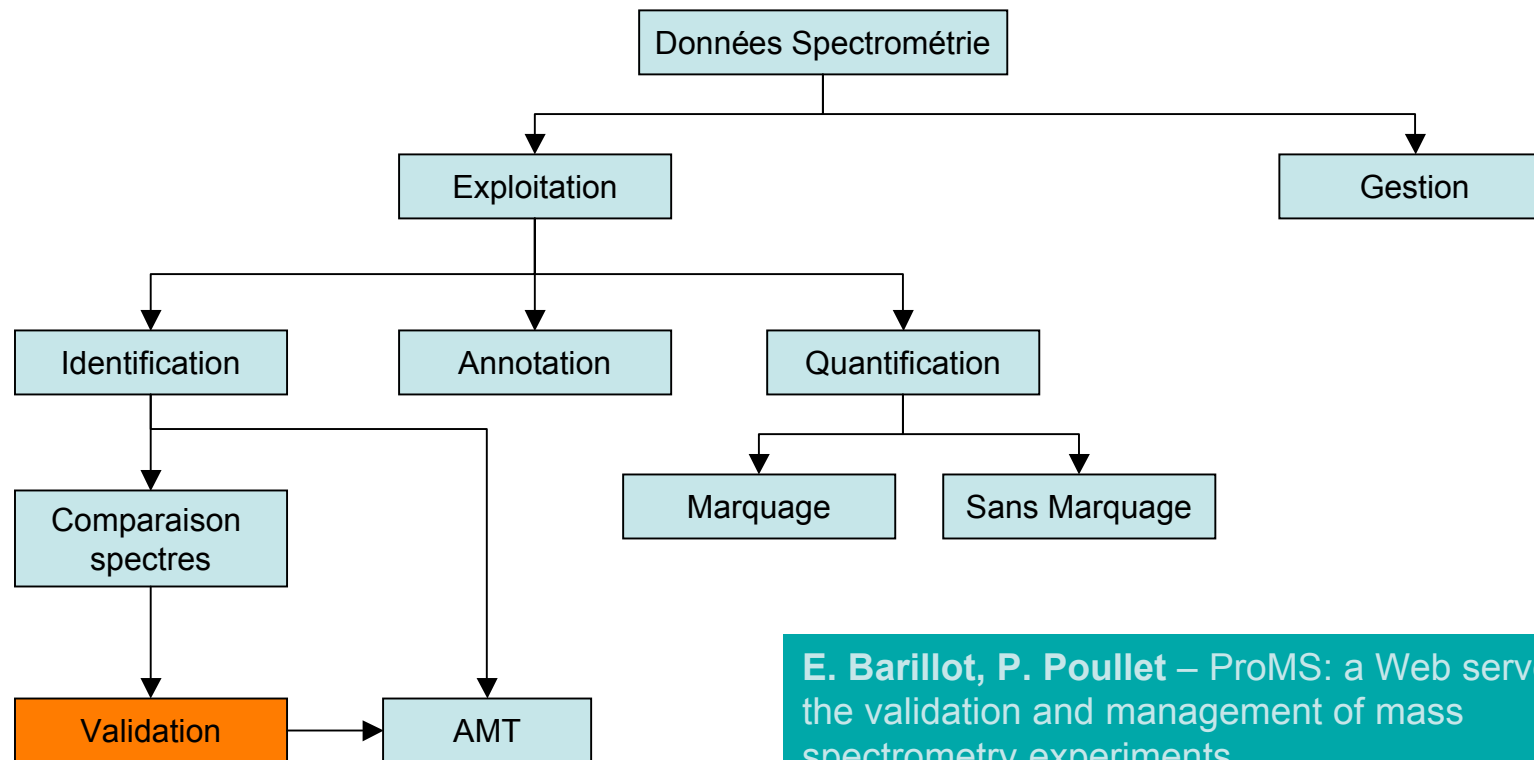
Validation des identifications



- **Besoins**

- Filtrer les résultats proposés par Mascot®/Sequest®/ ...
 - Selon des critères pré-définis
 - manuellement
- Juger de l'adéquation spectre - séquence peptidique – protéine
 - Accès au spectre matché sur la séquence peptidique
 - Accès à la description de la protéine (NCBI, SwissProt, etc)
- Re-soumettre une recherche pour affiner un résultat
- Rendre des résultats d'identification complets, cohérents et validés
 - Maintien de la cohérence des regroupements peptides – protéine
 - Re-calcul des pourcentages de couverture peptides – protéine

