



**HAL**  
open science

## Spatialized Metabolomic Annotation Combining MALDI Imaging and Molecular Networks

Marceau Levasseur, Edith Nicol, Nicolas Elie, Emeline Houël, Véronique Eparvier, David Touboul

► **To cite this version:**

Marceau Levasseur, Edith Nicol, Nicolas Elie, Emeline Houël, Véronique Eparvier, et al.. Spatialized Metabolomic Annotation Combining MALDI Imaging and Molecular Networks. *Analytical Chemistry*, 2023, 96 (1), pp.18-22. 10.1021/acs.analchem.3c03482 . hal-04426943

**HAL Id: hal-04426943**

**<https://hal.science/hal-04426943v1>**

Submitted on 10 Oct 2024

**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.

# Spatialized metabolomic annotation combining MALDI imaging and molecular network.

Marceau Levasseur,<sup>1</sup> Edith Nicol,<sup>2</sup> Nicolas Elie,<sup>1</sup> Emeline Houël,<sup>3</sup> Véronique Eparvier,<sup>1</sup> David Touboul<sup>1,2,\*</sup>.

<sup>1</sup> Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301, 91198, Gif-sur-Yvette, France; <sup>2</sup> Laboratoire de Chimie Moléculaire (LCM), CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France; <sup>3</sup> Sorbonne Université, CNRS, Laboratoire de Biodiversité et Biotechnologies Microbiennes, LBBM, Observatoire Océanologique, 66 650 Banyuls-sur-Mer, France.

**ABSTRACT:** MALDI mass spectrometry imaging has gained a major interest in the field of chemical imaging. This technique makes it possible to locate tens to hundreds of ionic signals on the sample surface without any a priori. One of the current challenges is still the limited ability to annotate signals in order to convert  $m/z$  values into probable chemical structures. At the same time, data obtained by LC-MS/MS have benefited from the development of numerous chemo-informatics tools, in particular molecular networks, for their efficient annotation. For the first time, we present here the combination of MALDI-FT-ICR imaging to molecular networks from MALDI-MS/MS data directly acquired on plant tissue sections. Annotation improvements are demonstrated paving the way of new annotation pipelines for MALDI imaging.

Mass spectrometry imaging (MSI) has expanded rapidly over the past 25 years, thanks to its ability to locate and quantify ionic signals on a variety of organic and/or inorganic surfaces, including biological tissue sections, heritage objects and inorganic materials.<sup>1,2,3</sup> This technology has benefited from numerous instrumental developments in terms of ionization sources and analyzers. Nevertheless, signal annotation, *i.e.*, the ability to transform  $m/z$  data in MS and MS/MS modes into putative chemical structures, remains a major challenge. From the earliest MALDI MSI articles on peptide/protein distribution, signal attribution was problematic due to the limited correlation between MALDI MSI data and attributions after extraction followed by classical proteomic approaches.<sup>4,5</sup> For the analysis of small endogenous or exogenous molecules, high-resolution mass spectrometry (TOF,<sup>6</sup> Orbitrap<sup>7</sup> or FT-ICR<sup>8</sup> analyzer) provides access to exact mass measurements, enabling the generation of a small number of coherent raw formulae, typically for values below  $m/z$  500. In the field of lipidomics, which is one of the major applications of MALDI MSI, the combination of high-resolution MS and MS/MS data with the search of very rich and well-curated databases enables to propose relatively reliable annotations, even for signals of low intensity.<sup>9,10</sup> In the case of the analysis of natural compounds on biological surfaces (leaves, fruit, microorganism cultures . . .), this workflow proves more complex due to the multiplicity of databases to be queried and the absence of reference MS and/or MS/MS spectra for many compounds.<sup>11,12</sup>

