Etude de la spéciation chimique de la collection nationale de violettes et mise en place d'un agro-raffinage de la violette de Toulouse

Justine Chervin

▶ **To cite this version:**

Justine Chervin. Etude de la spéciation chimique de la collection nationale de violettes et mise en place d'un agro-raffinage de la violette de Toulouse. Santé. Institut National Polytechnique de Toulouse - INPT, 2018. Français. NNT : 2018INPT0119 . tel-04222907

HAL Id: tel-04222907

https://theses.hal.science/tel-04222907

Submitted on 29 Sep 2023

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.

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Chimie Biologie Sante

Présentée et soutenue par :

Mme JUSTINE CHERVIN

le vendredi 9 novembre 2018

Titre :

Etude de la spéciation chimique de la collection nationale de violettes et mise en place d'un agro-raffinage de la violette de Toulouse

---

Ecole doctorale :

Sciences de la Matière (SDM)

Unité de recherche :

Laboratoire de Chimie Agro-Industrielle (L.C.A.)

Directeur(s) de Thèse :

M. THIERRY TALOU

M. GUILLAUME MARTI

Rapporteurs :

Mme ANNE-EMMANUELLE HAY-DE BETTIGNIES, UNIVERSITE LYON 1

M. XAVIER FERNANDEZ, UNIVERSITE DE NICE SOPHIA ANTIPOLIS

Membre(s) du jury :

M. PIERRE CHAMPY, UNIVERSITE PARIS 11, Président

M. GUILLAUME MARTI, UNIVERSITE TOULOUSE 3, Membre

Mme MARTINE URRUTIGOITY, INP TOULOUSE, Membre

M. RIMANTAS VENSKUTONIS, UNIVERSITE DE TECHNOLOGIE DE KAUNAS, Membre

M. THIERRY TALOU, INP TOULOUSE, Membre

Résumé

Le projet « Viola Tolosa » a pour objectif de valoriser une plante produite en Occitanie, la violette et plus particulièrement l'emblématique violette de Toulouse, pour des domaines essentiellement non alimentaires tels que la chimie des substances naturelles et la cosmétique. Les violettes appartiennent au genre *Viola* qui comprend plus de 500 espèces. Aujourd'hui, leurs usages sont principalement limités à des aspects ornementaux et culinaires. Néanmoins, l'intérêt croissant de la part des acteurs de la filière (industriels, cultivateurs et académiques) a conduit la région Occitanie à mettre en place le projet Viola Tolosa intitulé « Spéciation chimique de la collection nationale des violettes et mise en place d'un agro-raffinage de la violette de Toulouse ». Il comporte quatre aspects interdisciplinaires associant aspects fondamentaux et applicatifs.

La caractérisation de la centaine de plants de la collection de violettes détenue par les serres municipales de Toulouse, identifiée à 80% par des noms de cultivars ou vernaculaires, a été réalisée par l'intermédiaire d'études génétique et chimiotaxonomique. Une première étude génétique basée sur les séquences des espaces internes transcrits a permis de classer 58% de la collection au rang d'espèce. Cette étude phylogénétique a été complétée par une étude chimiotaxonomique à l'aide des profils chimiques des fractions volatiles des fleurs et non-volatiles des parties aériennes de la collection. Une projection orthogonale de structures latentes a permis d'indexer 96% de l'ensemble des plants par un nom d'espèce.

L'étude des métabolites secondaires non volatils des feuilles a été entreprise dans le but d'étudier le potentiel biologique des violettes, notamment les activités antioxydante, antifongique et inductrice des réponses immunitaires des plantes. L'étude détaillée d'un extrait hydroalcoolique de la violette de Toulouse a permis d'identifier huit composés antioxydants de la famille des flavonoïdes et des coumarines, dont trois ont été caractérisés par RMN 1D et 2D et deux *de novo* dérépliqués par réseau moléculaire. L'application sur l'ensemble de la collection a ensuite permis d'identifier six composés antioxydants, dont deux coumarines et quatre flavonoïdes, prépondérants chez deux espèces. Une relation espèce-activité a donc été mise en évidence. Au niveau des activités antifongiques, réalisées sur cinq souches de champignons, et de défenses végétales, par l'intermédiaire de l'étude de l'expression du gène marqueur « pathogenesis-related protein 1 », les résultats sont plus ambigus. Cependant, certaines espèces ont présenté une activité plus prononcée que les autres et ce criblage a permis de poser une hypothèse forte quant à l'implication des cyclotides.