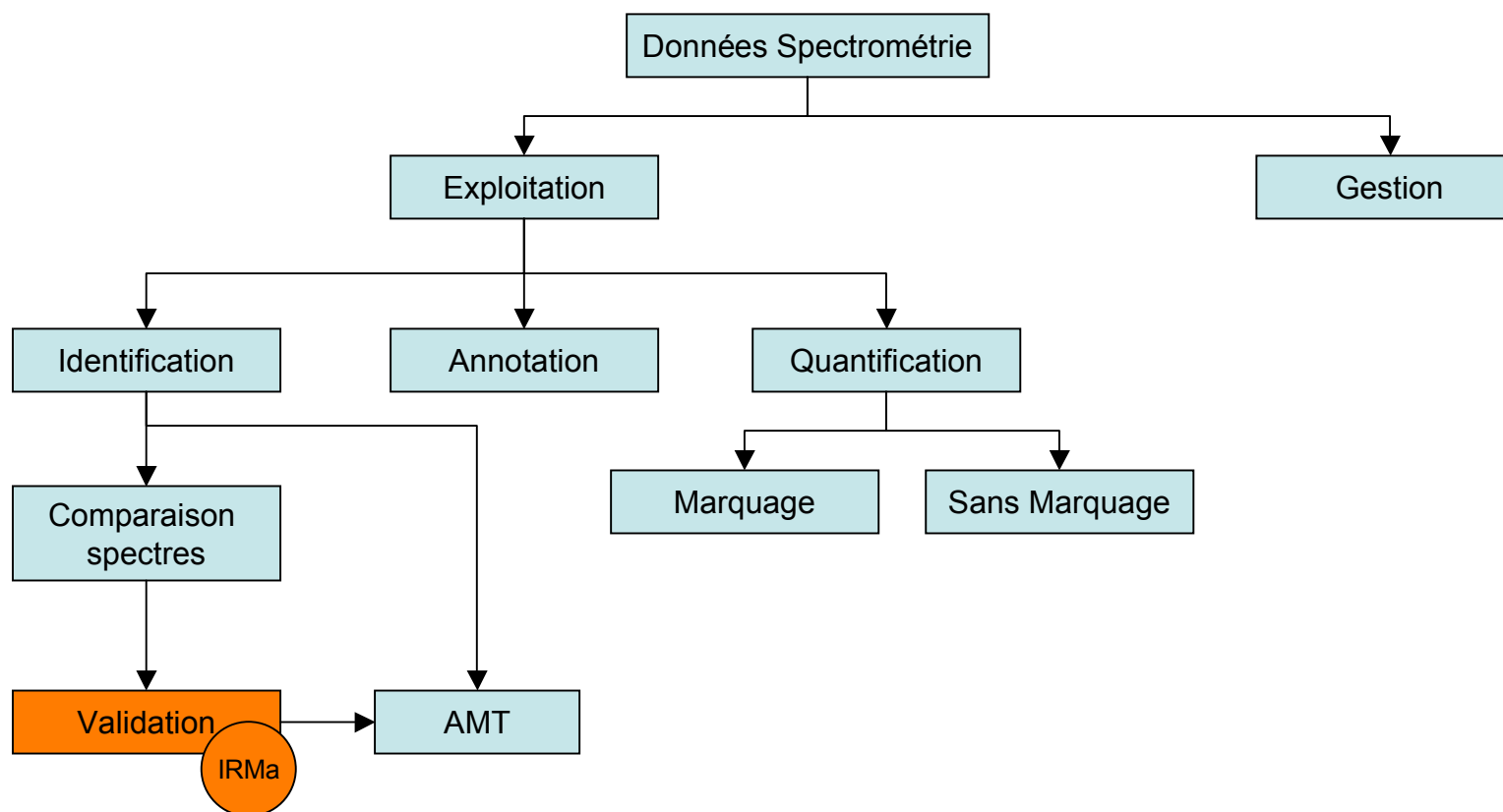
Informatique pour la spectrométrie de masse