In the field of metabolomic based on LC-HRMS/MS analysis, an elegant method for propagating annotation from MS and MS/MS data is to represent the data as a molecular network. This method is based on the pairwise comparison of experimental MS/MS spectra and the calculation of a cosine score reflecting the spectral similarity. This concept was first proposed by M.L. Gross<sup>13</sup> and implemented via the GNPS web platform

by P. Dorrestein's team 10 years later.<sup>14</sup> Another tool, called MetGem,<sup>15</sup> was also developed by our team, and enables molecular networks to be generated under a t-SNE-type visualization and without online data deposit. In both cases, data are annotated by comparing experimental data with an aggregation of ESI-MS/MS spectra pooled in public databases. One of the latest advances in this analytical methodology is the ion identity, which enables different adducts of the same ionic species to be connected in a molecular network.<sup>16</sup> Although this methodology was first applied to micro-organisms, it has become a standard for the analysis of specialized plant metabolites.<sup>17</sup> Very recently, Schmid *et al.* proposed the integration of MS imaging and LC-Ion Mobility-MS datasets for spatial metabolomics analyses with increased annotation confidence using MZmine3 software.<sup>18</sup>

For the first time, we propose a two-step analysis methodology with the acquisition of ion mapping by MSI followed by the acquisition of MALDI-MS/MS spectra directly from biological tissues to generate molecular networks. This methodology makes it possible to annotate the specialized metabolites detected and understand their distribution in biological tissue.

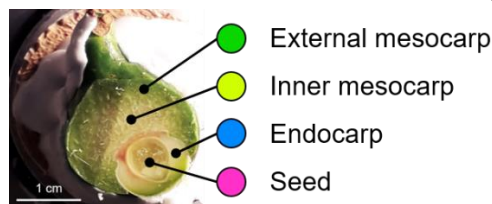


Figure 1. *Sextonia rubra* fruit anatomy. Different tissues analyzed and corresponding color code.

The chosen sample is a fruit of *Sextonia rubra* (Figure 1). *Sextonia rubra* (Mez.) van der Werff (Lauraceae) is a tropical

tree endemic to the Guiana Shield and the Brazilian Amazon.<sup>17</sup> Its heartwood shows a natural durability leading to its use as a building material. Although the description of this species dates back to 1997, the chemical composition of its fruits has never been studied which makes it an ideal model for demonstrating proof of concept since fruits contain numerous families of specialized metabolites.<sup>20-21</sup> Fruits commonly enclose specialized metabolites involved in the attraction of pollinators or the repulsion of pathogens/predators, thus involved in their dispersal and defense.

The general workflow to generate MALDI-MS imaging data and MS/MS data is introduced in Figure 2. First, HRMS data of each anatomical part of the fruit, *i.e.*, external and internal mesocarps, endocarp, and seed, have been recorded. These spectra have been used to generate a chemical formula for each ion in the positive ion mode. Then MALDI-MS imaging experiments have been performed using the continuous accumulation of selected ions (CASI) mode.<sup>18</sup> It has been demonstrated that this method permits the selective ejection of the most abundant ion species prior to ion accumulation, leading to a significant improvement of the sensitivity by adapting the charge capacity of the ion accumulation cell while enhancing dynamic range. Finally, for each detected ions, MS/MS spectra were manually acquired with an optimization of the collision energy in order to generate data of the best quality, *i.e.* fragments and neutral losses well distributed over the mass range without privileging low or high  $m/z$  fragment. MS/MS acquisitions were performed at similar locations (area < 100  $\mu\text{m}^2$ ), according to each tissue analyzed, in order to limit positional deviation.

