

Fiche de données

“Métabolomique”

Objectif: décrire les types de données de métabolomique manipulés lorsqu'on veut exploiter ces données pour différentes analyses ou faire de l'intégration de “omique”. Cette fiche est volontairement orientée “donnée” et pas “analytique”.

Rédactrices/Rédacteurs

Franck Giacomoni - MetaboHUB Clermont Ferrand - CATI EMPREINTE - INRAE

Marie Tremblay-Franco - MetaboHUB Toulouse - CATI EMPREINTE - INRAE

Mélanie Pétéra - MetaboHUB Clermont Ferrand - CATI EMPREINTE - INRAE

Définitions

Le métabolome désigne l'ensemble des substances chimiques à petites molécules présentes dans un échantillon biologique. L'échantillon biologique peut être une cellule, un organite cellulaire, un organe, un tissu, un extrait de tissu, un biofluide ou un organisme entier. Les petites molécules chimiques trouvées dans un métabolome donné peuvent inclure à la fois des métabolites endogènes qui sont naturellement produits par un organisme (tels que des acides aminés, des acides organiques, des acides nucléiques, des acides gras, des amines, des sucres, des vitamines, des cofacteurs, des pigments, des antibiotiques, etc.) ainsi que des produits chimiques exogènes (tels que des médicaments, des contaminants environnementaux, des additifs alimentaires, des toxines et d'autres xénobiotiques) qui ne sont pas naturellement produits par un organisme. L'identification et la quantification des métabolites permettront de connaître les modulations métaboliques engendrées par l'exposition à un facteur d'intérêt (régime alimentaire, médicament, xénobiotique...).

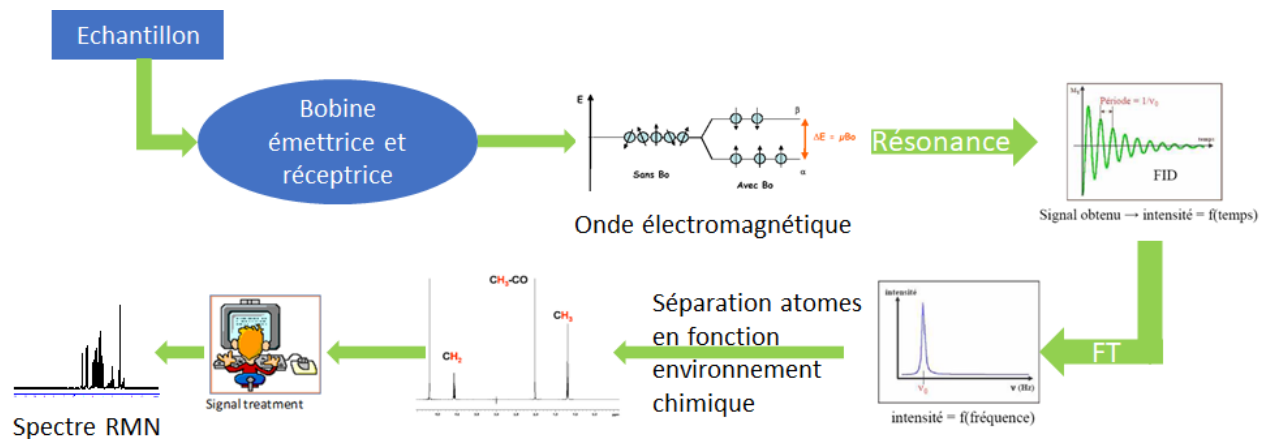
La métabolomique est la dernière-née des sciences 'omiques', apparue au début des années 2000. Elle est définie comme l'analyse de l'ensemble des petites molécules présentes et accessibles dans un système biologique. L'intérêt de cette approche réside dans le fait qu'elle permet d'avoir une vision intégrée du métabolisme, proche du phénotype. Cependant, elle consiste en un processus multi-étape, nécessitant des compétences multidisciplinaires, depuis le design expérimental, en passant par l'analyse des échantillons biologiques, le traitement des données générées, pour aller jusqu'à l'exploitation des résultats.

Description des données de métabolomique

Les principales technologies :

La spectroscopie par résonance magnétique nucléaire (RMN) et la spectrométrie de masse couplée à la chromatographie liquide (LC-MS) ou gazeuse (GC-MS) sont les deux principales technologies utilisées actuellement en métabolomique.