Finalement, l'ensemble de ces travaux a permis d'obtenir une carte d'identité des violettes de la collection (identification génétique, profil chimique, potentiel biologique) et une description semi-quantitative de l'ensemble des groupes chimiques est proposée par combinaison des données chromatographiques du détecteur Corona (CAD) et des données spectrales.

Différentes méthodes d'extraction (électroporation, micro-ondes, CO_2 supercritique et extraction hydroalcoolique) répondant aux préceptes de la chimie verte ont ensuite été comparées afin de sélectionner celle présentant le meilleur compromis entre le cahier des charges cosmétiques et l'enrichissement en molécules d'intérêt, en vue d'un transfert technologique.

Mots-clés : Violette, Chimiotaxonomie, Metabolomique, Activité antioxydante, Valorisation industrielle

Abstract

The "Viola Tolosa" project aims to promote a plant produced in Occitanie region, the violet and especially the emblematic violet of Toulouse, essentially for non-food fields such as the chemistry of natural compounds and cosmetics. Violets belong to *Viola* genus including more than 500 species. Today, their uses are mainly limited to ornamental and culinary aspects. Nevertheless, the growing interest of the actors of the sector (industrials, growers and academicals) led the Occitanie region to implement the Viola Tolosa project entitled "Chemical speciation of the national collection of violets et establishment of an agro-refining of the violet of Toulouse ". It comprises four interdisciplinary aspects associating fundamental and applicative aspects.

The characterization of the 100 or so plants in the violet collection owned by the Toulouse municipal greenhouses, including 80% identified by cultivar or vernacular names, was carried out through genetic and chemotaxonomic studies. A first genetic study based on internal transcribed spacers conducted to classify 58% of the collection as a species. This phylogenetic study was completed by chemotaxonomic studies of chemical profiles of flowers volatile fractions and non-volatile aerial parts of the collection. Discriminant analysis of orthogonal projection to latent structure model finally allowed indexation of 96% of all plants with a species name.

Study of non-volatile secondary metabolites of leaves has also been undertaken to study the biological potential of violets, including antioxidant, antifungal and defense inducer. The detailed study of a hydroalcoholic extract of the violet of Toulouse allowed the identification of eight antioxidant compounds belonging to flavonoids and coumarins. Three of them have been characterized by 1D and 2D NMR and two were *de novo* dereplicated through molecular network. The application to the whole collection conducted to highlight six antioxidant compounds, including two coumarins and four flavonoids, predominant in two species. A species-activity relationship was therefore highlighted. Regarding antifungal activities carried out on five fungal strains, and defense inducer through the study of pathogenesis-related protein 1, the results are more ambiguous. However, some species showed better activity than others and this screening led to a strong hypothesis regarding the involvement of cyclotides.

Finally, all this work led to the establishment of an identity card of violets of the collection (genetic identification, chemical profiling, et biological potential) et a semi-quantitative description of all the species is considered by combining chromatographic data based on corona detector et spectral data.

Different methods of extraction (electroporation, microwaves, supercritical CO_2 et hydroalcoholic extraction) corresponding to green chemistry precepts were then compared in order to select the one presenting the best compromise between cosmetic specifications et enrichment in molecules of interest, for technological transfer.

Key words: Violet, Chemotaxonomy, Metabolomics, Antioxidant activity, industrial valorization

Résumé vulgarisé

La violette de Toulouse (aka *Viola alba subsp. dehnhardtii*) est emblématique de la ville rose, prêtant ainsi son image à des produits allant de la cosmétique à la confiserie. Le projet « Viola Tolosa » est destiné à valoriser les violettes dans les domaines de la chimie agro-industrielle et la dermo-cosmétique. Une collection nationale rassemblant une centaine de variétés de violettes identifiées par des noms vernaculaires a constitué le point de départ des travaux. Une première étude phylogénétique a été complétée par l'analyse des profils chimiques afin de caractériser finement l'ensemble de la collection. Les composés volatils des fleurs à maturité et leur impact olfactif ont été examinés afin de différencier la collection sur la base de profils spécifiques et d'identifier des notes odorantes d'intérêt. Le fractionnement des feuilles et l'étude des métabolites secondaires non volatils ont constitué le deuxième aspect en vue de l'obtention d'extraits présentant diverses activités biologiques. L'issue de ces recherches porte sur la proposition d'un nouvel extrait naturel pour le domaine dermo-cosmétique avec l'établissement d'une filière « Production-Extraction-Formulation ».