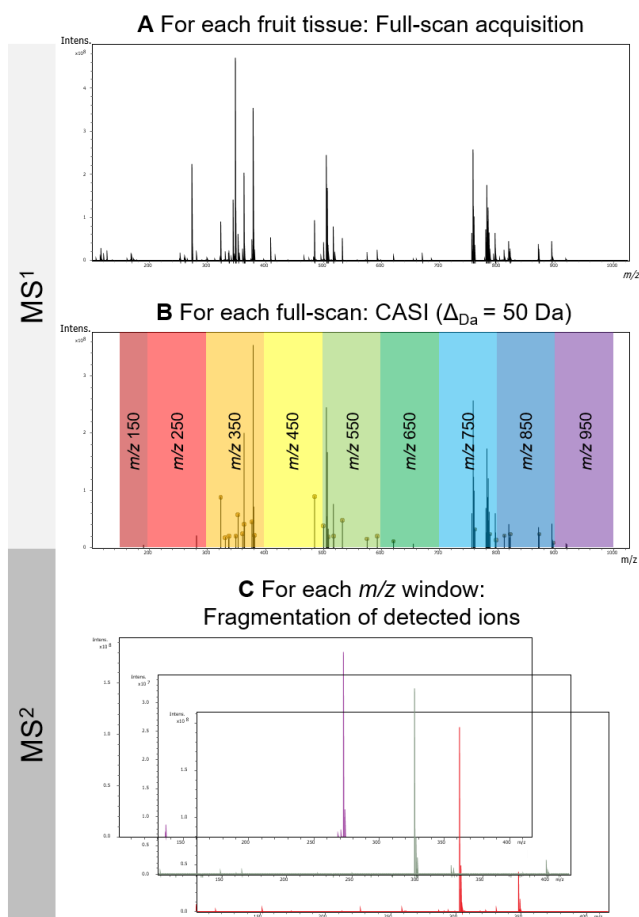


Figure 2. Workflow of the data acquisition.

Figure 3 shows MALDI-FT-ICR ion images obtained at a resolution of 20  $\mu\text{m}$  in the positive ion modes on the sections of a *Sextonia rubra* fruit using 2,5-dihydroxybenzoic acid (DHB) as matrix. The MS image at  $m/z$  758.5677 (Figure 3e) allows to easily discriminate the different types of tissue to correlate the optical image (Figure 3a) and ion images.

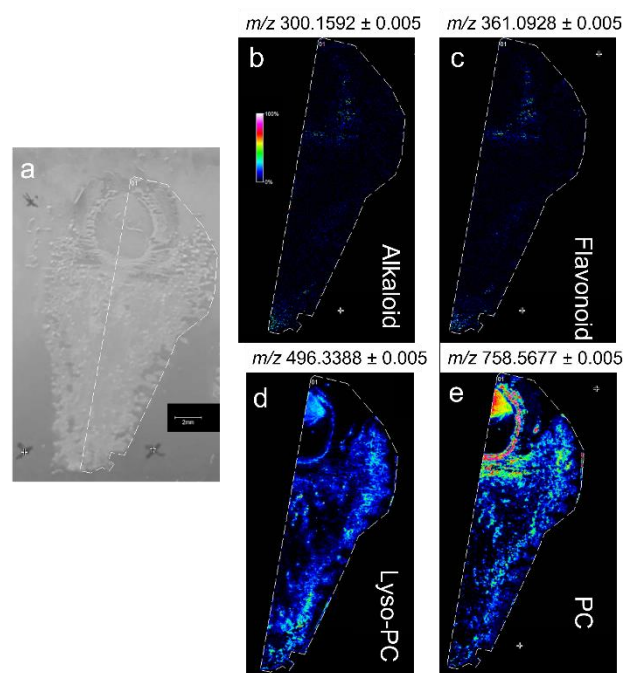


Figure 3. MALDI-FT-ICR images in positive ion mode with a spatial resolution of 20  $\mu\text{m}$ . a- optical image of the fruit section. b,c,d,e- MS images at  $m/z$  300.1592,  $m/z$  361.0928,  $m/z$  496.3388 and  $m/z$  758.5677, respectively.

It must be noted that odd and even ions have been detected on the different parts of the fruit section indicating the presence of nitrogen-containing compounds, such as pyrroles, piperidines, diketopiperazines, isoquinolines or phospholipids. As many classes of natural products can correlate with a single molecular formula, it is thus required to perform MS/MS experiments to get more structural information for each detected species. The collision energy was optimized for each precursor ion to obtain at least two fragments. 51 MS/MS spectra fulfilled this criterion. Then a molecular network from MALDI-MS/MS spectra have been generated and depicted using MetGem software (Figure 4) to annotate detected specialized metabolites. A crucial step resides in the conversion of the MALDI spectra to create a .mgf file containing the intensity of the parent ions and their fragments, then a .csv file containing the intensity of each parent ion, their  $m/z$  value, the collision energy used and in which tissue the ion was detected. The code is available at [https://github.com/m-levasseur/MALDI\\_FT\\_ICR\\_mgf](https://github.com/m-levasseur/MALDI_FT_ICR_mgf). When applying cosine-score threshold at 0.4, the molecular network clearly exhibits three distinct clusters, probably indicating three different molecular classes of detected endogenous natural products.

