IFSCC 2025 full paper (IFSCC2025-977)

"Metabolomic approaches in the development of cosmetic products : discriminative sourcing, sustainable and efficient ingredients."

Souhila Messaili[1], Salma Massis[1], Manon Massias[1], Doha Haggouch[1], Eldra Delannay[1,*]

[1] 75008, Sisley Paris, Paris, France

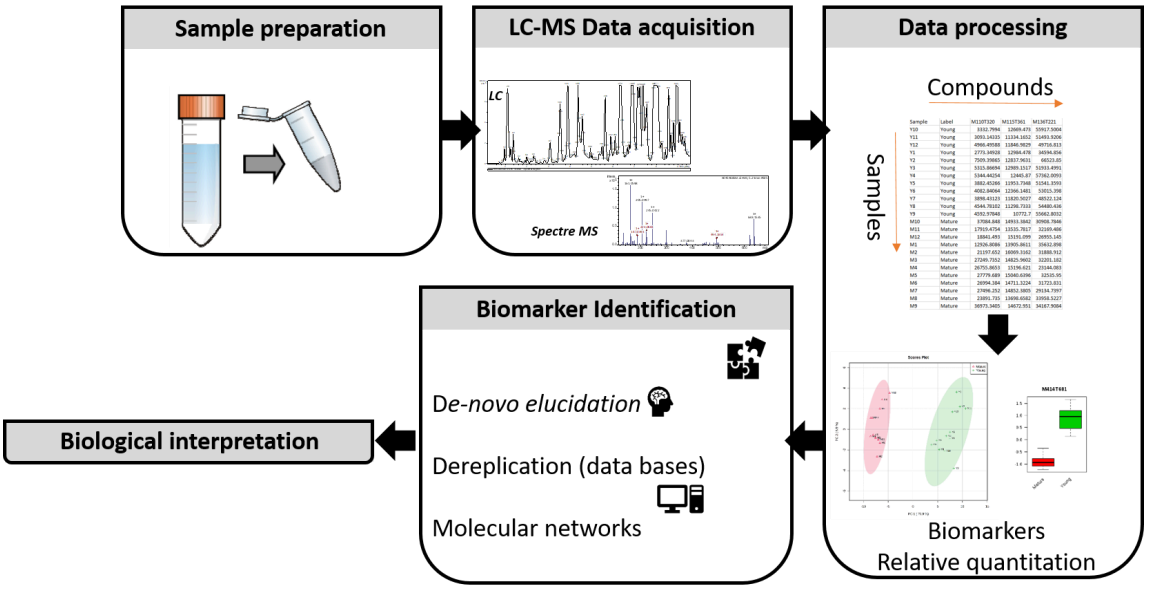
1. Introduction

The development of a cosmetic ingredient relies on a rigorous selection of raw materials and extraction processes to ensure its efficacy and safety. In this context, metabolomics approaches [1-2] play a key role by enabling an in-depth analysis of the chemical composition of plant, microbial, or other extracts. Thanks to these advanced analytical techniques, it becomes possible to discriminate biogeographical origins, evaluate the impact of cultivation methods (conventional agriculture, organic, hydroponic, etc.), and compare the effect of different extraction processes such as the use of alternative, green, and more sustainable solvents, ultrasounds, or enzymatic techniques. By offering a detailed view of the metabolic profile of the different extracts developed, metabolomics allows for the optimization of their quality, stability, and bioactive potential, thus contributing to innovation in the cosmetic field.

The species on which we focused for the development of a sustainable and effective cosmetic ingredient is from the Magnoliopsidae class and is based on several complementary aspects: - a phytochemical composition, a priori flavonoids and anthocyanins for antioxidant effects[3], and mucilages and polysaccharides for a moisturizing and soothing aspect - a certain role in ecology which confers a sustainability aspect, since it is a native European plant with fairly good resilience to environmental stresses (ideal for organic farming or the evaluation of indoor crops) - the possibility of using 'ecological' extraction processes (ultrasounds, bio-based solvents, etc.) to pursue the sustainable aspect and obtain active ingredients without a negative impact on the environment.

The term metabolomics emerged alongside other "omics" technologies in biological research (genomics, transcriptomics, proteomics, and metabolomics[4]), providing an overview of an organism's biochemical state that can be used to monitor significant metabolite variations. Indeed, as metabolites are the end products of cellular regulatory processes, their levels can be considered the ultimate response of biological systems to genetic or environmental changes. Unlike metabolite profiling, which focuses on compounds from a single metabolic pathway, this approach allows for the identification and quantification of all metabolites present in a biological sample. Two types of analysis can then be considered: global or untargeted analysis, which focuses on all compounds in the sample without prior selection; and targeted analysis, which focuses on a limited number of known metabolites, belonging to one or more chemical families, and predefined before analysis to determine their precise concentrations. These two strategies (metabolite profiling and metabolomics[5]), which can be complementary, are used either for a complete study of the chemical composition of a