Vulgarized abstract

The violet of Toulouse (aka *Viola alba subsp. dehnhardtii*) is emblematic of the pink city, lending its image to products ranging from cosmetics to confectionery. The project "Viola Tolosa" is intended to promote violets in the fields of agro-industrial chemistry and dermo-cosmetics. A national collection of about 100 varieties of violets identified by vernacular names was the starting point of this work. A first phylogenetic study was completed by the analysis of chemical profiles in order to characterize the entire collection. Volatile compounds of mature flowers and their olfactory impact were examined in order to differentiate the collection on the basis of specific profiles and to identify odorous notes of interest. Leaf splitting and the study of nonvolatile secondary metabolites were the second aspect in obtaining extracts with various biological activities. The objective is to propose a novel natural extract for the dermo-cosmetic field with the setting up of a "Production-Extraction-Formulation" sector.

Remerciements

Je tiens à remercier Mr. le Professeur Pierre Champy membre du laboratoire Chimie des Substances Naturelles – Pharmacognosie de l'Université Paris Sud pour avoir accepté de présider le jury de soutenance de ma thèse. Mes remerciements s'adressent également à Mme Anne-Emmanuelle Hay-de Bettignies, maître de conférence à l'Institut des Sciences Pharmaceutiques et Biologiques de l'Université de Lyon I en phytochimie, pharmacognosie et phytothérapie et membre du laboratoire de Pharmacognosie, Phytochimie, Botanique et Homéopathie ; ainsi qu'à Mr. le Professeur Xavier Fernandez effectuant sa recherche et ses enseignements au sein de l'Université Sophia Antipolis de Nice pour avoir accepté d'être les rapporteurs de cette thèse. Je remercie très sincèrement Mme la Professeure Martine Urrutigoïty membre du Laboratoire de Chimie de Coordination de l'ENSIACET, ainsi que Mr. le Professeur Rimantas Venskutonis responsable de l'équipe de recherche Functional Food and Materials du département des Sciences technologiques de l'Université de Technologie de Kaunas (Lituanie) pour leur présence au sein du jury de thèse.

Ce travail a été réalisé en majeure partie au sein du laboratoire de Pharmacologie et de Biologie pour le Développement (Pharma-Dev), unité mixte de recherche entre l'Université Paul Sabatier III (UPS) et l'Institut de Recherche pour le Développement (IRD), sous la direction de Mr. le Professeur Nicolas Fabre et au sein du Laboratoire de Chimie et de Coordination (LCA), unité mixte de recherche entre le CNRS et l'Institut Polytechnique de Toulouse (INP), sous la direction de Mr. Gérard Vilarem. Les analyses génétiques et les essais phytosanitaires ont été réalisés au sein du Laboratoire de Recherche en Sciences Végétales (LRSV), unité mixte de recherche en l'UPS et le CNRS, sous la direction de Mr. Bernard Dumas.
L'ensemble du projet a été soutenu par la région Occitanie avec l'implication des serres municipales de Toulouse ainsi que par un industriel régional, le Groupe Berdoues Parfums et Cosmétique.

Bien évidemment, je tiens à remercier le Groupe Berdoues pour son implication dans le projet en particulier Mme Sophie Berdoues-Coudouy, présidente et Mme Valérie Bonzon-Ponnet, directrice R&D pour m'avoir apporté la confiance nécessaire à la réalisation de mes travaux au sein des laboratoires universitaires. Valérie, merci pour votre gentillesse et votre disponibilité et ce malgré les nombreuses responsabilités qui vont incombent.

J'adresse également mes remerciements à la région Occitanie pour le financement apporté à ce travail (projet CLE 13053062) ainsi qu'à l''Association Nationale de la Recherche et de la Technologie (ANRT), et aux serres municipales de Toulouse, en particulier à Mme Christelle Bringaud, responsable technique des serres et Mme Nadine Rossini, responsable des violettes, pour leur accueil et leur disposition à chacune de mes visites.

