



HAL
open science

Les réseaux moléculaires, une approche bio-informatique globale pour interpréter les données de spectrométrie de masse tandem

Louis-Félix Nothias-Scaglia, Mélissa Esposito, Jean Costa, Julien Paolini,
David Touboul, Marc Litaudon

► To cite this version:

Louis-Félix Nothias-Scaglia, Mélissa Esposito, Jean Costa, Julien Paolini, David Touboul, et al.. Les réseaux moléculaires, une approche bio-informatique globale pour interpréter les données de spectrométrie de masse tandem. *Spectra Analyse*, 2015, 307, pp.73-78. hal-01254171

HAL Id: hal-01254171

<https://hal.science/hal-01254171>

Submitted on 11 Jan 2016

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.

Louis-Félix NOTHIAS-SCAGLIA^{1,2*}, Mélissa ESPOSITO^{1,2}, Jean COSTA², Julien PAOLINI²,
David TOUBOUL¹, Marc LITAUDON^{1*}

Les réseaux moléculaires, une approche bio-informatique globale pour interpréter les données de spectrométrie de masse tandem

RÉSUMÉ

L'interprétation des données issues de l'analyse par spectrométrie de masse tandem (MS²) d'un mélange complexe d'origine naturelle est une tâche ardue, généralement effectuée de manière manuelle par un spécialiste. Récemment, une approche bio-informatique innovante, nommée « réseaux moléculaires » (*molecular networking*), a été développée par un laboratoire américain de l'Université de Californie de San Diego. Cette approche est capable d'organiser et de visualiser des données MS² sous la forme de carte de similarité spectrale, mettant en lumière l'existence de groupes spectraux ainsi que leurs degrés de similarité. Grâce à l'utilisation de réseaux moléculaires, des mélanges complexes peuvent être comparés *in silico* et annotés, facilitant l'étude du métabolome d'organismes vivants. Avec la mise en ligne de la plateforme GNPS (<http://gnps.ucsd.edu>), cette approche est désormais accessible à l'ensemble de la communauté scientifique. Cette fiche d'application présente le potentiel des réseaux moléculaires, au travers de deux exemples qui s'appuient sur l'analyse par chromatographie liquide couplée à la spectrométrie de masse tandem (LC-MS²) d'extraits végétaux du genre *Euphorbia*.

MOTS-CLÉS

Réseaux moléculaires – spectrométrie de masse tandem – MS/MS – MS² – spectrométrie de masse multi-niveaux – MSⁿ – bioinformatique – GNPS – *Euphorbia*

Molecular networks, a bioinformatic tool for tandem mass spectrometry data interpretation

SUMMARY

Data interpretation from analysis by tandem mass spectrometry (MS²) of a complex mixture of natural origin is a daunting task, usually done manually by a specialist. Recently, a novel bioinformatics approach, called «molecular networks» or «molecular networking», was developed by an American laboratory at the University of California San Diego. This approach is able to organize and view MS² data in the form of a spectral similarity map, highlighting the presence of spectral groups and their degree of similarity. Through the use of molecular networks, complex mixtures may be compared in silico and annotated, facilitating the study of living organisms metabolome. With the launch of the platform GNPS (<http://gnps.ucsd.edu>), this approach is now available to the entire scientific community. This application note shows the potential of molecular networks, through two examples that are based on the analysis by liquid chromatography coupled with tandem mass spectrometry (LC-MS²) of plant extracts from Euphorbia genus.

KEYWORDS

Molecular networking – tandem mass spectrometry – MS/MS – MS² – multiple-stage mass spectrometry – MSⁿ – bioinformatics – GNPS – Euphorbia

Abréviations

LC : liquid chromatography – chromatographie en phase liquide.
Métabolome : ensemble des métabolites primaires et secondaires présents dans un échantillon.
LC-(HR)MS : liquid chromatography coupled to high resolution mass spectrometry – chromatographie liquide couplée à la spectrométrie de masse haute résolution
MS² : tandem mass spectrometry or tandem mass spectrum – spectrométrie de masse tandem ou spectre de masse tandem.
MSⁿ : multiple-stage mass spectrometry – spectrométrie de masse multi-niveaux.
GNPS : Global Natural Products Social molecular networking (<http://gnps.ucsd.edu>)

¹ Institut de Chimie des Substances Naturelles, CNRS, ICSN UPR 2301, Labex CEBA, Université Paris-Saclay, 91198, Gif-sur-Yvette, France