La **résonance magnétique nucléaire** est une technique de spectroscopie appliquée aux particules ou ensembles de particules atomiques qui ont un spin nucléaire non nul. Le noyau de l'atome considéré absorbe les rayonnements électromagnétiques d'une fréquence spécifique en présence d'un fort champ magnétique.



La **spectrométrie de masse** est une technique physique d'analyse permettant de détecter et d'identifier des molécules par mesure de leur masse, et de caractériser leur structure chimique. Son principe réside dans la séparation en phase gazeuse ou liquide de molécules chargées (ions) en fonction de leur rapport masse/charge (m/z).

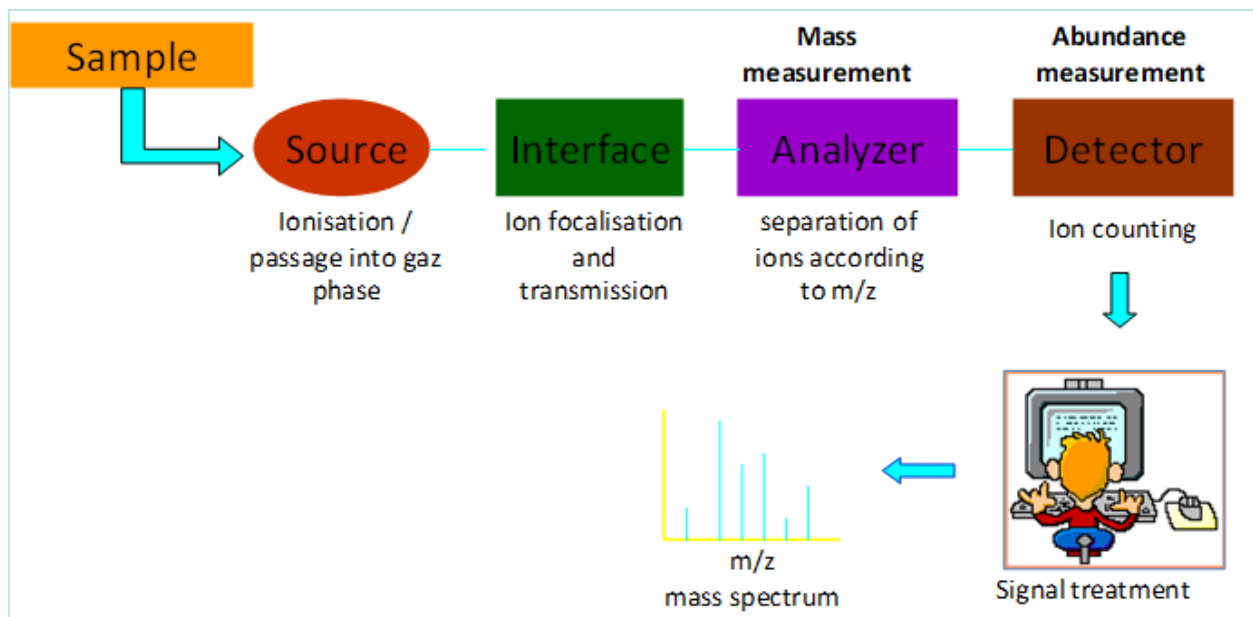


Tableau global

Tableau général décrivant globalement une expérience en métabolomique.

	Description
<i>Echantillon</i>	<p>Source : espèce, variété, génotypes sauvage/mutant, phénotypes ;</p> <p>Extraction : organisme entier, organe/tissu/biofluides/lignées cellulaires, stade de développement/âge, ...</p> <p>Préparation : elle va être spécifique de trois facteurs :</p> <ul style="list-style-type: none"> ● la technique analytique ● la famille chimique ciblée ● le type de matrice (fluide / organe / gaz...)
<i>Conditions expérimentales</i>	<p>Type d'expériences : comparaison de traitement, régime alimentaire, génotype, cinétique, ...</p> <p>Facteur(s) : dépendant des expérimentations : généralement à 1 facteur, mais de plus en plus multifactoriel et longitudinal</p>
<i>Protocoles</i>	Voir quel niveau on veut ?