Mes remerciements se tournent ensuite vers mon directeur de thèse, Mr. Thierry Talou, pour son soutien et sa gentillesse en toutes circonstances. Je remercie très sincèrement mon co-directeur de thèse, Mr. Guillaume Marti, tant pour ses qualités scientifiques qu'humaines. Son dynamisme a insufflé un vrai souffle à ce projet et son encadrement toujours optimiste a rendu le travail sous sa tutelle très agréable et instructif. Merci à vous deux pour votre encadrement et la confiance que vous m'avez apportée dès le départ.

Je remercie Mr. Bernard Dumas, et tous les membres de l'équipe Immunité Végétale et Effecteurs, en particulier Mme Marie-Thérèse Esquerré-Tugayé pour son implication dans les travaux de recherche et sa grande culture scientifique ainsi que Mr. Laurent Camborde pour son soutien dans l'encadrement de mes stagiaires, son dynamisme et sa bonne humeur constante malgré les nombreux problèmes survenus.

Il semble que je t'ai inspiré l'envie de faire une thèse alors je te remercie d'avoir contribué à la mienne et te souhaite une bonne continuation dans la tienne !

Je souhaite ensuite remercier toutes les personnes qui ont directement été impliquées dans mes travaux.

Au sein du laboratoire Pharma-Dev, je remercie Nicolas Fabre, pour son engagement dans le projet, son accueil au sein de l'unité et ses conseils toujours constructifs ainsi que Françoise Nepveu pour son implication et son regard bienveillant sur les résultats de thèse. Merci à Pierre Perio, pour sa grande disponibilité et son incommensurable gentillesse. Pierre m'a toujours apporté son aide en toutes circonstances, que ce soit pour un appui technique ou autre. Merci pour tes conseils !

Je remercie Mme Isabelle Fourasté pour son aide dans l'étude botanique des violettes. Malgré la distance entre chez elle et le laboratoire, et les mouvements de grèves des cheminots à cette période, elle a fait le déplacement deux fois par semaine pour m'encadrer dans ce domaine que je découvrais, toujours dans la joie de vivre et je l'en remercie gracieusement.

Au sein du laboratoire Synthèse et Physico-Chimie de Molécules d'Intérêt Biologique (SPCMIB), je tiens à adresser mes remerciements à Mme Isabelle Fabing pour son dynamisme et ses conseils avisés en matière de purification des composés naturels.

Au sein de l'Institut de Chimie de Toulouse (ITC) mes remerciements s'adressent tout particulièrement à Mr. Marc Vedrenne et Mr. Pierre Lavedan su service commun de RMN pour leur expertise et leur patience dans la réalisation de mes analyses RMN.

Au sein du LCA, je remercie Sarah Grivot pour sa gentillesse et son aide dans la réalisation des essais sur les composés volatils. Je remercie Romain Armengau pour sa bonne humeur et ses encouragements ainsi qu'Anne Lung pour sa bienveillance lors de ma journée broyage.

J'adresse mes remerciements aux différentes stagiaires que j'ai eu le plaisir d'encadrer au cours de ces trois ans, Juliette, Johanna, Audrey, Léa et Marjorie. Merci pour votre investissement et votre motivation ayant permis de faire avancer les travaux de thèse dans les différentes thématiques.

Je souhaite également remercier l'ensemble des collègues de travail avec lesquels j'ai passé ces trois ans.

Je remercie tout d'abord l'ensemble des équipes Peps et Redstress, dont Marieke, Anne-Cécile, Karine, Yann, Polo, Geneviève, Mohamed et nouvellement arrivée Hélène, pour les instants partagés au laboratoire et en dehors lors des lunchs et barbecues improvisés chez les uns et les autres. Merci à l'équipe Biocid, en particulier Sandra, Matthias, Guillaume, Maylis et Stéphane pour m'avoir acceptée dans leur laboratoire sur la fin de ma thèse et pour leur gentillesse.