² Laboratoire de Chimie des Produits Naturels, UMR CNRS SPE 6134, Université de Corse, 20250, Corte, France

* Pour correspondance : louisfelix.nothias@gmail.com, marc.litaudon@cnrs.fr

I - Introduction

L'analyse globale de la composition chimique d'extraits naturels complexes issus d'organismes vivants est une tâche laborieuse qui nécessite la mise en œuvre de techniques analytiques modernes. La chromatographie en phase liquide couplée à la spectrométrie de masse tandem (LC-MS²) est devenue l'une des techniques phares utilisées pour l'analyse de tels mélanges, représentatifs des métabolomes des espèces considérées (1). Lors d'une analyse MS², le spectromètre de masse tandem sélectionne un ion précurseur avec le premier analyseur (généralement un quadripôle), puis cet ion est fragmenté par dissociation induite par collision et les fragments ainsi formés sont analysés ; soit en mode basse résolution en masse (instrument de type quadripôle ou trappe d'ion, en particulier) ou haute résolution en masse (instrument de type temps de vol ou orbitrap) afin d'obtenir un spectre MS². Ce dernier est composé uniquement d'ions produits par la fragmentation de l'ion précurseur et est donc riche en informations structurales, contrairement aux spectres MS (dans lesquels tous les ions intacts sont détectés). Toutefois, l'analyse d'un mélange complexe en mode acquisition automatique de spectres MS² conduit à l'obtention de plusieurs centaines, voire plusieurs milliers de spectres de ce type. L'interprétation de ces données est habituellement réalisée par un spécialiste qui, malgré son expertise et sa détermination, ne pourra analyser que partiellement les résultats. Des outils bio-informatiques ont été développés ces dernières années pour faciliter l'annotation et la comparaison des analyses LC-(HR)MS pour l'analyse de petites molécules (2). Cependant, aucune ne prend en charge la dimension MS² contrairement aux algorithmes utilisés en protéomique où des peptides de digestion sont séquencés en MS² et où les interrogations en base de données sont réalisées en routine depuis dix ans (3). C'est dans ce contexte que s'est développé le recours à des réseaux moléculaires (*molecular networking*), une approche bio-informatique permettant l'interprétation globale des données (4–7). Conçu par les groupes de Pieter C. Dorrestein et de Nuno Bandeira (Université de Californie, San Diego), cet outil innovant repose sur un postulat simple : les molécules structurellement proches partagent des voies de fragmentations similaires en spectrométrie de masse conduisant à des fragments ou des pertes de neutre en commun. Ainsi, dans les réseaux moléculaires, les spectres MS² d'un jeu de données sont comparés deux à deux par alignement spectral, puis un indice de corrélation, appelé similarité cosinus (*cosine score*, CS), est attribué entre chaque paire de spectres. L'algorithme d'alignement spectral, sur lequel reposent les réseaux moléculaires, est dérivé d'outils bio-informatiques développés originellement pour l'analyse des protéines. Les

données sont préalablement simplifiées par la fusion des spectres considérés comme identiques (MS-Cluster), puis l'alignement spectral est effectué sur chaque paire de spectre MS² sur la base des différences entre les valeurs des m/z , et de leur intensité relative (4). Deux spectres superposables de ce type ont un CS de 1, alors que deux spectres sans ion ou perte de neutre commun ont un CS de 0. Les spectres ayant un CS supérieur à une certaine valeur (typiquement 0,7) ont une forte similarité spectrale ; ils sont alors considérés comme issus de la fragmentation de constituants chimiquement apparentés. Les résultats sont ensuite interprétés graphiquement sous la forme de réseaux de similarité spectrale, dénommés par extrapolation réseaux moléculaires (terminologie employée par l'équipe de l'UCSD). Les premières applications ont été réalisées à partir d'un programme écrit en langage MATLAB et du logiciel de visualisation Cytoscape (4–7).

L'utilisation de réseaux moléculaires (ou réseaux de similarité spectrale) permet : (i) de détecter un ensemble d'analogues structuraux à partir du spectre MS² d'une molécule de référence ; (ii) d'identifier des groupes spectraux MS² ; et (iii) de permettre la comparaison fine entre les analyses MS² de différents échantillons. Sa puissance réside dans sa capacité à organiser les données y compris en condition *ex nihilo*, c'est-à-dire sans connaissance préalable de l'échantillon (tel que des spectres MS² de composés de référence ou des informations sur les métabolites présents).