plant extract of interest or in combination with advanced multivariate data analysis in the search for biomarkers responsible for the discrimination between plant extracts. When these statistical analyses are performed on a sufficient number of biological replicates, they make it possible to discriminate and classify samples into groups and to follow the modifications of the metabolome composition that are linked to a given physiological state, to the influence of a stress or a stimulus, to the genetic modification, to the interaction with other organisms, etc. In our case study, which focuses on the development of a cosmetic ingredient according to its production method, it allows for the evaluation of biochemical changes related to the plant's growth environment and its stage of maturity (controlled environment, open field, and different geographical territories), biological efficacy, toxicity, and the effectiveness of extraction treatments (nature of solvents, ratio, techniques such as reflux, maceration, ultrasound, enzymes...) with metabolic interactions and structure-activity relationships. By combining advanced analytical technologies (such as mass spectrometry[6]) with computational tools, metabolomics offers a comprehensive view of metabolism and helps to analyze bioactive compounds and their effects on the skin in depth. The diagram below presents the different steps that generally comprise a metabolomic analysis by LC-HRMS[7] and that we implemented for the study of extracts from the selected Magnoliopsida species.



2. Materials and Methods

Following the collection of plant samples, which must be rigorously performed as it partly determines the relevance of the study and the results obtained, the extraction process development, the first step of the analysis , is sample preparation for metabolomic studies. Its objective is to extract as completely as possible the metabolites present in a sample. This preparation is dependent on the type of matrix studied. The second step is the acquisition of data provided by the separation of analytes/metabolites in the chromatographic system and the acquisition of mass spectra by the spectrometer. At this stage, different approaches are studied to improve the quality and quantity of information acquired, notably through the use of high-resolution mass spectrometers. After acquisition, the data are processed. This step is crucial to extract relevant and usable information from the acquired data. The statistical analysis step allows for the identification of discriminant variables between the groups of individuals studied. This step can highlight potential biomarkers, and the set of these biomarkers constitutes the metabolic signature related to the studied condition/perturbation. The final step consists of the biological interpretation of the obtained metabolic signature. This step is currently the most complex.

2.1 Plant raw material

The selected *Magnoliopsidae* species is cultivated using two different methods: a conventional method in open fields (distinct European geographical origins, to favor short supply chains), under sustainable agriculture practices, and a vertical farming method (also within short supply chains). For optimal characterization of the cosmetic interest, different maturity stages (vegetative and flowering) were sampled. The harvested biomass was dried using conventional methods (oven <40°C or racks in the sun) and ground into different particle sizes. For the purposes of this study, with flowering stages and 0.8-1mm size particles, the following nomenclature was adopted: Vb-pf1 (Variety b in country 1), V-Pb1 (Variety a in country 2), and V-T1 (Variety a, indoor cultivation, in country 1).

2.2 Extraction method

Various extraction techniques were explored, including maceration, microwave-assisted extraction, ultrasound-assisted extraction, and fermentation. For this specific study, we selected an acidic aqueous extraction (2% citric acid w/w) by maceration at 40°C, with continuous stirring at 600 rpm for 6 hours and 30 minutes. The plant-to-solvent ratio used was 1:25.

2.3 UHPLC-UV-ESI-QTOF-MS/MS analysis

MS and MS/MS experiments were conducted on an Agilent G6545B UHR-Q-TOF mass spectrometer (Agilent, Les Ulis, France) using data-dependent acquisition (DDA) mode with an Agilent Jet Stream (AJS) ionization source operating in positive ion mode. Nitrogen was employed as the drying gas at a flow rate of 10 L/min and 325°C and as the nebulizing gas at a pressure of 20 psi (1.38 bar). Mass spectra were recorded over the m/z range of 100-1700 at a rate of 2 spectra/second and over the m/z range of 100-1700 at a rate of 5 spectra/s respectively for MS and MS/MS. The capillary voltage was set to 4 kV. Three precursor ions were selected per cycle and fragmented at collision energy of 15 eV. All MS data were processed using MassHunter 10.0 software (Agilent). Molecular formulas were generated using elemental composition (C, H, O) up to infinite, Na ≤ 1, and N ≤ 5, S ≤ 5, with a mass accuracy of ≤ 5 ppm.
Chromatographic analysis was conducted using an Agilent 1290 Infinity II system equipped with an auto sampler, a quaternary pump, a thermostat column compartment, and a DAD detector. The column employed was an UHPLC Eclipse Zorbax column (150 x 2.1 mm; 1.8 µm) (Agilent, Les Ulis, France). The mobile phase consisted of water (A) and methanol (B), both acidified with 0.1% and 0.08% formic acid, respectively, at a flow rate of 400 µL/min. The mobile phase gradient was programmed as follows: 0-1 min 10% (B), 1-8 min 80% (B), 8-9 min 80% (B), 9-11 min 100% (B), 11-15 min 100% (B), 15-15.1 min 10% (B), 15.1-18 min 10% (B). Compound separation was performed at 40°C. Samples were diluted 100 times in water, and 2 µL were injected into the system.