Je tiens à exprimer mes remerciements tout particuliers aux deux gestionnaires de Pharma-Dev, Franck et Ingrid, toujours de bonne humeur et toujours disposés à répondre à nos attentes. Ingrid, merci pour tes conseils et ton écoute durant les derniers mois de ma thèse. J'apporte aussi une dédicace toute particulière à Valérie C. pour ta bonté d'âme et ton extrême gentillesse. Merci pour ton soutien sans faille, ton optimisme et ta bonne humeur constante qui font de toi l'une des âmes de ce laboratoire !

Enfin, je ne saurai terminer ce tour en remerciant tous les doctorants de Pharma-Dev, Fatima, Asih, Monivan, Chio et bien entendu Manon. Monivan et Chio, j'ai été très heureuse de vous avoir comme collègues de bureau et je vous remercie de m'avoir supportée pendant deux ans et d'avoir fait vivre une bonne entente au sein de nos quelques mètres carrés. Manon, nous avons supporté ces trois années

ensemble du début à la fin, à coup de mots fléchés entre midi et deux pour nous détendre et oublier nos tracas. Une belle amitié s'est tissée qui je l'espère continuera encore longtemps autour de cours de couture pour mettre un an à confectionner une robe par exemple (et j'espère bien d'autres) !

Et pour finir, je remercie toutes les personnes qui me suivent depuis le début

Marie, notre belle et profonde amitié n'a cessé de se renforcer depuis 6 ans déjà ! L'on se suit depuis l'école d'ingénieur et par chance nos laboratoires se trouvaient à quelques rues l'un de l'autre. Coïncidence ? Je parlerai plutôt de destin ! Nous nous sommes épaulées toutes deux dans nos thèses respectives et t'avoir à mes côtés pour les coups durs et les moments heureux de la thèse et de la vie en général a été très important pour moi. J'espère t'avoir apporté autant que toi tu l'as fait. Nos sessions sport et nos sorties détente au théâtre, au cinéma et j'en passe m'ont apporté l'équilibre nécessaire pour affronter ces trois années. Je ne cesse de le répéter, mais nous étions faites pour se rencontrer !

Je remercie également ma famille, en particulier ma mère et mes frères. J'ai conscience de vous avoir largement rabattu les oreilles avec mon projet souvent au cœur de nos discussions et même si je voyais que vous ne suiviez pas tout ce que je disais, vous portiez toujours autant d'attention à mes propos et je souhaiterais vous en remercier. Josselin, nos conversations téléphoniques me remontaient toujours le moral dans les coups durs, et ce même si tu as pris la fâcheuse habitude de téléphoner quand tu étais sur les routes lors de tes déplacements pour que le temps passe plus vite. Malgré la distance, votre soutien m'a été très précieux !

Et pour finir je souhaiterais te remercier Tristan. Je n'ai pas les mots pour t'exprimer à quel point ton soutien, ton intérêt pour mon travail (à toi aussi je t'ai rabattu les oreilles pendant trois ans avec mes violettes) et ta présence à mes côtés au cours de ces trois années m'ont été plus que nécessaire pour suivre le cap. Nous avons traversé quelques épreuves difficiles mais ton sang-froid, tes encouragements et ta confiance m'ont permis de ne pas flancher.

Communications scientifiques

- **Posters**

Colloque Recherche 2016 Chimie et Procédés du Végétal, Montpellier, 27-29 janvier 2016. **A comprehensive chemodiversity screening of *Viola* genus in Midi-Pyrennees by a holistic approach**, Justine Chervin, G. Marti, P.Perio, N. Martins, K. Reybier, F. Nepveu, N. Fabre, G. Vilarem, T. Talou,

Renewable Resources et Biorefineries. Ghent, Belgique, 30Mai-1erJuin 2016. **Metabolomics as a tool for biorefinery: case study of a comprenhensive chemotaxonomy screening of violets**. Justine Chervin, G. Marti, P.Perio, N. Martins, K. Reybier, F. Nepveu, N. Fabre, G. Vilarem, T. Talou.

9th Joint Natural Products Conference 2016, Copenhague, Danemark, 24-27 Juillet. **LC-MS based metabolomic approach applied to bioactive molecules identification: case study of a comprehensive chemodiversity screening of *Viola* genus**. Justine Chervin, G. Marti, P. Perio, N. Martins, K. Reybier, F. Nepveu, N. Fabre, G. Vilarem, T. Talou.