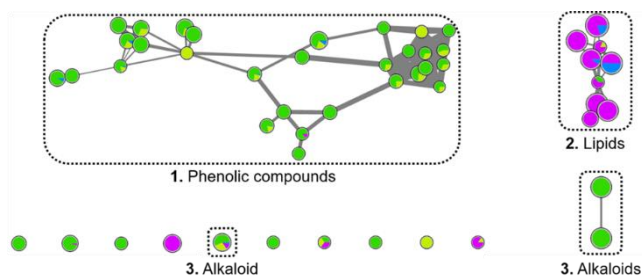


Figure 4. Molecular networks calculated from the MALDI-MS/MS spectra directly acquired on the fruit section. Pie charts are corresponding to the relative ion intensities in the different fruit parts (see Figure 1 for corresponding colors).

In order to get a first annotation of the dataset, comparison of experimental data with public databases (GNPS, MS-Dial, MassBank, LipidBlast) has been performed by searching standards and analogs ( $\pm 100$  Da). The largest cluster is gathering ions highly abundant in the external mesocarp of the fruits. The database queries indicated that all the ions belong to phenolic compounds. In particular, the nodes outlined in blue in Figure 5 are corresponding to ions with structures related to flavonoids. Flavonoids are specialized metabolites of vascular plants, exhibiting the same basic structure formed by two aromatic rings linked by three carbons: C6-C3-C6, a chain usually closed by a hexagonal or pentagonal oxygenated heterocycle. They are pigments involved in the coloring of petals, mesocarps and pericarps, showing strong antioxidant activities and photoprotection effects of the fruits. Thus, the distribution and the annotation are fully relevant with the known chemistry of a fruit. The same cluster is also related to other compounds derived from 2,5-DHB. In fact, 2,5-DHB was used as a matrix for MALDI imaging and the structure of these compounds (protonated, sodium- and potassium-adducts, clusters, dehydrated 2,5-DHB and derivatives) are structurally related to flavonoids due to hydroxybenzoic ring. Moreover, flavonoids have already been reported to form ion clusters (protonated, sodiated or potassiated) with the dehydrated DHB moiety.<sup>23</sup> Thus, molecular networks allow to easily eliminate all the matrix background by a fast annotation of all these species and identify interfering signals that are difficult to identify using conventional techniques. This methodology makes it easier to select features that need further investigation, e.g. signals considered as artifacts can be easily removed from a dataset when searching for biomarkers.

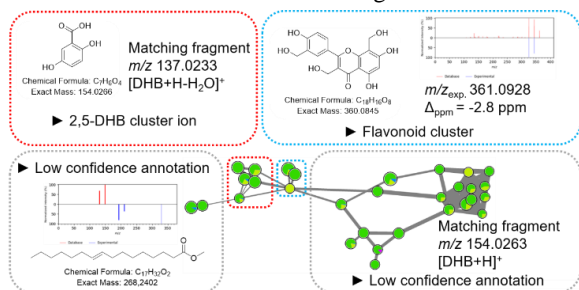


Figure 5. Annotation details of the largest cluster from Figure 1.

The second-largest cluster (Figure 6) related to molecules highly abundant in the seed was annotated as lipids by comparing experimental data with LipidBlast database, which can be interrogated via MetGem. Glycerophosphatidylcholines (PCs)

and lysoglycerophosphatidylcholines (LPCs) detected as sodium and potassium adducts were clearly annotated. High content of PCs has been reported for fruit seeds indicating that our observations are in good agreement with the literature.<sup>24</sup>

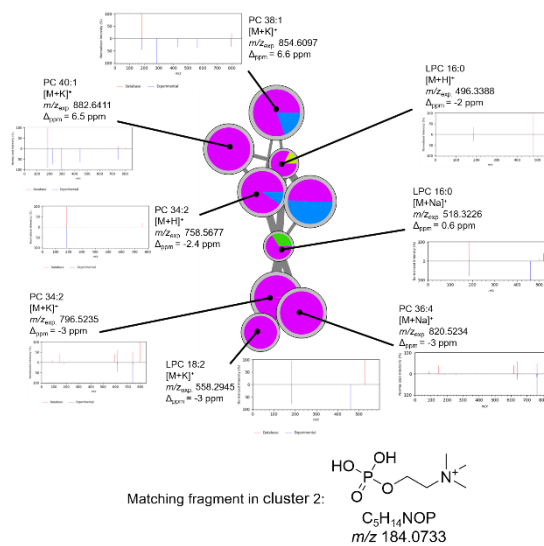


Figure 6. Annotation details of the lipid cluster from Figure 1.