E. Barillot, P. Poulet – ProMS: a Web server for the validation and management of mass spectrometry experiments.

E. Mouton, A. Gonzalez de Peredo – Tri et Quantification d'Analyse Proteomiques.

Informatique pour la spectrométrie de masse

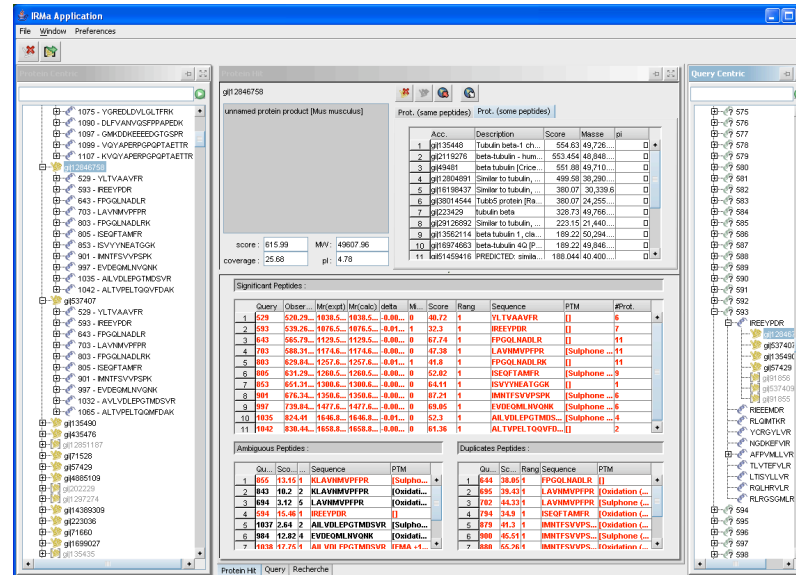


Validation des identifications

- IRMa

- Fonctionnalités

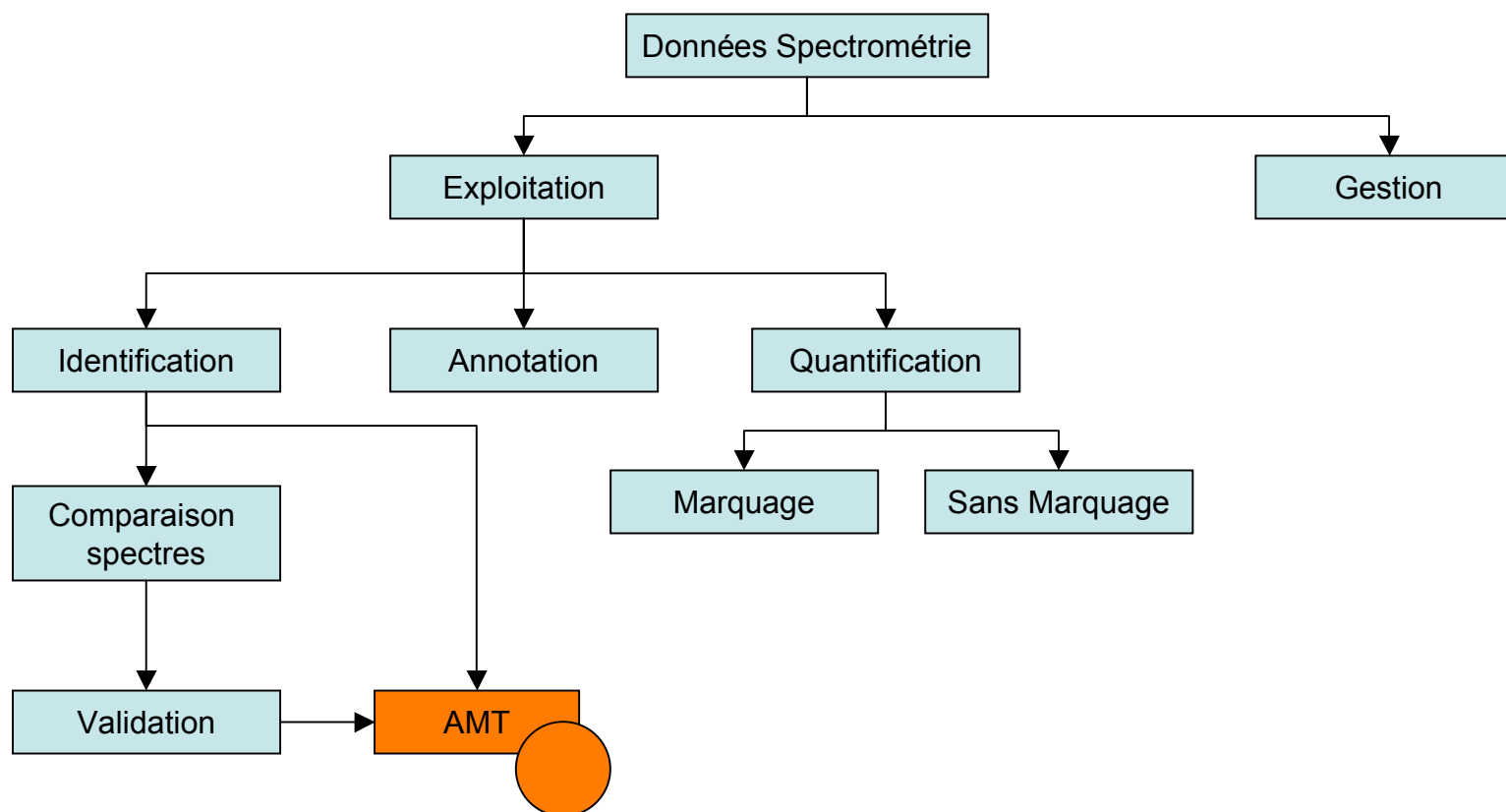
- Relecture directe des fichiers de résultats Mascot
 - Filtrage automatique / manuel
 - Possibilités d'exports