47th International Symposium on Essential Oils (ISEO2016), Nice, France, 11-14 September 2016. **Chemotaxonomy vs Genotaxonomy of Viola genus: Can native volatile compounds of flowers be genuine tracers for violets differentiation?** Justine Chervin, B. Dumas, L. Camborde, M-T. Esquerré-Tugayé, F. Nepveu, N. Fabre, G. Marti, T. Talou.

Projet Viola Tolosa. **Biodiversité des fragrances de violettes et valorisation d'un patrimoine végétal**. Fête de la violette, 4-5 mars 2017, place du Capitole, Toulouse.

Dixièmes journées de l'école doctorale Science de la Matière, Ensiacet, Toulouse, 17-18 mai 2017. **Phylogeny of violets through a multiplexed approach: metabolomics vs chemotaxonomy vs genotaxonomy of *Viola* genus**. Justine Chervin, B. Dumas, L. Camborde, M-T. Esquerré-Tugayé, F. Nepveu, N. Fabre, G. Marti, T. Talou.

Projet Viola Tolosa. **Valorisation du patrimoine végétal de la région Occitanie**. Journée portes ouvertes des serres municipales de Toulouse, mai 2018.Communication orales

- **Communication orale**

Trends in Natural Product Research – PSE Young Scientists' Meeting Lille 2017 Natural Products in Health, Agro-Food and Cosmetics, Lille, France, 28 juin-1 juillet 2017. **Deciphering the phylogeny of violets through a multiplexed approach: genetic markers combined to metabolomics and native volatiles profiling.** Justine Chervin, Bernard Dumas, Marie-Thérèse Esquerre, Nicolas Fabre, Karine Reybier, Guillaume Marti, Thierry Talou.

- **Publications scientifiques**

Chervin, Justine, Pierre Perio, Nathalie Martins-Froment, Chiobouaphong Pharkeovilay, Karine Reybier, Françoise Nepveu, Nicolas Fabre, Thierry Talou, Valérie Bonzon-Ponnet, et Guillaume Marti. 2017. '**Dereplication of Natural Products from Complex Extracts by Regression Analysis and Molecular Networking: Case Study of Redox-Active Compounds from *Viola alba subsp. dehnhardtii***'. *Metabolomics* 13 (8). https://doi.org/10.1007/s11306-017-1227-6.

Justine Chervin, Thierry Talou, Marjorie Audonnet, Bernard Dumas, Laurent Camborde, Marie-Thérèse Esquerré-Tugayé, Christophe Roux, Guillaume Cabanac, Guillaume Marti. '**Deciphering the phylogeny of violets based on multiplexed genetic and metabolomic approaches**' (en reviewing dans *Phytochemistry*).