The smallest cluster containing two nodes, with even  $m/z$  ratio corresponding to a particular class of molecules bearing an odd number of nitrogens, has been investigated (Figure 7). The two connected compounds share a neutral loss ( $-45$  Da) and a common fragment at  $m/z$  253 (Figure 7A). Dereplication process suggests that these compounds could be annotated as alkaloids of the aporphine subfamily, more specifically boldine derivatives, according to database search (Figure 7B, C). However, no aporphine alkaloid can possess the calculated chemical formula determined from HRMS data. Furthermore, the neutral loss of 45 Da corresponds to a secondary amine  $C_2H_5NH_2$ , that is not possible from boldine protonated species. Nevertheless, the phytochemical diversity of Lauraceae is well documented in the literature<sup>25</sup> in order to refine the annotation. This plant family is able to synthesize alkaloids of the benzyloquinoline subfamily. This subfamily meets structural criteria consistent with the determined chemical formula (Figure 7D). It must be noted that this molecule has already been isolated from tree leaves of a closely related taxonomic species: *Ocotea paranapiacabensis* Coe-Teix (Lauraceae).<sup>26</sup>



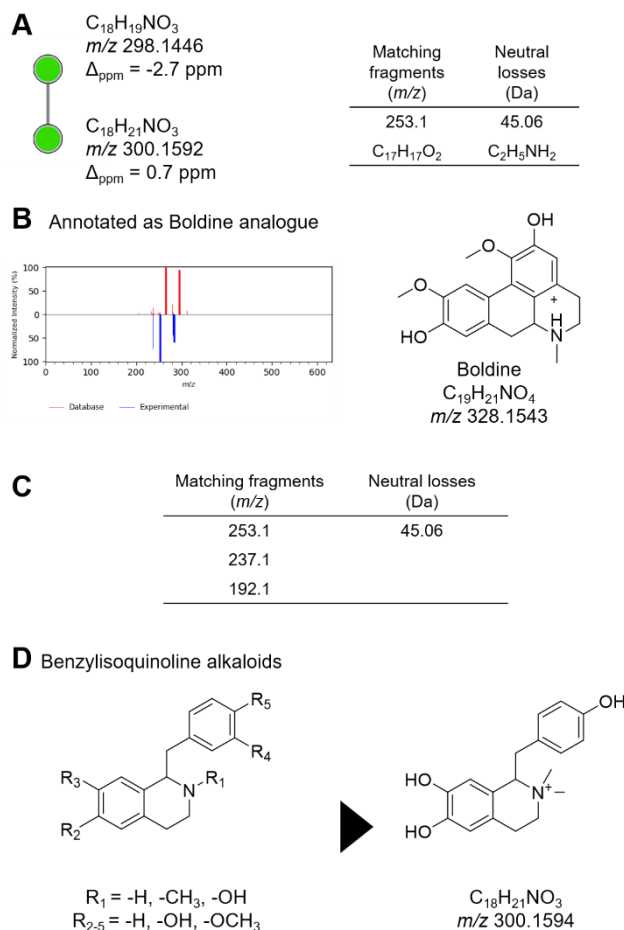


Figure 7. Annotation details of the alkaloids cluster from the Figure 1. A- nodes annotated as alkaloids with diagnostic fragment and neutral loss, B- mirrored spectra between the ion at  $m/z$  300.1592 and boldine, a standard in queried databases and its structure, C- diagnostic fragments and neutral loss of ion at  $m/z$  300.1592 involved in boldine annotation, D- annotation correction.