2.4 Statistical analysis

A total of 700 features were generated from UHPLC-HRMS analyses by Profinder. Each feature is identified by its most abundant mass-to-charge ratio (*m/z*) in its mass spectrum (including parent ions or fragment ions or adducts), along with its retention time in minutes. This large number of variables could lead to instability in the statistical model, resulting in irrelevant statistical outcomes. To address this issue, statistical tests were combined to obtain a smaller dataset from the initial matrix. To this end, the starting matrix was entered into the online software Mass Professional Profiler; the variables were centered and reduced. Partial Least Squares Discriminant Analysis (PLS-DA) was applied to obtain the Variable Important in Projection (VIP) scores of the variables, allowing for the selection of features and the elimination of those that were not discriminant between groups. Only 36 variables (number of variables = 3 x number of individuals) with the highest VIP scores according to component 1 were chosen to create a sub-matrix containing the most relevant variables. The dataset of 36 variables and 12 individuals was thus extracted from the initial matrix and used for statistical analysis to conduct the comparative study and determine the biomarkers discriminating the

aerial parts grown in vertical farms and those in fields, as well as to discriminate the different origins (countries 1 and 2). The reduced matrix was entered into Mass Professional Profiler, and the data were centered and reduced to normalize the dataset. Univariate (Volcano Plot) and unsupervised multivariate analyses, including Principal Component Analysis (PCA) and Hierarchical Clustering Analysis (HCA), as well as supervised PLS-DA, were used to identify the differences between the three groups of plants.

2.5 Compounds Characterization

To identify biomarkers of the studied aerial flowering parts, the mass tolerance was set at 4 ppm for MS and 7 ppm for MS/MS or small m/z values. The elements C, H, O, N, P, and S were selected for the elemental composition calculation, and only coherent chemical formulas were taken into account. Furthermore, chromatographic characteristics and HRMS/MS data, as well as bibliographic information, were used for identification. Possible molecular formulas for each selected peak were searched in public databases, namely KNapSAcK, Dictionary of Natural Products, HMBD, Super Natural II, and PubChem to characterize known natural products. In addition, some compounds were unambiguously identified by comparing their retention time and spectral data with those of reference standards. Finally, knowledge of the molecular family of biomarkers through molecular networking allowed for the selection and confirmation of chemical formulas when several were coherent.

2.6 In-tubo biochemical activities/assays

The first biochemical activities were evaluated in-tubo (acellular assay) using DPPH and Tyrosinase assays. This was determined by measuring the absorbance of the reagent in the presence of different samples (solvent for blank, Trolox and kojic acid for positive controls, total extract or fractions). The assessment of activity was conducted in 96-well plates, with each sample deposited. Water was used to dilute the samples and as the negative control. All assays were performed in triplicate. All assays were performed on a BMG Labtech ClariostarPlus microplate reader using Clariostar Mars software to read and process the data. Below is a table 1 summarizing the biochemical tests that were performed, along with their objectives and principles.

Table 1. Cosmetic activity screening of Magnoliaceae aerial part extracts : a summary of chemical and enzymatic assays.

Performed assays	Objective / Principle	Reference
DPPH radical scavenging activity assay	The antioxidants contained in an extract reduce DPPH, a blue-colored radical cation, and the color of the reactive medium turns pink depending on the antioxidant concentration.	Lee *et al.*[8]
Tyrosinase assay	Quantification of dopachrome, a colored product (dark orange) formed during the reaction catalyzed by tyrosinase, This production of dopachrome can be measured by visible spectrophotometry at 490 nm.	Lim *et al.*[9]
COX-2	To measure the activity of cyclooxygenase (COX), a key enzyme in the production of prostanoids, biological mediators involved in inflammation and pain. The presence of prostaglandin, which is fluorescent, and therefore the activity of COX, is quantified by fluorimetry at 535/590 nm or 535/587 nm. A decrease in fluorescence indicates inhibition.	Kit assay[10]
Hyaluronidase assay	The sample is brought into contact with the hyaluronidase enzyme, then hyaluronic acid is added and potentially digested if the enzyme is not inhibited. The addition of bovine serum albumin (BSA) precipitates the remaining hyaluronic acid, and if the optical density (OD) measured at 400 and 600 nm is higher, it indicates stronger inhibition.	Bralley *et al.*[11]