En janvier 2014, les groupes de Pieter C. Dorrestein et de Nuno Bandeira ont mis en ligne la plateforme internet *Global Natural Product Social molecular networking* (GNPS), dédiée à la déréplication et l'interprétation des données MS² (<http://gnps.ucsd.edu>). Cette plateforme regroupe différents outils pour l'annotation automatique, en particulier des banques pour ce type de spectres. Elle permet de générer des réseaux de similarité spectrale à partir de n'importe quelle analyse MS². La plateforme GNPS héberge également un répertoire public de dépôt de données (MassIVE, *Mass spectrometry Interactive Virtual Environment*). Ces dernières peuvent être annotées par la communauté des utilisateurs et régulièrement ré-analysées (*continuous identification*) afin de bénéficier de l'enrichissement continu des banques spectrales. Récemment, un outil (en version beta) pour réaliser une comparaison entre les spectres MS² expérimentaux et les spectres MS² simulés (*in silico*) a été ajouté. À ce jour, l'article qui décrit le développement de la plateforme GNPS n'est pas encore publié. Néanmoins, celle-ci semble rencontrer un véritable engouement dans la communauté scientifique (plus de 8 000 utilisateurs en septembre 2015) (8).

Dans le but de présenter la plateforme GNPS à la communauté francophone, deux essais de réseaux moléculaires, réalisés à partir d'analyse LC-MS² de plantes appartenant au genre *Euphorbia*, sont présentés ci-dessous. Le premier exemple traite de la détection d'analogues structuraux dans un extrait d'*Euphorbia amygdaloides* subsp. *semiperfoliata* (sélectionné comme échantillon de référence). Douze diterpènes, de type esters de jatrophane (noté JE), ont été isolés à partir de cette espèce et leur voie de fragmentation étudiée par MSⁿ (9,10). Ces composés sont répartis en deux groupes (JE-A et JE-B) selon la substitution du squelette diterpénique. L'étude de la fragmentation de ces constituants par MSⁿ a montré que les molécules de ces deux groupes se fragmentaient de façon comparable (adduit [M+Na]⁺ majoritaire, élimination des esters par perte neutre) mais qu'elles avaient des empreintes spectrales différentes au niveau du noyau diterpénique. Le second exemple vise à comparer les profils métabolomiques LC-MS² de deux espèces du genre *Euphorbia* (*Euphorbia amygdaloides* subsp. *semiperfoliata* et *Euphorbia dendroides*).

II - Matériels et méthodes

1. Matériel végétal et extraction

Voir articles (9,13).

2. Analyse LC-MS²

Un spectromètre de masse hybride à trappe d'ion linéaire associé à un détecteur Orbitrap (LTQ-XL Orbitrap, Thermo Fisher Scientific, Les Ulis, France) équipée d'une source ESI opérant en mode ionisation positive a été utilisé. Cet appareil est capable de réaliser des expériences de spectrométrie de masse multi-niveaux en haute résolution. Les spectres MS² ont été acquis sur la plage m/z 150 à 1 000. Dans le mode balayage complet en MS, la résolution de la distance à mi-hauteur de l'analyseur Orbitrap a été fixée à 30 000, et à 15 000 pour les spectres MS². Le mode d'acquisition automatique de spectres a été paramétré pour réaliser des spectres MS² sur les trois ions les plus intenses en MS, suivi d'une exclusion de 40 sec après huit répétitions. Les paramètres instrumentaux ont été réglés comme suit : *source voltage* : 5 kV, *lens 1 voltage* : -15 V, *capillary temperature* : 275 °C, *gate lens voltage* : -35 V, *capillary voltage* : 25 V, *tube lens voltage* : 65 V. Les paramètres de CID ont été fixés : énergie de collision fixée à 30 % du maximum avec un temps d'activation de 30 ms. La LC a été réalisée avec un système HPLC Ultimate 3000 system (Dionex, Voisins-le-Bretonneux, France) composé d'un dégazeur, d'une pompe quaternaire, d'un «autosampler», d'un four à colonne, d'un détecteur à barette de diodes. La séparation a été réalisée sur

une colonne C₁₈ (Sunfire, 150 × 2.1 mm × 3.5 μm; Waters, Guyancourt, France), équipée d'une pré-colonne. La température de la colonne a été fixée à 25 °C. L'élution a été conduite avec une phase mobile composée d'eau + 0.1 % d'acide formique (AF) (A) et d'acétonitrile + 0.1 % AF (B), avec un gradient 5 à 95 % B en 40 min, puis d'un pallier 100 % B de 10 min à un débit de 250 μL/min. Le volume d'injection a été fixé à 10 μL.

