HAL
open science

Développement d'outils informatiques pour le traitement et l'analyse des données en RMN-1H, dans un contexte d'expériences métabolomiques

Daniel Jacob, Catherine Deborde, Stéphane Bernillon, Annick Moing

To cite this version:

Daniel Jacob, Catherine Deborde, Stéphane Bernillon, Annick Moing. Développement d'outils informatiques pour le traitement et l'analyse des données en RMN-1H, dans un contexte d'expériences métabolomiques. 5. Journées scientifiques du Réseau Français de Métabolomique et Fluxomique, Labo/service de l'auteur, Ville service., May 2011, Paris, France. hal-02747402

HAL Id: hal-02747402

https://hal.inrae.fr/hal-02747402v1

Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

O8

Développement d'outils informatiques pour le traitement et l'analyse des données en RMN-^{1}H, dans un contexte d'expériences métabolomiques.

Daniel Jacob[1, 2], Catherine Deborde[1], Stéphane Bernillon[1], Annick Moing[1]

1- PMFB –UMR 1332, INRA, F-33140 Villenave d'Ornon
2- CBIB – Université Bordeaux 2 – F-33076 Bordeaux Cedex

Objectif : Visualiser un jeu de données métabolomiques complet et permettre d'identifier les groupes d'échantillons et les variables analytiques pertinentes (VAP) correspondantes pouvant expliquer leur discrimination.

Les techniques analytiques, dont la RMN-^{1}H, produisent des signatures spectrales contenant l'information utile recherchée. Outre cette information, c'est aussi et surtout ses variations en fonction des facteurs du plan expérimental choisi que l'on souhaite mettre en évidence.

Il convient tout d'abord de gérer ces données, en leur associant des métadonnées (ontologies) permettant de les classer selon leur groupe d'appartenance. Ensuite, afin d'extraire l'information de la signature spectrale, des traitements sur les spectres (lignes de bases, bruits, réalignement) sont nécessaires. Selon l'objectif visé, les traitements seront différents :

1/ identification de métabolites : une déconvolution permettra de séparer et de quantifier les pics (lorentziennes) cachés notamment sous des épaulements ;

2/ recherche des VAP sur les signatures: extraction de la partie informative des spectres à l'aide de la transformée en ondelettes discrète (DWT) ; découpage des spectres en zones distinctes (non chevauchantes), réalignement des spectres par zone, découpage des zones en régions (buckets) de tailles variables (1 VAP = 1 bucket).

Pour interpréter un jeu de données métabolomiques complet, deux stratégies complémentaires sont utilisées : 1/ une visualisation interactive des spectres superposés et coloriés selon différents critères (basés sur les métadonnées) ; 2/ un scénario d'analyses biostatistiques incluant le traitement des données manquantes, la normalisation des données, la sélection (réduction) éventuelle des VAP et des analyses multivariées (PCA, ICA, ou PLS) ou univariées.

Pour chaque étape de ces deux stratégies, il existe de nombreux outils (« home made toolboxes », logiciels gratuits ou commerciaux) nécessitant de nombreuses manipulations afin de constituer un pipe-line complet. Une base de données de gestion de profils métabolomiques à l'instar de MeRy-B [1] (plantes, http://www.cbib.u-bordeaux2.fr/MERYB), permet déjà de gérer l'ensemble des données et des métadonnées. En associant les outils de traitements précités, une telle base de données se transforme en un véritable outil de travail performant en simplifiant l'ensemble des tâches tout le long de la chaine des traitements.

Référence Bibliographique :

[1] MeRy-B: a web knowledgebase for the storage, visualization, analysis and annotation of plant NMR metabolomic profiles. Hélène Ferry-Dumazet, Laurent Gil, Catherine Deborde, Annick Moing, Stéphane Bernillon, Dominique Rolin, Macha Nikolski, Antoine de Daruvar and Daniel Jacob. BMC Plant Biology, 2011, *in press.*