Sugar and a total polyphenols dosages	Total sugars and total phenols were quantified using colorimetric methods against glucose and gallic acid standards, respectively. Sugars formed a yellow complex with phenol and sulfuric acid, measured at 490 nm. Polyphenols formed a blue complex with Folin reagent via a redox reaction, measured at 735 nm.	Dubois et *al*.[12]

2.7 Molecular Network design

The molecular network was constructed using the GNPS (Global Natural Products Social Molecular Networking) platform [http://gnps.ucsd.edu]. All MS/MS spectra obtained were exported to the GNPS platform in the form of peak list (.mgf) containing comprehensive analysis information: mass spectrometer properties, precursor ions (m/z), retention times, fragment ions (m/z, intensity, charge). Subsequently, the MS/MS spectra were pairwise compared to search for spectral similarities, *i.e.* identical fragment ions and/or neutral losses. The optimal parameters were as follows: tolerance of 0.02 Da for parent and fragment, identical fragment ions and/or neutral losses; cosine score ≥ 0.7; minimum matched peaks ≥ 6; topK network 10; maximum connected component size 100; minimum cluster size 2, with no MSCluster execution for one molecular network and yes for another one. The results were downloaded and exported for visualization using the Cytoscape 3.8.2 software [https://cytoscape.org]. Compound identification relied on the spectral libraries within GNPS. The MS/MS spectra of compounds were compared to the MS/MS spectra of compounds contained in the GNPS library platform, using the following parameters: library search minimum matched peaks 6; score threshold 0.7; maximum analog search mass difference 100.

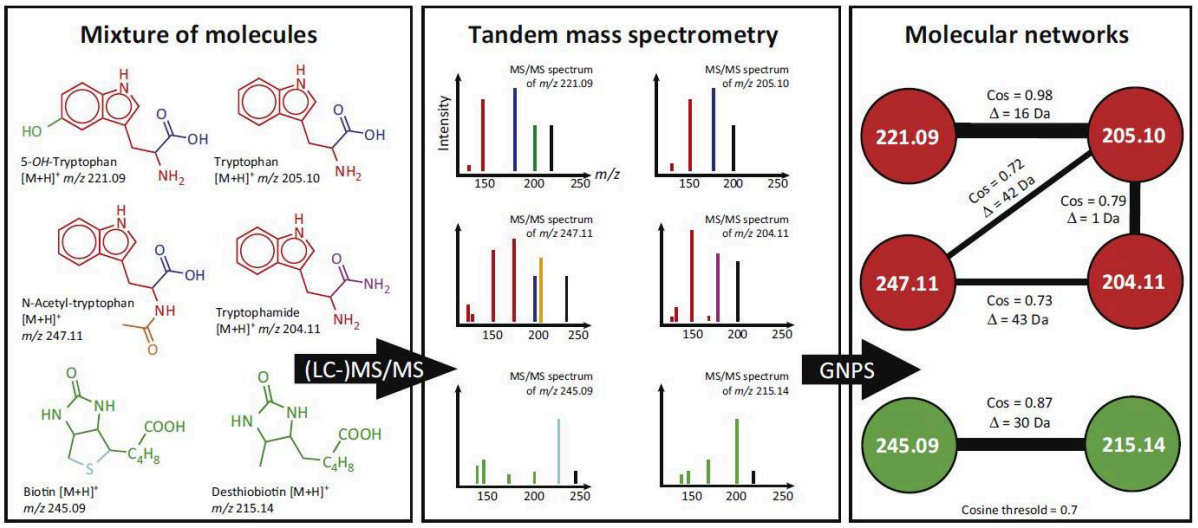


Figure 1. Diagram of the molecular network creation process, adapted from R.A. Quinn et al. (2017)[13].

3. Results

The presented study therefore focuses on the metabolomic analysis of a species from the Magnoliopsidae class for the development of a sustainable and effective cosmetic ingredient, and allows for the evaluation of the impact of geographical origin, cultivation practices, and extraction methods.

(a) (b)