3. Réalisation de réseaux moléculaires avec GNPS et Cytoscape

Les réseaux moléculaires MS² ont été générés sur la plateforme internet GNPS (<http://gnps.ucsd.edu>). Au préalable, les données des analyses sont converties au format .mzXML ou .mgf, avec le logiciel MS-Convert (ProteoWizard, <http://proteowizard.sourceforge.net/>), puis elles sont envoyées sur GNPS à l'aide d'un client FTP. Les spectres ont été traités en éliminant les ions situés dans un intervalle de +/- 17 Da autour du m/z du précurseur. Les données ont ensuite été simplifiées en utilisant MS-Cluster avec une tolérance de 0,05 Da pour l'ion précurseur et de 0,05 Da pour les ions produits en MS² pour créer un spectre consensus. Les spectres consensus contenant moins de deux spectres n'ont pas été retenus. L'ensemble des résultats a été comparé aux spectres de la banque spectrale GNPS. Les spectres de la banque spectrale ont été traités de la même manière que les spectres de l'analyse MS². Les données générées ont été déposées sur le MassIVE Public dataset de la plateforme GNPS sous les références : MSV000079384 / MSV000079385.

Ensuite, des réseaux moléculaires sont générés et les résultats peuvent être pré-visualisés directement en ligne sur GNPS. Les données sont ensuite importées hors ligne et représentées dans le logiciel Cytoscape 3.1.1 (<http://www.cytoscape.org/>) (11). Le calque de représentation AllegroLayout (<http://allegroviva.com/allegrolayout2/>) a été utilisé pour organiser la disposition spatiale du réseau dans Cytoscape, configuré avec un champ de force répulsif (mode *weak clustering* pondéré par le *cosine score*). Les spectres MS² ayant une forte similarité spectrale (CS > 0,7) sont représentés sous forme de perles (*node* en anglais). Ainsi, les liaisons entre ces perles apparaissent sur les réseaux moléculaires dès lors que le score CS est supérieur à 0,7 et qu'au moins six ions communs sont détectés. Les liaisons inter-spectre MS² ont été conservées seulement si, pour chacun des deux spectres concernés, elles figuraient mutuellement parmi les 10 spectres MS² les plus similaires. La taille des perles est proportionnelle à l'intensité du précurseur du spectre et l'épaisseur des connections inter-perles est proportionnelle à la valeur du CS. Pour chaque perle, les spectres peuvent être visualisés sur GNPS à l'aide d'un lien internet associé.