SOMMAIRE ILLUSTRE

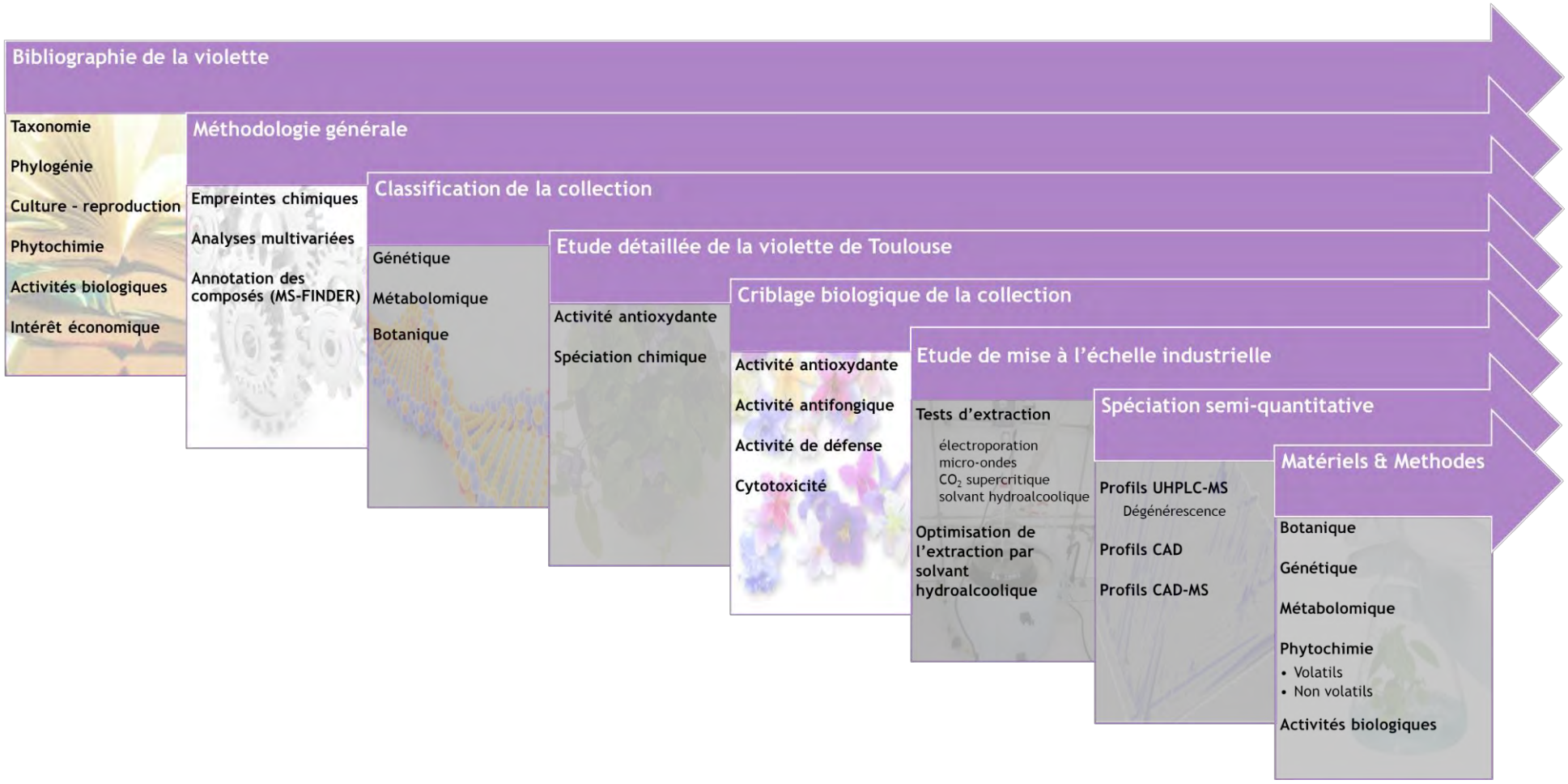


Table des matières

Résumé 2
Résumé vulgarisé 4
Vulgarized abstract 4
Remerciements 5
Communications scientifiques 8
Table des matières 10
Liste des abréviations 15
Table des illustrations 17
Introduction 23
PARTIE I ***Etat de l'art*** 26
I.1 Taxonomie de la violette 27
I.1.1 Appartenance botanique 27
I.1.2 Répartition géographique 27
I.1.3 Caractéristiques morphologiques 28
I.1.4 Caractéristiques caryologiques 29
I.2 Phylogénie des violettes 30
I.3 Culture et reproduction 30
I.4 Phytochimie des violettes 31
I.4.1 Composés volatils de la violette 31
I.4.1.1 Les fleurs 32
I.4.1.2 Les feuilles 33
I.4.2 Composés non volatils de la violette 33
I.5 Propriétés biologiques des *Viola* 37
I.5.1 Propriétés antioxydantes 38
I.5.1.1 Stress oxydant 38
I.5.1.2 Antioxydants naturels 38
I.5.1.2.1 Défenses enzymatiques 38
I.5.1.2.2 Défenses non enzymatiques 39
I.5.2 Propriétés phytosanitaires 41
I.6 Intérêts économiques de la violette 42
PARTIE II Résultats & Discussions 45
II.1 Méthodologie générale de la thèse 46
II.2 Classification des violettes de la collection 47
II.2.1 Contexte 47
II.2.2 Classification par analyse génétique 47
II.2.2.1 Séquences des régions ITS 47
II.2.2.2 Construction des arbres phylogénétiques 49
II.2.2.2.1 Arbre de distance 49
II.2.2.2.2 Reconstruction phylogénétique 49