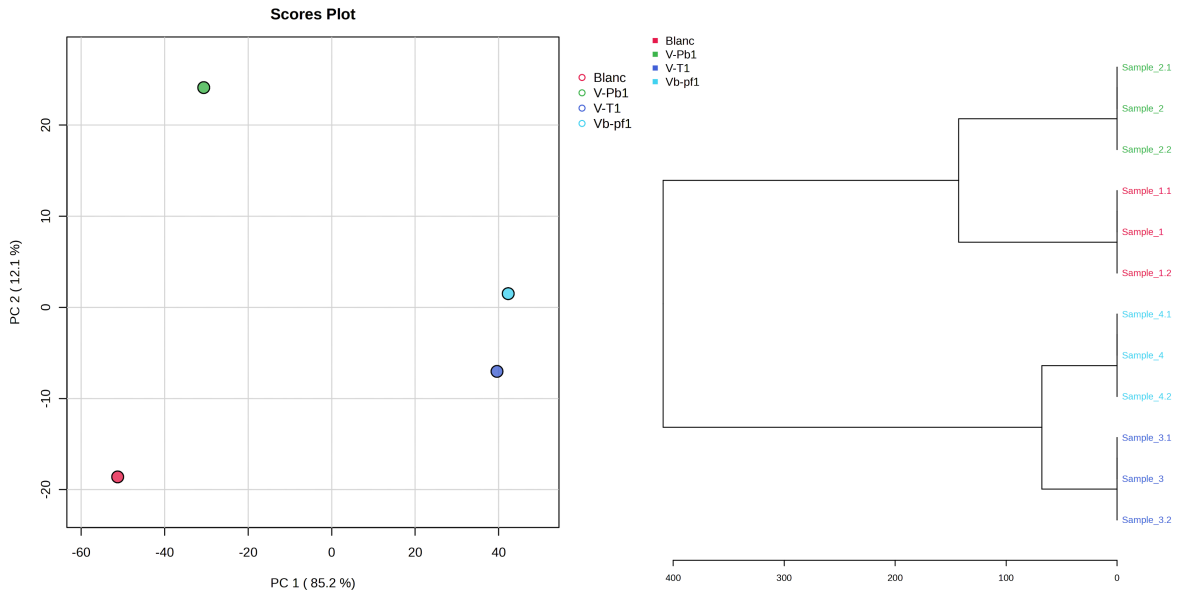
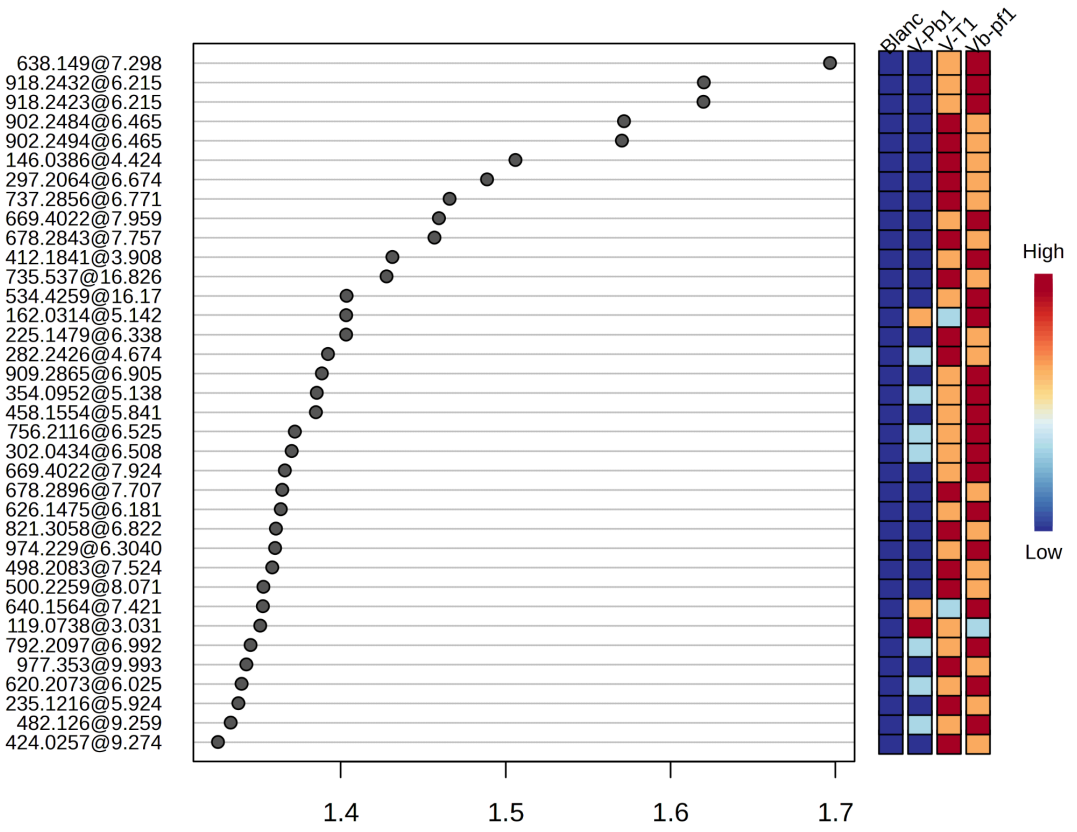


Figure 2. Unsupervised multivariate analysis of Magnoliopsida extracts : (a) PCA (Principal Component Analysis) and (b) HCA (Hierarchical Clustering Ascendant) discriminate geographical and cultivation origins.