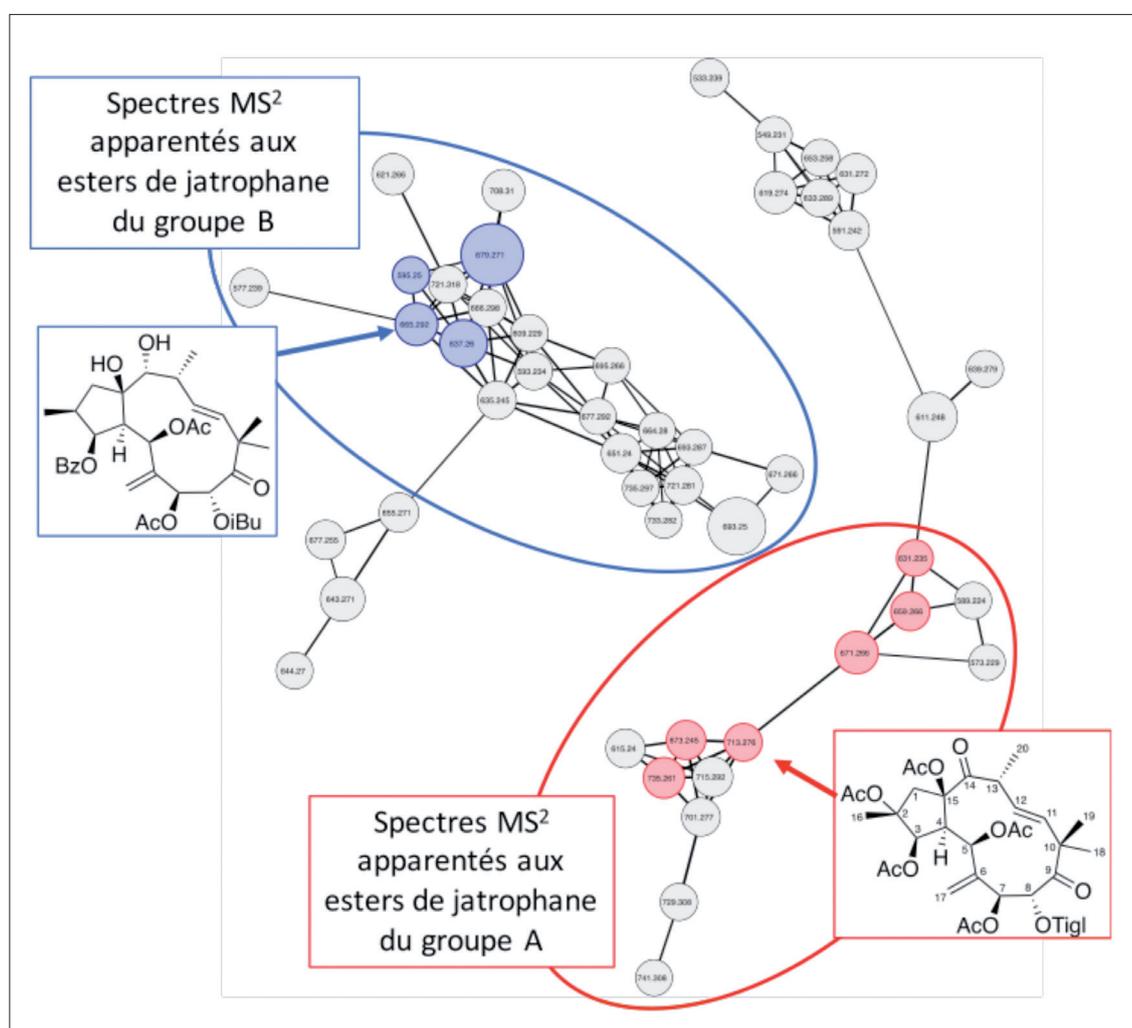
III - Illustration du potentiel des réseaux moléculaires sur des données MS² issues de l'analyse LC-MS² d'extraits de plantes du genre *Euphorbia*

1. Détection d'analogues structuraux chez *E. amygdaloides* subsp. *semiperfoliata*

L'extrait d'*E. amygdaloides* subsp. *semiperfoliata* ainsi que les douze esters de jatrophane JE (composés de références) isolés à partir de cette espèce ont été analysés par LC-MSⁿ en mode MS² dynamique, avec un spectromètre hybride associant une trappe d'ion linéaire et un détecteur Orbitrap (LTQ XL Orbitrap, Thermo Fisher Scientific). Les données correspondantes ont été déposées dans MassIVE (titre : *Euphorbia amygdaloides* ssp. *semiperfoliata*). Le résultat de la génération de réseaux de similarité spectrale est consultable en ligne sur GNPS. Les deux principaux réseaux moléculaires de ce jeu de données sont présentés sur la Figure 1. Les perles rouges correspondent aux spectres MS² de références de JE-A tandis

que les perles bleues à ceux de références de JE-B (9,10). Grâce à l'ajout des références, des spectres MS² apparentés ont été détectés comme étant des analogues probables sur la base de la forte similarité spectrale. Par ailleurs, l'annotation de ces analogues a pu être confortée par l'observation des fragmentations caractéristiques dans les spectres MS² mais aussi par la détermination de formules brutes compatibles avec des JEs. Cependant, il faut souligner qu'un certain nombre de difficultés est apparue lors de l'interprétation de ces réseaux de similarité spectrale. En premier lieu, des couples de spectres séparés par une unité de masse et retrouvés dans le même groupe spectral ont été observés. Ces spectres MS² ont été interprétés comme étant ceux de l'ion mono-isotopique et celui de son premier isotope, correspondant à cette même molécule ayant un atome sous la forme isotopique contenant un neutron supplémentaire (deutérium pour le proton, ¹³C pour le ¹²C, ¹⁷O pour l'¹⁶O,...). Cette interprétation a été confirmée par l'observation de temps de rétention similaire et l'inspection des données MS pour ces couples d'ions. Certains spectromètres de masse tandem sont capables d'exclure les ions du massif isotopique de façon dynamique et peuvent limiter ces faux positifs. De plus, certains logiciels

Figure 1
Réseaux moléculaires générés à partir d'une analyse LC-MS² du latex d'*Euphorbia amygdaloides* subsp. *semiperfoliata* réalisée avec un LTQ-XL Orbitrap. Les perles colorées correspondent à des spectres MS² identifiés par comparaison avec des standards d'esters de jatrophane.



propriétaires sont en mesure de supprimer post-acquisition ces spectres artéfactuels.