In conclusion, we demonstrated that MALDI-MS/MS spectra acquired in situ on a plant tissue surface can be efficiently used for the data annotation by the careful interpretation of the resulting molecular network. This paves the way of a new annotation pipeline in MALDI imaging to better correlate images from the ion intensity at a particular  $m/z$  value to the chemical information. Although the resolution conferred by MALDI-FT-ICR analysis is important, future work will need to be carried out using a MALDI-Orbitrap to automate the HR-IMS and MS/MS acquisition process.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

The raw data, ion cartography, script and molecular network are available here: <https://zenodo.org/deposit/8186791#>

Plant material; sample preparation; MS analyses; Data pre-processing; Molecular network calculation; Panel of MS images (PDF)

## AUTHOR INFORMATION

### Corresponding Author

\* David Touboul, [orcid.org/0000-0003-2751-774X](https://orcid.org/0000-0003-2751-774X) ; Email: [david.touboul@cnrs.fr](mailto:david.touboul@cnrs.fr).

## Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript.

## ACKNOWLEDGMENT

We thank the CIRAD team, Saint-Omer Cazal, Michel Baisie, Soepe Koesse, Jean-Martin Odan, Nadine Amusant, the CNRS engineer Yannick Estevez, and the INRAE tree climber Jocelyn Cazal for the pruning and felling of the studied tree. This work has received financial support from the CNRS through the 80|Prime program. Financial support from the IR INFRANALYTICS FR2054 for conducting the research is gratefully acknowledged.

## REFERENCES

- Caprioli, R. M.; Farmer, T. B.; Gile, J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **1997**, *69*(23), 4751-4760.
- Guo, X.; Wang, X.; Tian, C.; Dai, J.; Zhao, Z.; Duan, Y. Development of mass spectrometry imaging techniques and its latest applications. *Talanta* **2023**, *264*, 124721.
- Tuck, M.; Blanc, L.; Touti, R.; Patterson, N.H.; Van Nuffel, S.; Villette, S.; Taveau, J. C.; Römpp, A.; Brunelle, A.; Lecomte, S.; Desbenoit, N. Multimodal Imaging Based on Vibrational Spectroscopies and Mass Spectrometry Imaging Applied to Biological Tissue: A Multiscale and Multiomics Review. *Anal. Chem.* **2021**, *93*(1), 445-477.
- Delcourt, V.; Franck, J.; Quanico, J.; Gimeno, J.P.; Wisztorski, M.; Raffo-Romero, A.; Kobeissy, F.; Roucou, X.; Salzert, M.; Fournier, I. Spatially-Resolved Top-down Proteomics Bridged to MALDI MS Imaging Reveals the Molecular Physiome of Brain Regions. *Mol. Cell. Proteomics* **2018**, *17*(2), 357-372.
- Ryan, D. J.; Spraggins, J. M.; Caprioli, R. M. Protein identification strategies in MALDI imaging mass spectrometry: a brief review. *Curr. Opin. Chem. Biol.* **2019**, *48*, 64-72.
- Wu, J.; McAllister, H. Exact mass measurement on an electrospray ionization time-of-flight mass spectrometer: error distribution and selective averaging. *J. Mass Spectrom.* **2003**, *38*(10), 1043-1053.
- Nolting, D.; Malek, R.; Makarov, A. Ion traps in modern mass spectrometry. *Mass Spectrom. Rev.* **2019**, *38*(2), 150-168.
- Gosset-Erard, C.; Aubriet, F.; Leize-Wagner, E.; François, Y. N.; Chaïmbault, P. Hyphenation of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) with separation methods: The art of compromises and the possible - A review. *Talanta* **2023**, *257*, 124324.
- Fu, T.; Oetjen, J.; Chapelle, M.; Verdu, A.; Szesny, M.; Chaumot, A.; Degli-Esposti, D.; Geffard, O.; Clément, Y.; Salvador, A.; Ayciriex, S. In situ isobaric lipid mapping by MALDI-ion mobility separation-mass spectrometry imaging. *J. Mass Spectrom.* **2020**, *55*(9), e4531.
- Cerruti, C. D.; Touboul, D.; Guérineau, V.; Petit, V.W.; Laprévote, O.; Brunelle, A. MALDI imaging mass spectrometry of lipids by adding lithium salts to the matrix solution. *Anal. Bioanal. Chem.* **2011**, *401*(1), 75-87.
- Wolfender, J. L.; Litaudon, M.; Touboul, D.; Queiroz, E.F. Innovative omics-based approaches for prioritisation and targeted isolation of natural products - new strategies for drug discovery. *Nat. Prod. Rep.* **2019**, *36*(6), 855-868.
- Beniddir, M. A.; Kang, K. B.; Genta-Jouve, G.; Huber, F.; Rogers, S.; van der Hooft, J. J. J. Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches. *Nat. Prod. Rep.* **2021**, *38*(11), 1967-1993.
- Wan, K. X.; Vidavsky, I.; Gross, M. L. Comparing Similar Spectra: From Similarity Index to Spectral Contrast Angle. *J. Am. Soc. Mass Spectrom.* **2002**, *13*(1), 85-88.

14. Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*(8), 828-837.
15. Olivon, F.; Elie, N.; Grelier, G.; Roussi, F.; Litaudon, M.; Touboul, D. MetGem Software for the Generation of Molecular Networks Based on the t-SNE Algorithm. *Anal. Chem.* **2018**, *90*(23), 13900-13908.
16. Schmid, R. *et al.* Ion Identity Molecular Networking for Mass Spectrometry-Based Metabolomics in the GNPS environment. *Nat. Commun.* **2021**, *12*(1), 3832-3844.
17. Li, D.; Gaquerel, E. Next-Generation Mass Spectrometry Metabolomics Revives the Functional Analysis of Plant Metabolic Diversity. *Annu. Rev. Plant Biol.* **2021**, *72*(1), 867-891.
18. Schmid, R. *et al.* Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* **2023**, *41*(4), 447-449.
19. van der Werff H. Sextonia, a New Genus of Lauraceae from South America. *Novon.* **1997**, *7*, 436-439.
20. Crecelius, A. C.; Hölscher, D.; Hoffmann, T.; Schneider, B.; Fischer, T. C.; Hanke, M. V.; Flachowsky, H.; Schwab, W.; Schubert, U. S. Spatial and temporal localization of flavonoid metabolites in strawberry fruit (*Fragaria* × *ananassa*). *J. Agric. Food Chem.* **2017**, *65*(17), 3559-3568.
21. Fujita, R.; Hayasaka, T.; Jin, S.; Hui, S. P.; Hoshino, Y. Comparison of anthocyanin distribution in berries of Haskap (*Lonicera caerulea* subsp. *edulis* (Turcz. Ex. Herder) Hultén, Miyamaquisukagura (*Lonicera gracilipes* Miq.), and their interspecific hybrid using imaging mass spectrometry. *Plant Science.* **2020**, *300*, 110633-110642.
22. Prentice, B. M.; Ryan, D. J.; Grove, K. J.; Cornett, D. S.; Caprioli, R. M.; Spraggins, J.M. Dynamic Range Expansion by Gas-Phase Ion Fractionation and Enrichment for Imaging Mass Spectrometry. *Anal. Chem.* **2020**, *92*(19), 13092-13100.
23. Madeira, P. J. A.; Florêncio, M.H. Flavonoid-Matrix Cluster Ions in MALDI Mass Spectrometry. *J. Mass Spectrom.* **2009**, *44*(7), 1105-1113.
24. Feenstra, A. D.; Hansen, R. L.; Lee Y. J. Multi-Matrix, Dual Polarity, Tandem Mass Spectrometry Imaging Strategy Applied to a Germinated Maize Seed: Toward Mass Spectrometry Imaging of an Untargeted Metabolome. *Analyst*, **2015**, *140*(21), 7293-7304.
25. Silva Teles, M. M. R.; Vieira Pinheiro, A. A.; Da Silva Dias, C.; Fachine Tavares, J.; Barbosa Filho, J. M.; Leitão Da Cunha, E. V. Alkaloids of the Lauraceae. In *The Alkaloids: Chemistry and Biology*; Elsevier, 2019; Vol. 82, pp 147-304.
26. Freitas, L. D.; Valli, M.; Dametto, A. C.; Pennacchi, P. C.; Andricopulo, A. D.; Maria-Engler, S. S.; Bolzani, V. S. Advanced Glycation End Product Inhibition by Alkaloids from *Ocotea paranapiacabensis* for the Prevention of Skin Aging. *J. Nat. Prod.* **2020**, *83*(3), 649-656.