<i>Données</i>	<p>Objectifs :</p> <p><i>Approche ouverte</i> : identifier les métabolites à partir d'une stratégie d'analyse <i>sans a priori</i> sur les familles chimiques recherchées dans le but :</p> <ul style="list-style-type: none"> • soit de mettre en évidence les variables discriminantes d'un état par rapport à un autre (exemple patient humain malade/sain ou consommateurs fort de bananes vs faible) • soit d'étudier le réseau métabolique d'un organisme (approche mécanistique) <p><i>Approche ciblée</i> : Identifier en présence / absence de métabolites pré-définis</p> <p>Type de données :</p> <ul style="list-style-type: none"> • Spectres prétraités selon la technique analytique : intensités relatives (données continues positives) <p>Caractéristiques majeurs des données :</p> <ul style="list-style-type: none"> • RMN : ? • MS : Hétérogénéité des redondances entre variables, effets analytiques connus, forte sensibilité à des facteurs externes non nécessairement maîtrisables ou même mesurables (facteurs environnementaux par exemple), difficultés liées aux bruits (signaux parasites, signaux bruités à niveau de bruit variable...).
----------------	--

Tableau sur les données de métabolomique

Tableau décrivant les principaux types de données générés des expériences classiques de métabolomique et jugées comme d'intérêt pour mener une démarche d'intégration multi-omique.

Données	Types de données	Description	Mesures
Intensité relative en RMN	<i>FLOAT</i> : sous forme de table - une valeur par bucket pour chaque échantillon biologique	Mesure d'intensité relative de signal: aire sous la courbe calculée pour chaque bucket (fenêtre de largeur fixe) et normalisée (divisée par l'intensité totale du spectre, poids d'échantillon...)	intensité normalisée par échantillon biologique

Intensité relative en MS	<i>FLOAT</i> : sous forme de table - une valeur par couple "Temps de rétention; Masse sur charge" pour chaque échantillon biologique	Mesure d'intensité relative de signal : surface ou hauteur des pics chromatographiques associés à chaque ion extrait ; les ions peuvent parfois être filtrés ou poolés ensemble par composé selon les méthodes ; si nécessaire, les intensités sont corrigées d'effets analytiques ou d'effets majeurs de design	intensité (potentiellement normalisée) par échantillon biologique
--------------------------	--	--	---

Banques et dépôts en métabolomique :

Quelques banques de références en métabolomique

[PubChem](#) - NIH (USA)

[ChEBI](#) - EBI Chemical Entities of Biological Interest

[LipidMaps](#) - Wellcome Trust

[COCONUT](#) - COLleCtion of Open Natural ProdUcTs

Human Metabolome DataBase ([HMDB](#)) et l'ensemble des [base de données](#) du Wishart laboratory

[MassBank.eu](#) & MassBank of North America ([MoNA](#))

[PeakForest.org](#) - MetaboHUB (FR)

Dépôts internationaux en métabolomique

[Metabolights](#) - An EBI database for Metabolomics experiments and derived information (EU)

[Metabolomics Workbench](#) - An NHI database (USA)

Exemples de données

Exemple de workflows de traitement de données de référence

Sur la page <https://workflow4metabolomics.org/ref-workflows-histories> les historiques notés W4M000x sont des historiques de références et possèdent un DOI. Il peuvent être notamment consultés pour le prétraitement des spectres bruts et l'obtention de la matrice de données générée pour les analyses statistiques.

Publication associée : Guitton Y., Tremblay-Franco M., Le Corguillé G., Martin J.F., Pétéra M., Roger-Mele P., Delabrière A., Goulitquer S., Monsoor M., Duperier C., Canlet C., Servien R., Tardivel P., Caron C., Giacomoni F., Thévenot E.A., Create, run, share, publish, and reference your LC–MS, FIA–MS, GC–MS, and NMR data analysis workflows with the Workflow4Metabolomics 3.0 Galaxy online infrastructure for metabolomics, *The International Journal of Biochemistry & Cell Biology*, 2017, ISSN 1357-2725, <http://dx.doi.org/10.1016/j.biocel.2017.07.002>

Ce papier est aussi disponible sur [open archive HAL](#).