The analysis of the score plot and the associated dendrogram reveals a clear separation into two distinct groups. The first group, distinguished by the individuals in green, corresponds to a variety originating from country 2. The second group, represented by the individuals in blue, comprises two distinct varieties, a (indoor cultivation) and b, both originating from country 1. This initial level of results suggests that the 1st growth parameters in indoor cultivation evaluated are well-suited to maintain adequate phytochemical quality (similar to pen fields cultivation,in the same country). This allows us to validate the interest in continuing research on indoor cultivation, particularly to achieve significant gains in terms of water consumption and biomass yield. Indeed, we anticipate a reduction in water consumption of approximately 90% (according to NASA data published in Nature and Food in 2020) and a potential increase of 10 to 20 times in the biomass produced on a reduced surface area.

(a)



(b)

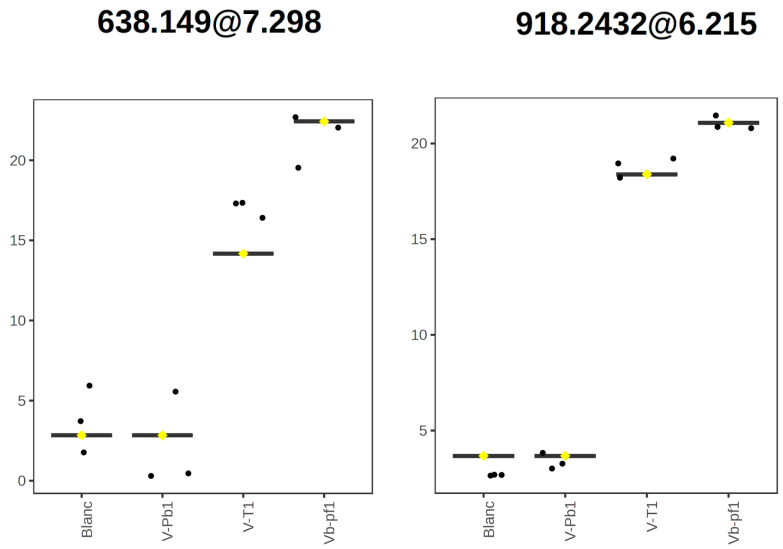


Figure 3. (a) Variable Importance in Projection (VIP) Analysis. (b) Box Plots of key discriminatory metabolites between sample groups.

The analysis of VIP scores, a key statistical indicator, reveals a significantly higher abundance of metabolites in the indoor and local circuit open-field crops in country 1 compared to the other group in country 2. In the context of this study, and to further our metabolomic research and primary efficiency analyses, we have chosen to focus on the

open-field crops of countries 1 and 2. To visualize and compare the levels of the two most discriminating variables among individuals i.e. 638.149@7.298 and 918.2432@6.215, particularly those from the two distinct geographical origins (V-Pb1 and Vb-pf1), we generated box plots. These graphical representations clearly illustrate the significant differences observed.

The most discriminating variables identified correspond to the molecular formulas $C_{28}H_{30}O_{17}$ (error 0.11 ppm) and $C_{42}H_{46}O_{23}$ (error 0.38 ppm). These molecular formulas, in conjunction with their retention times and fragmentation spectra, suggest that they belong to the family of glycosylated flavonoids. More specifically, they are likely to be Herbacetin 7-(6''-quinoylglucoside) and Kaempferol 3-(2'''-(E)-p-coumarylsophoroside)-7-glucoside, respectively.

To gain a comprehensive overview of the variation of the 700 features in our dataset, a Fold Change analysis was performed.

