

Workshop "Analyse et interprétation des données en protéomique et métabolomique : points communs et spécificités"

**Mercredi 28 mars 2018
Lyon (Forum Labo)**

La protéomique et la métabolomique sont deux sciences "omiques" complémentaires pour l'étude du vivant et la médecine personnalisée, qui partagent notamment une technique analytique commune : la spectrométrie de masse. Nous sommes heureux de vous accueillir à ce workshop qui est constitué de deux parties visant à présenter un état de l'art sur l'analyse et l'interprétation des données dans chaque discipline, et à faire émerger des axes thématiques transverses

<https://framaforms.org/workshop-analyse-et-interpretation-des-donnees-en-proteomique-et-metabolomique-points-communs-et>

Programme :

10h-12h : Etat de l'art

- Stratégies d'acquisition/interprétation des données en protéomique [Christophe Bruley, CEA-Grenoble]
- Traitement et analyse des données en métabolomique [Etienne Thévenot, CEA-Saclay]
- Banques et dépôts de données en protéomique [Christine Carapito, CNRS-Strasbourg]
- Banques de données, bases de connaissance et dépôts de référence en métabolomique [Franck Giacomoni, INRA, Clermont Ferrand]
- Traitement et analyse des données en protéomique [Mélisande Blein-Nicolas, INRA-Moulon]
- Analyse des données omiques dans le contexte des réseaux métaboliques [Nathalie Poupin, INRA-Toulouse]

14h-16h : Ateliers participatifs pour dégager des axes de convergence

16h-16h30 : Omics integration and modeling: Prediction of proteins lifetime of tomato fruit [Isma Belouah, INRA-Bordeaux]

14h-16h
QUESTIONS OUVERTES

Intégration multi-omiques : questions générales

De l'intérêt d'intégrer les données métabolomiques et protéomiques ?

→ corrélation avec un phénotype, objectif descriptif (ex thématique des biomarqueurs), émettre des hypothèses, restreindre la liste de « candidats », approches dynamiques

Imagerie par spectrométrie de masse : retour d'expérience de Regis avec des données de lipidomique/metabolomique (imagerie, traitement de données avec metaspace) + des données de protéomique. Besoin de corréler ces données → possibilité d'utiliser MetExplore par exemple (outil assez simple d'utilisation mais nécessité d'avoir les bons identifiants) pour visualiser dans quelle voie on a des métabolites discriminants et des gènes différentiellement exprimés.

Question des biologistes à propos des biomarqueurs sélectionnés : est-ce qu'on n'a rien oublié ?

→ il peut être intéressant de positionner les acteurs moléculaires dans un réseau pour évaluer si certaines molécules qui auraient dû être présentes ont effectivement été vues ; il est important également de considérer que certaines molécules ont pu être écartées à différentes étapes du workflow (molécules mal retenues ou mal ionisées lors de l'acquisition, signaux à la limite du bruit lors de la détection des pics, variables redondantes lors de la sélection).

Corrélation entre protéomique et transcriptomique : ça ne converge pas ! Importance de prendre en compte les temps de réponse selon les types moléculaires

Outils permettant l'intégration multi-omiques

MetExplore : compartimentation cellulaire apparaît-elle dans MetExplore (pour les protéines) ?

→ Pour celles qui ont des localisations multiples il y a duplication des réactions.

Intégration dynamique possible un jour avec MetExplore ?

→ Il existe aussi MetaboAnalyst (peut faire du pathway enrichissement ; multi-omics).

ThermoFischer Cloud : intégration transcriptomique/protéomique (enrichissement par pathways).

Reactome fait enrichissement de pathways (attention : les annotations ne sont pas bonnes pour toutes les espèces !).

MixOmics (stats, R package): statistiques il faut que les jeux de données aient des comportements similaires (ex loi normale). Etape de transformation dans le workflow. Enjeu de la normalisation.

Pathway studio (text mining). Existe-t-il un outil équivalent (très productif mais très cher, équivalent avec IPA) ?

→ pas à notre connaissance

Banques de données

Difficile de se retrouver dans les banques de données en microbiologie

→ Avoir une démarche de choix des séquences organisme par organisme, choisir des banques spécialisées.

Remarque : Intérêt des bases de données 'espèce –spécifique » (ex : NextProt). Importance de la curation et de la qualité des données, des metadata (source des données).

Utiliser des banques de données organisme spécifique ou pas ?

→ Attention en biologie médicale on s'intéresse de plus en plus aux interactions microbes-hôte. Cf aussi pharmaco-toxicologie où on peut passer à côté d'éléments qui n'ont rien à voir avec un organisme de référence (ex : humain).

Questions diverses

Existe-t-il des approches de type « Fluxomique » en protéomique ?

→ time course avec marquage des protéines

Comment annoter des listes de milliers de métabolites ?

→ en fait on ne les annote pas tous, différentes approches (interrogation de banques, prédiction de structure, fragmentation MS/MS notamment) sont nécessaires suivant le type de métabolites

Question conceptuelle : ça ne ferait pas plus sens de coupler les métabolites et les protéines (déphasage entre transcriptomique) ?

→ tout dépend de la question

Intérêt des approches ciblées.

Quels critères pour fixer le nombre de répliques biologiques pour la métabolomique?

→ design en métabolomique ciblé sur une catégorie de métaboliques. Stratégie avec des pools. ça dépend des études.

Données brutes, que sont-elles ? Pour la métabolomique et la protéomique : fichier qui sort de l'instrument (il n'y a pas eu de prétraitement appliqué).

Outils qui permettraient de nous orienter sur les méthodes de normalisation et d'imputation ? Est-ce que ça dépend de l'instrument ?

→ Pas d'outil existant. Faire d'abord la description des données pour faire le bon choix de méthodes, mieux comprendre les aspects stats, comparer les effets des méthodes sur les données réelles (simuler le cas échéant certaines des données comme manquantes)

Dépôts de données

Lipidomique : dépôt des données brutes. Va-t-on forcer les gens à déposer ces données ? [musique !]
→ A priori dans MetaboLights. Il y a encore des progrès à faire en métabolomique. A l'EBI ils font le maximum mais les journaux ne forcent pas ça.

En protéomique une solution un peu souple a été mise en place pour le dépôt de données (soumission complète ou partielle)

Y-a-t-il un consortium à l'international au sujet du dépôt de données?

→ Il y a des avancées ex : Workflows Galaxy et MetaboLights compatibles (projet phenomenal)

Concernant W4M les données sur l'instance pourraient être directement versées dans le dépôt MetaboLights ?

→ en cours, en connexion avec l'EBI (discussion aussi avec les constructeurs pour la gestion des formats).

ProteomeXchange : est ce que l'origine de données est mentionnée. En principe oui. enjeu des metadata : à renseigner au maximum.

About Elixir (<https://www.elixir-europe.org/>)

Vise à harmoniser les initiatives au niveau européen. La métabolomique et la protéomique font maintenant partie d'une communauté. Volonté de fédérer, ressources maintenues, offre de formations pour les utilisateurs et développeurs.

Mission autour de l'interopérabilité nécessaire pour que le biologiste n'ait pas à se soucier des problèmes de format, d'utilisation des données

Remarques diverses

Félicitations pour cette journée !

Groupes (pas eu besoin !)

Banques de données /repositories

Stats et réseaux

Integration proteo/metabo