Une autre difficulté résulte de la simplification des données par la fusion de spectres, lors de l'utilisation de MS-Cluster, aboutissant à la création de spectres consensus. En effet, dans une analyse LC-MS², plusieurs spectres MS² sont en général acquis pour chaque ion sélectionné. L'utilisation de MS-Cluster va donc simplifier les données en fusionnant les spectres identiques en un même spectre consensus. Il en résulte qu'une grande majorité des perles sont composées de plusieurs spectres issus du même pic chromatographique, mais peuvent aussi être accompagnées de spectres ayant des temps de rétention différents. Dans ce dernier cas, il s'agit donc de plusieurs isomères ayant des caractéristiques spectrales identiques. Par ailleurs, des spectres MS² de molécules isobariques peuvent parfois être fusionnés dans un même spectre consensus, aboutissant à des spectres artefacts de nature « chimérique », c'est-à-dire affichant des caractéristiques spectrales hybrides entre les deux isobares (données non présentées).

2. Comparaison des profils métabolomiques LC-MS² de deux espèces du genre *Euphorbia*

Les réseaux moléculaires peuvent être utilisés pour comparer des données LC-MS² provenant d'échantillons différents. La figure 2 représente les réseaux moléculaires générés à partir de données issues de l'analyse de l'extrait d'*E. amygdaloides* subsp. *semiperfoliata* et de celui d'*Euphorbia*

dendroides (titre : *Euphorbia dendroides*). Les résultats montrent que : (i) peu de spectres MS² sont communs aux deux échantillons, (ii) aucun spectre correspondant à des spectres d'esters de jatrophane (de type JE-A et JE-B) n'est détecté dans l'extrait d'*E. dendroides*. Toutefois, l'observation de réseaux de similarité spectrale, regroupant des perles provenant des deux extraits, suggère l'existence d'analogues structuraux en commun non annotés, ou du moins de molécules ayant des caractéristiques spectrales similaires.

3. Vers une terminologie française adaptée

L'un des enjeux liés à l'émergence de cette méthode réside dans la création d'une terminologie adaptée à la langue française. Celle-ci doit prendre en compte la complexité inhérente à cette approche novatrice, tout en permettant l'emploi de termes destinés à être compris par un public non-spécialiste. La terminologie anglo-saxonne choisie par l'équipe de P. C. Dorrestein (UCSD) met en avant le fait qu'un spectre MS² est produit par un ion-molécule (*molecular networks*). Bien que ce choix semble judicieux pour communiquer auprès d'un public non-spécialiste, celui-ci est susceptible d'induire des contre-sens puisqu'il est basé sur une extrapolation qui peut être vue comme réductrice par le spécialiste. De plus, au-delà de l'apparence cartographique des réseaux moléculaires, cette approche est par nature constituée de plusieurs facettes conceptuelles (mathématique, spectrale et moléculaire). Ainsi, plutôt que d'utiliser une