- Statut actuel

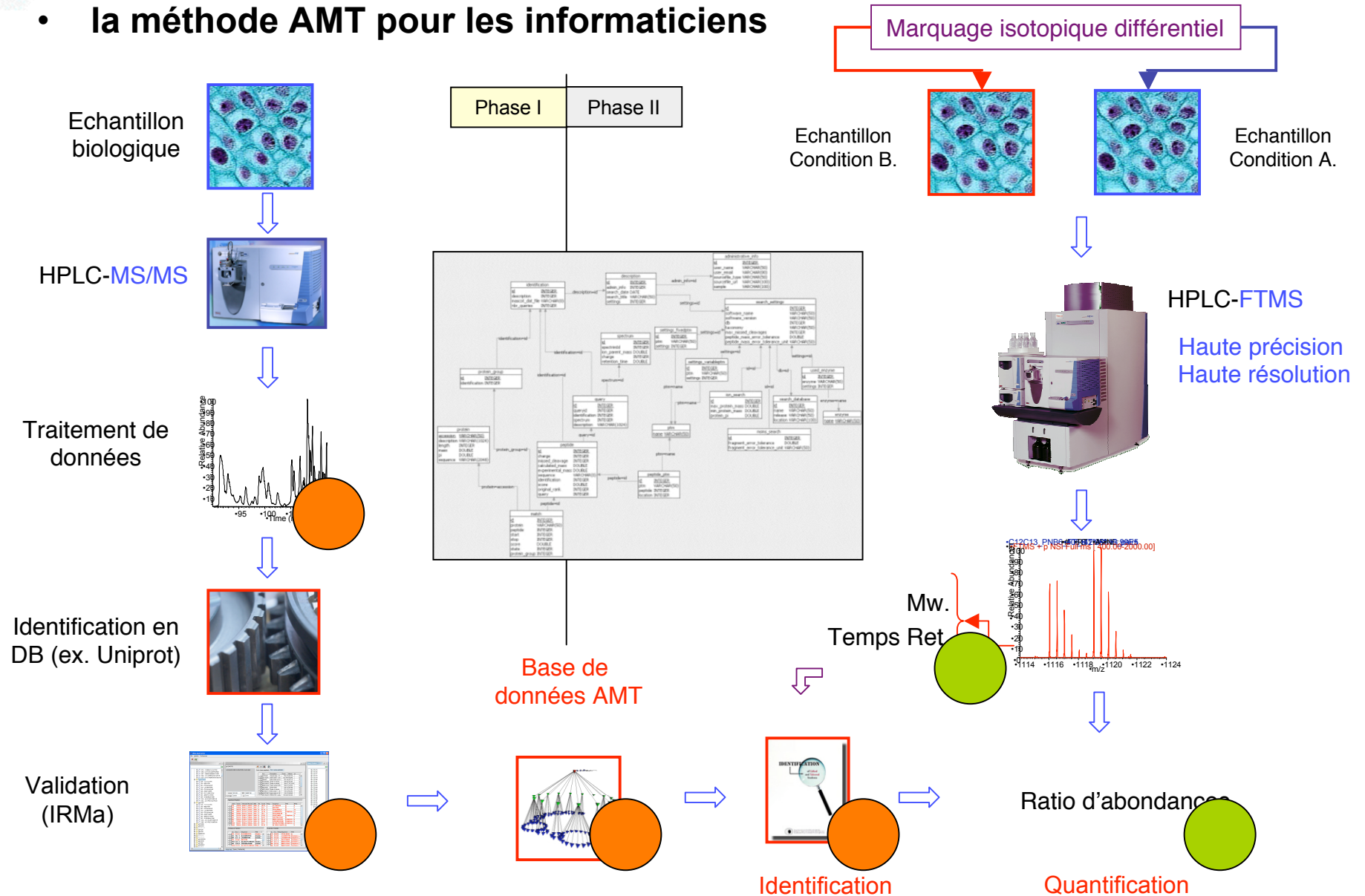
- Utilisé en routine sur la plate-forme
 - Ajout fréquent de nouvelles fonctionnalités
 - Installation au Service de Biochimie post-génomique & Toxicologie Nucléaire (CEA Marcoule)