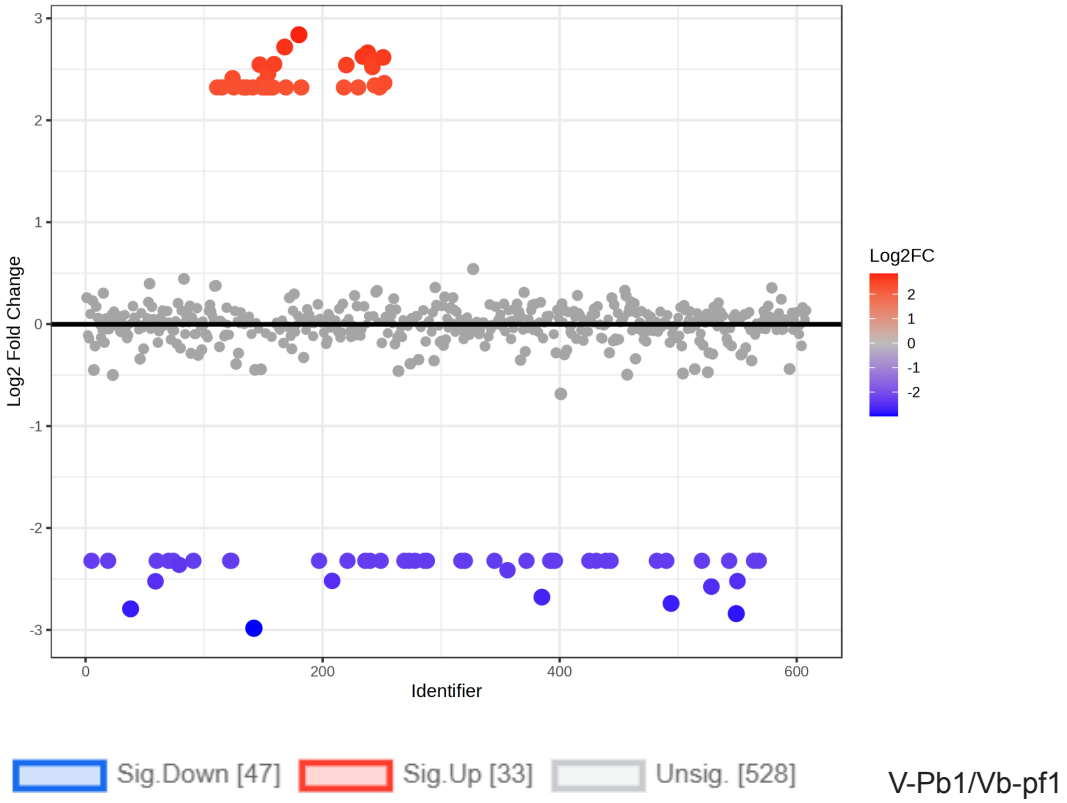


Figure 4. Metabolites FoldChange analysis : comparing two geographical locations in conventional field samples.

The FoldChange analysis allowed us to identify the most significantly discriminating compounds : 47 compounds that are overexpressed in Vb-Pf1 (Country 1), 33 compounds that are overexpressed in V-Pb1 (Country 2), and the other compounds with non-significant variation.

Table 2. Example of results differentiating between different origins using the same extraction process

Biochemical assays	V-Pb1	Vb-Pf1
%MS	2.2	1.9
% DPPH	23	32
% Tyrosinase inhibition	35	35
%COX-2 inhibition	68	74
%Hyaluronidase inhibition	15	15

Sugar dosage g/L	14	14
Polyphenols dosage g/L	0.5	0.43

However, initial efficacy tests conducted revealed no significant differences between the geographical origins. Faced with this observation, three main hypotheses can be considered: 1/ The two origins present, for this biochemical assays, equivalent quality and are therefore interchangeable, 2/ The extraction method selected for this metabolomic analysis is not the most suitable for discriminating the specific metabolites of geographical origins, 3/ The activity tests performed need to be further investigated and supplemented by other evaluations (such as hydration, collagen synthesis, etc.). To verify hypothesis 1, we explored various extraction processes (ultrasound, maceration, microwaves) with conventional and alternative solvents of varying polarities. The same analyses (heat maps, FoldChange, enzymatic efficacy tests) were applied to identify distinct and relevant extracts for a cosmetic application.

4. Discussion

The initial results of our metabolomic study on the Magnoliopsidae species reveal significant differences in the extracted metabolic profiles, depending on the extraction methods employed. Our preliminary observations suggest that solvent polarity plays a predominant role in extraction efficiency, potentially more decisive than the technique itself, whether maceration or ultrasound. Regarding antioxidant activity, our preliminary data indicate a correlation with a charge effect, as well as a notable presence of polyphenolic compounds. This observation highlights the importance of these compounds in the antioxidant potential of the studied species. However, the relationship between antioxidant activity and depigmenting activity remains uncertain at this stage. We have not been able to establish a definitive link between these two properties, which requires further investigation. To clarify these aspects and gain a more precise understanding of the bioactive compounds involved, it is imperative to continue our study with a detailed metabolomic analysis. This includes conducting detailed fragmentation spectra and developing molecular networks[14]. These approaches will allow us to more precisely identify and characterize the extracted metabolites, and to better understand their respective roles in the observed activities. In perspective, these in-depth analyses will help us to: - Confirm or refute our initial hypotheses regarding the influence of solvent polarity and extraction technique. - Validate the indoor cultivation parameters, to ensure that the phytochemical quality is similar to or different from that of open-field cultivation, for certain ecological advantages. - Identify the specific polyphenolic compounds responsible for the antioxidant activity and evaluate their potential as cosmetic ingredients. - Elucidate the relationship between antioxidant and depigmenting activity, determining if a link exists and identifying the underlying mechanisms. - Discover new bioactive compounds potentially interesting for dermocosmetic applications. - Identify if the most discriminating variables identified could be an interest for dermocosmetic activity, to orient the extraction process. In summary, these initial observations open promising avenues for the development of a sustainable and effective cosmetic ingredient from this Magnoliopsidae species. The upcoming in-depth metabolomic analyses will be crucial to validate these preliminary results and fully exploit the potential of this natural resource.