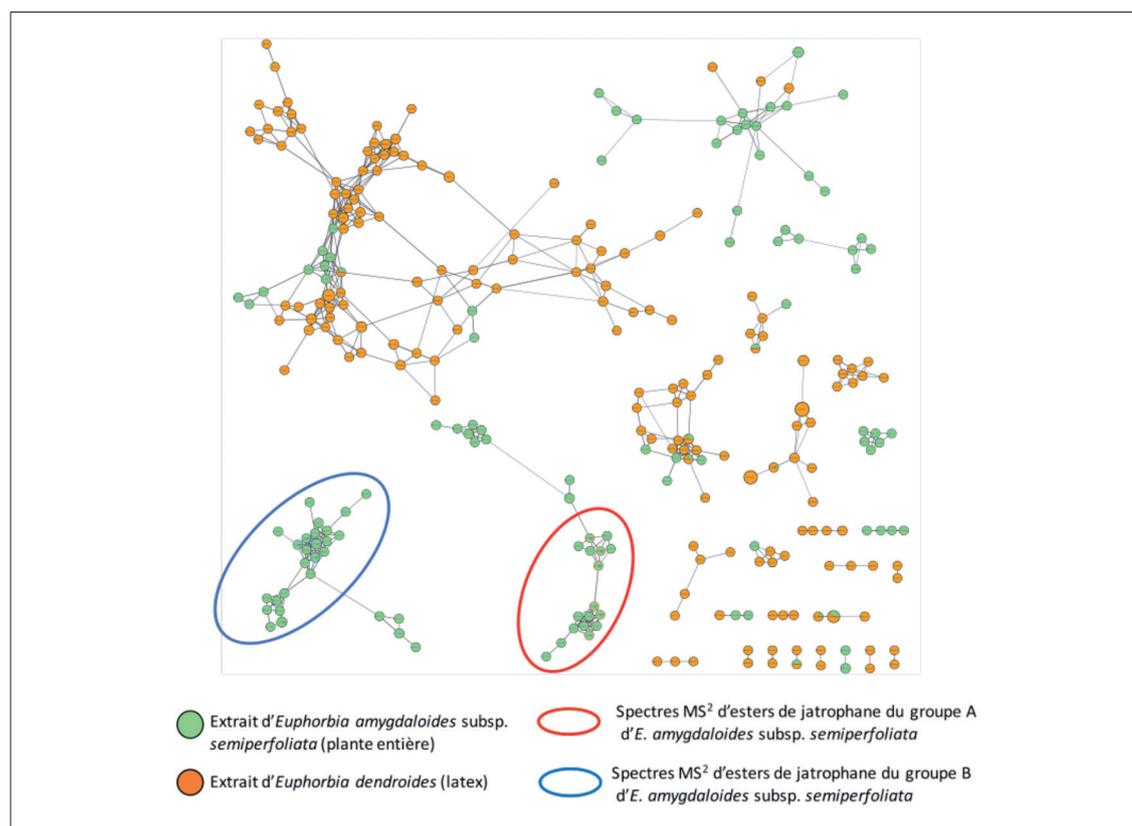
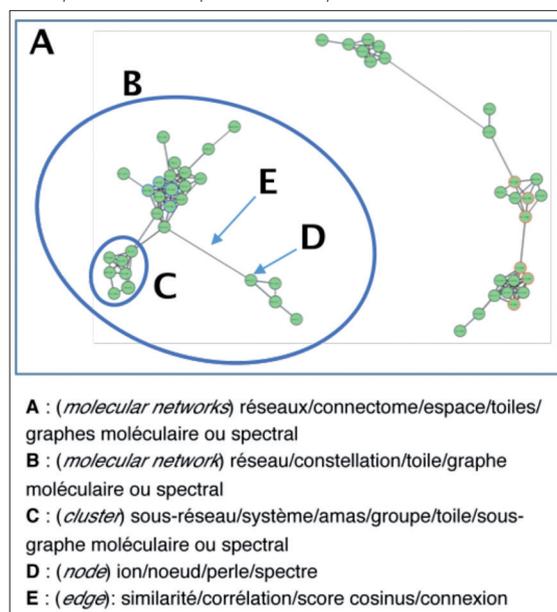


Figure 2
Comparaison des profils métabolomiques de deux espèces d'*Euphorbia* générés par analyse LC-MS², à l'aide de réseaux moléculaires

Figure 3
Variations terminologiques pouvant être associées aux réseaux moléculaires. Les termes anglo-saxons correspondants sont précisés entre parenthèses.



sémantique basée sur la traduction littérale des noms anglo-saxons inspirés de la terminologie informatique (*molecular networking*, *molecular network*, *node*, *edge*), nous suggérons l'utilisation d'une terminologie multidimensionnelle, où différents termes pourraient se substituer selon la facette à mettre en exergue. Par exemple, la facette mathématique pourrait être décrite par la terminologie issue de la géométrie dans l'espace (graphe, sommet, arête) et l'aspect

cartographique par analogie avec l'astronomie (espaces, constellations, système/amas). De plus, il faut remarquer l'existence du terme « connectome » en neurobiologie, qui pourrait être utilisé pour désigner l'ensemble des relations de similarité spectrale/moléculaire de données. Un sondage informel a été diffusé à l'Institut de Chimie des Substances Naturelles et via le Réseau Francophone de Métabolomique et Fluxomique, qui a permis de préciser les possibilités terminologiques s'offrant à la communauté (12). Les principales propositions sont présentées dans la Figure 3.