Informatique pour la spectrométrie de masse

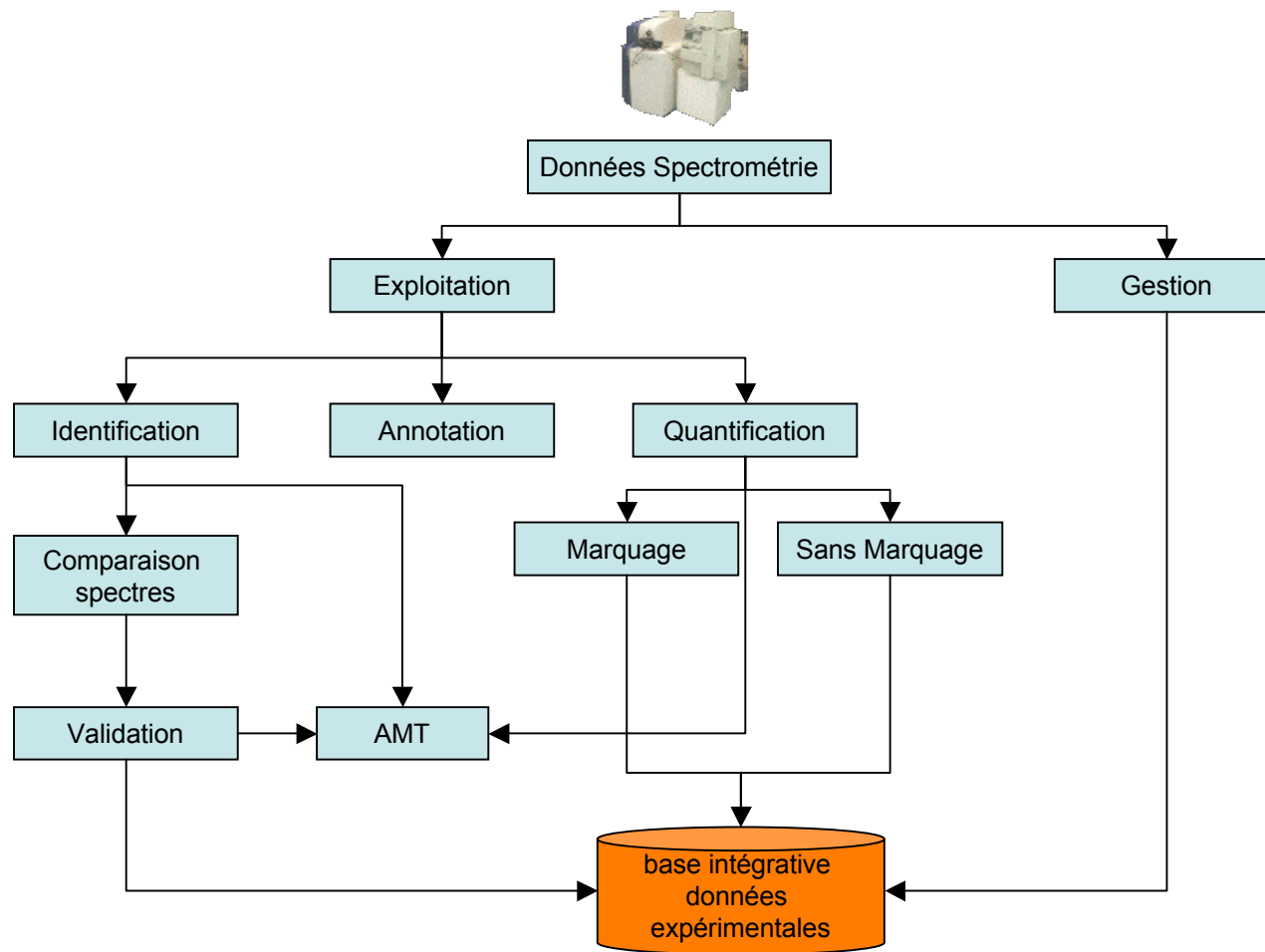


Méthode AMT

la méthode AMT pour les informaticiens



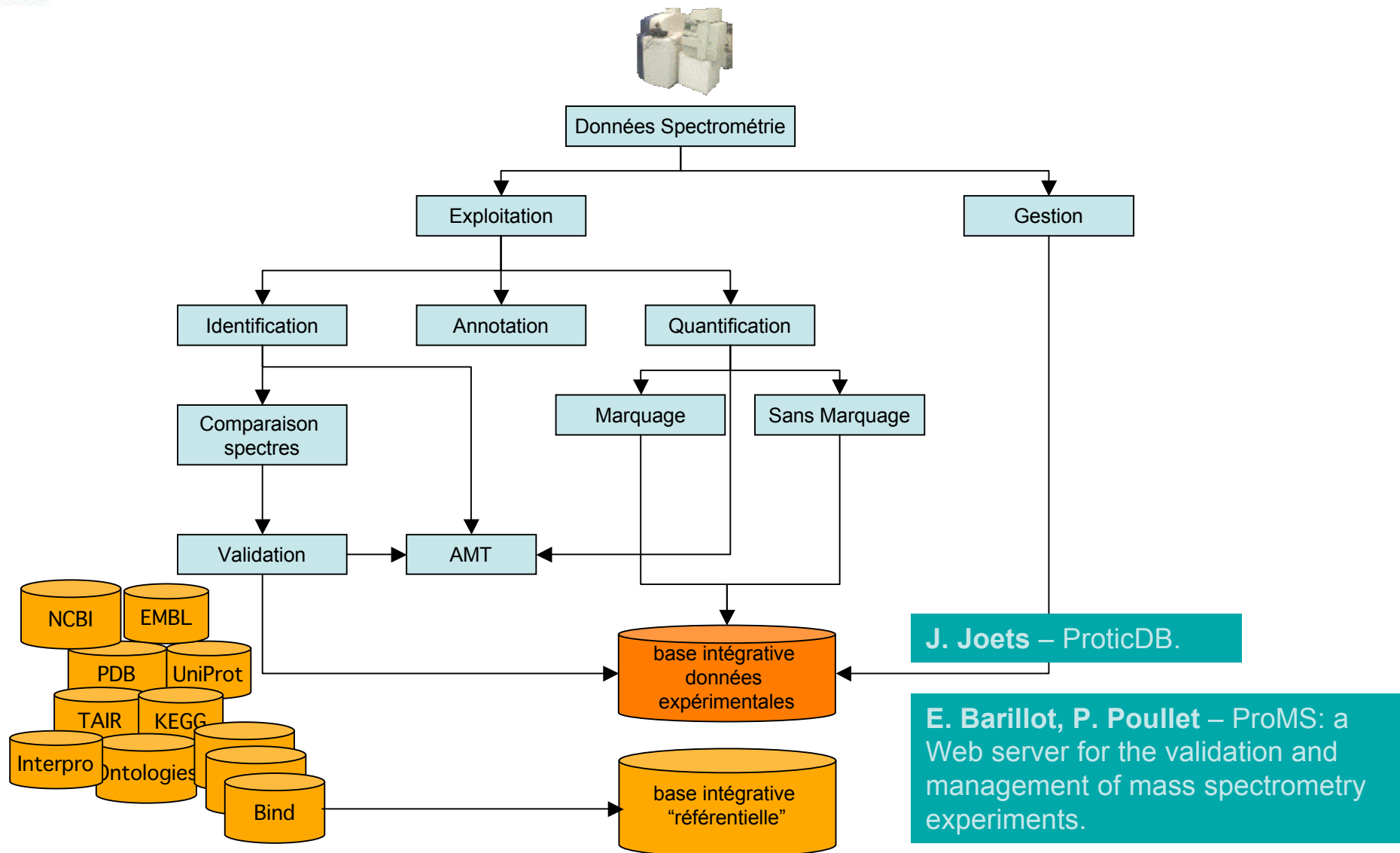
Intégration de données



- **Besoins**

- Études à l'échelle d'un organisme ou d'un compartiment cellulaire
- Analyse exploratoire (fouille de données)
- Collecte et structuration des données relatives à un projet
 - En spectrométrie de masse
 - *Identifications en spectrométrie de masse*
 - *Caractéristiques des échantillons*
 - *Résultats de quantification*
 - D'autres natures
 - *Génomique*
 - *Métabolique*
 - *Transcriptomique*
 - ...

Intégration de données



Laboratoire d'Etude de la Dynamique des Protéomes

ERM 0201 CEA/INSERM/UJF
Jérôme GARIN

CePRA

**Service
Administratif**

Céline FLEURY

**Equipe
« Lysosome »**

Agnès CHAPEL
Lucie CREPIN
Magali CHEMALI
Alain DUPUIS
Anaïs JENSEN
Agnès JOURNET

**Equipe
« Protéomique »**

Annie ADRAIT
Sabine BRUGIERE
Virginie BRUN
Magali COURT
Myriam FERRO
Maighread GALLAGHER
Michel JAQUINOD
Sylvie JAQUINOD-KIEFFER
Alexandra KRAUT
Lauriane KUHN
Mathilde LOUWAGIE
Marlène MARCELLIN
Christophe MASSELON
Laurent KELLNER

**Equipe
« Informatique et
Bio-informatique »**

Christophe BRULEY
Véronique DUPIERRIS
Mimi GREBICI
Jérémy TURBET Marianne
TARDIF