5. Conclusion

The preliminary study highlighted an equivalence between indoor and open-field cultivation. However, it is highly probable that by adjusting the cultivation parameters (elicitation for a phytochemical target of interest) and/or exploring different extraction processes, we can identify distinct advantages for indoor cultivation, going beyond the already established ecological benefits. These adjustments could potentially reinforce the interest for this specific origin. The results obtained also seem to indicate that solvent polarity plays a predominant role in extraction efficiency, potentially more so than the technique itself (maceration or ultrasound). The initial data indicate a correlation between antioxidant activity and a charge effect, as well as a notable presence of polyphenolic compounds, highlighting the importance

of these compounds for the antioxidant potential of the studied species. However, the link between antioxidant activity and depigmenting activity remains uncertain and requires further investigation. The next steps include a detailed metabolomic analysis with fragmentation spectra and the interpretation of molecular networks to more precisely identify and characterize the extracted metabolites and understand their respective roles. The perspectives include confirming the initial hypotheses, identifying the polyphenols responsible for the antioxidant activity, elucidating the relationship between antioxidant and depigmenting activities, and discovering new bioactive compounds.

References list

[1] Pratami, Diah Kartika, et al. “A Systematic Review of Metabolomics Studies on Metabolite Profiling and Phytogeographical Discrimination of Propolis.” Journal of Functional Foods, vol. 123, Dec. 2024, p. 106602. DOI.org (Crossref), https://doi.org/10.1016/j.jff.2024.106602.
[2] Hou, Jinli, et al. “Metabolomics Analysis of Astragali Radix in Shanxi Province: Investigating the Impact of Various Cultivation Methods and Growth Years on Metabolite Profiles.” Food Chemistry, Dec. 2024, p. 142492. DOI.org (Crossref), https://doi.org/10.1016/j.foodchem.2024.142492.
[3] Li, Xianbao, et al. “Untargeted Metabolomics Profiling of Purple Rice Phenolics and Their Antioxidant Activities.” LWT, vol. 214, Dec. 2024, p. 117127. DOI.org (Crossref), https://doi.org/10.1016/j.lwt.2024.117127.
[4] Sumner, L. W., Mendes, P., & Dixon, R. A. (2003). Plant metabolomics : Large-scale phytochemistry in the functional genomics era. Phytochemistry, 62(6), 817-836. https://doi.org/10.1016/s0031-9422(02)00708-2.
[5] Dettmer, K, et al. “Mass spectrometry-based metabolomics.” Mass Spectrometry Reviews, 2007. 26(1), 51-78. https://doi.org/10.1002/mas.20108.
[6] Dunn, W. B. (2008). Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. Physical Biology, 5(1), 011001. https://doi.org/10.1088/1478-3975/5/1/011001.
[7] Pierre Barbier. “Utilisation de la métabolomique pour l'étude de l'encéphalopathie hépatique et développement de nouvelles approches pour l'acquisition de données par spectrométrie de masse.” http://www.theses.fr/2019SACLS193. 2019.
[8] Lee, S. K, et al. “Evaluation of the antioxidant potential of natural products”. Combinatorial Chemistry & High Throughput Screening, 1998, 1(1), 35-46.
[9] Lim, T. Y., et al. “Evaluation of Antioxidant, Antibacterial and Anti-Tyrosinase Activities of Four Macaranga Species.” Food Chemistry, vol. 114, no. 2, May 2009, pp. 594–99. ScienceDirect, https://doi.org/10.1016/j.foodchem.2008.09.093.
[10] https://cdn.caymanchem.com/cdn/insert/701080.pdf. 28/04/2025.
[11] Bralley, Eve, et al. “Inhibition of Hyaluronidase Activity by Vitis Rotundifolia . (Muscadine) Berry Seeds and Skins.” Pharmaceutical Biology, vol. 45, no. 9, Jan. 2007, pp. 667–73. DOI.org (Crossref), https://doi.org/10.1080/13880200701545018.
[12] DuBois, Michel., et al. “Colorimetric Method for Determination of Sugars and Related Substances.” Analytical Chemistry, vol. 28, no. 3, Mar. 1956, pp. 350–56. DOI.org (Crossref), https://doi.org/10.1021/ac60111a017.
[13] Quinn, Robert A., et al. “Molecular Networking As a Drug Discovery, Drug Metabolism, and Precision Medicine Strategy.” Trends in Pharmacological Sciences, vol. 38, no. 2, Feb. 2017, pp. 143–54. DOI.org (Crossref), https://doi.org/10.1016/j.tips.2016.10.011.
[14] Messaili, Souhila, et al. “Metabolite Profiling and Molecular Network Shows Kinkeloids as Promoting of Collagen Synthesis from Combretum Micranthum.” Separations, vol. 11, no. 5, Apr. 2024, p. 132. DOI.org (Crossref), https://doi.org/10.3390/separations11050132.