IV - Conclusion

Les réseaux moléculaires constituent une approche novatrice pour faciliter l'interprétation et l'annotation des résultats d'analyses LC-MS². Il faut garder à l'esprit que l'interprétation de ces réseaux moléculaires doit être effectuée avec précaution et discernement, du fait de la présence d'artefacts et des limites inhérentes à la spectrométrie de masse. Par ailleurs, l'émergence de cette approche complexe impose à la communauté de réfléchir à la terminologie française qui doit être y associée. Grâce à son accessibilité sur la plateforme internet GNPS, l'utilisation des réseaux moléculaires devrait se généraliser à court terme. Cette plateforme a été conçue pour évoluer au fil des contributions des utilisateurs, mais aussi pour intégrer l'ajout de prochains outils bio-informatiques complémentaires (8). GNPS constitue donc un lieu incontournable pour les communautés de chercheurs utilisant la spectrométrie de masse tandem.

REMERCIEMENTS

Nous souhaitons remercier Isabelle Schmitz-Afonso (CNRS-ICSN) pour la réalisation des analyses LC-MS². Par ailleurs, nous souhaitons remercier toutes les personnes qui ont participé à la réflexion terminologique, notamment ceux de l'ICSN et les membres Réseau Francophone de Métabolomique et Fluxomique (RFMF). Ce travail a bénéficié d'un soutien financier «Investissement d'Avenir» de l'Agence National de la Recherche (CEBA, ref. ANR-10-LABX-25-01).

RÉFÉRENCES

- (1) WOLFENDER JL *et al.*, Current approaches and challenges for the metabolite profiling of complex natural extracts, *Journal of Chromatography A*, 2015, 1382, 136–164, doi:10.1016/j.chroma.2014.10.091.
- (2) COBLE JB *et al.*, Comparative evaluation of preprocessing freeware on chromatography/mass spectrometry data for signature discovery, *Journal of Chromatography A*, 2014, 1358, 155–164, doi:10.1016/j.chroma.2014.06.100.
- (3) RYU SY, Bioinformatics tools to identify and quantify proteins using mass spectrometry data, *Advances in Protein Chemistry and Structural Biology*, 2014, 94, 1–17, doi:10.1016/B978-0-12-800168-4.00001-9
- (4) WATROUS J *et al.*, Mass spectral molecular networking of living microbial colonies, *Proceedings of National Academy of Sciences*, 2012, 109, E1743–E1752, doi:10.1073/pnas.1203689109.
- (5) Yang JY *et al.*, Molecular Networking as a Dereplication Strategy, *Journal of Natural Products*, 2013, 76, 1686–1699, doi:10.1021/np400413s.
- (6) BOUSLIMANI A *et al.*, Mass spectrometry of natural products: current, emerging and future technologies, *Natural Products Reports*, 2014, 31, 718, doi:10.1039/c4np00044g.
- (7) GARG N *et al.*, Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures, *International Journal of Mass Spectrometry*, 2015, 377, 719–727, doi:10.1016/j.ijms.2014.06.005.
- (8) NOTHIAS-SCAGLIA LF, communication personnelle de DOR-RESTEIN PC (UCSD), (octobre 2015).
- (9) NOTHIAS-SCAGLIA LF *et al.*, Jatrophone Diterpenes as Inhibitors of Chikungunya Virus Replication: Structure–Activity Relationship and Discovery of a Potent Lead, *Journal of Natural Products*, 2014, 77, 1505–1512, doi:10.1021/np500271u.
- (10) NOTHIAS-SCAGLIA LF *et al.*, Insights on profiling of phorbol, deoxyphorbol, ingenol and jatrophone diterpene esters, *Journal of Chromatography A*, 2015, 1422, 128–139, doi: 10.1016/j.chroma.2015.09.092.
- (11) SHANNON P *et al.*, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Research*, 2003, 13, 2498–2504, doi: 10.1101/gr.1239303
- (12) NOTHIAS-SCAGLIA LF *et al.*, Sondage au sujet de la terminologie que la communauté francophone pourrait utiliser sur la thématique "molecular networking MS/MS?", 2015. https://docs.google.com/forms/d/1_ApBEgpHulDTjtrS_7AiyUJxYx39o_w5LyCSTs3OxfM/prefill (accessed November 30, 2015).
- (13) NOTHIAS-SCAGLIA LF *et al.*, LC-MS²-Based dereplication of Euphorbia extracts with anti-Chikungunya virus activity, *Fitoterapia*, 2015, 105, 202–209, doi: 10.1016/j.fitote.2015.06.021.