II.2.2.3 Conclusion 52
II.2.3 Extension de la classification par analyses phytochimiques 52
II.2.3.1 Intégration des profils volatils des fleurs 52
II.2.3.1.1 Optimisation de l'analyse HS-SPME par plan de Doehlert 53
II.2.3.1.2 Analyse métabolomique des données GC-MS 57
II.2.3.1.3 Identification de biomarqueurs des fleurs 59
II.2.3.1.4 Étude des profils olfactifs 62
II.2.3.2 Intégration des profils non volatils des parties aériennes 63
II.2.3.2.1 Analyse métabolomique des données UHPLC-HRMS/MS 63
II.2.3.2.2 Identification de biomarqueurs des parties aériennes 66
II.2.3.3 Combinaison des données ITS et UHPLC-MS 70
II.2.4 Monographie végétale 72
II.2.4.1 Description des espèces 72
II.2.4.2 Analyse morphologique 74
II.2.4.2.1 L'inflorescence 74
II.2.4.2.2 La feuille 76
II.2.4.3 Analyse anatomique 78
II.2.4.3.1 Les bractéoles 78
II.2.4.3.2 L'épiderme des feuilles 80
II.2.5 Conclusion 84
II.3 Etude détaillée de la violette de Toulouse 85
II.3.1 Contexte 85
II.3.2 Fractionnement sur cartouche SPE C_{18} 85
II.3.3 Analyse UHPLC-HRMS/MS 87
II.3.4 Potentiel antioxydant de l'extrait hydroalcoolique 88
II.3.4.1 Résultats des essais DPPH et RPE 88
II.3.4.2 Analyses statistiques des résultats antioxydants 89
II.3.4.3 Identification et déréplication des molécules d'intérêt 90
II.3.4.4 Validation de l'activité antioxydante 99
II.3.5 Conclusion 101
II.4 Potentiels biologiques des violettes de la collection 102
II.4.1 Contexte 102
II.4.2 Criblage de l'activité antioxydante 102
II.4.2.1 Résultats du test DPPH 102
II.4.2.2 ANOVA des données DPPH 103
II.4.2.3 Dosage des polyphénols et flavonoïdes totaux 104
II.4.2.4 ACP et OPLS des données DPPH 105
II.4.2.5 Dosage UHPLC-UV des flavonoïdes et coumarines antioxydants 107
II.4.2.6 Conclusion 109
II.4.3 Criblage des activités phytosanitaires 110
II.4.3.1 Criblage de l'activité antifongique 110
II.4.3.1.1 Résultats d'inhibition de la croissance de cinq souches 110
II.4.3.1.2 Analyse statistique des données antifongiques 115
II.4.3.1.3 Vérification de l'activité des cyclotides 119
II.4.3.2 Criblage de l'activité élicitrice de défense des plantes 126
II.4.3.2.1 Résultat de l'activité GUS 126
II.4.3.2.2 Analyse statistique des données SDP 127
II.4.3.3 Conclusion des activités phytosanitaires 129
II.4.4 Criblage de l'activité cytotoxique 131
II.4.4.1 Résultats du criblage cytotoxique 131
II.4.4.2 Analyse statistique des données de cytotoxicité 133

II.4.4.2.1 ANOVA 133
II.4.4.2.2 ACP et OPLS 134

II.5 **Etude de mise à l'échelle industrielle** 135
II.5.1 Description des méthodes d'extraction testées 135
II.5.2 Profils UHPLC-HRMS 136
II.5.3 Rendement et potentiel antioxydant 136
II.5.3.1 Rendement d'extraction 136
II.5.3.2 Dosage des polyphénols 137
II.5.3.3 Potentiel antioxydant 139
II.5.3.3.1 « Multivariate design » 139
II.5.3.3.2 Résultats des essais DPPH et RPE 140
II.5.4 Optimisation de l'extraction hydroalcoolique 141
II.5.4.1 Montage 142
II.5.4.2 Résultats d'optimisation 142
II.5.4.2.1 Rendement d'extraction 142
II.5.4.2.2 Potentiel antioxydant 143
II.5.4.2.3 Profils UHPLC-HRMS/MS 143
II.5.5 Conclusion 145

II.6 Spéciation semi-quantitative de la collection 145
II.6.1 Contexte 145
II.6.2 Principe de traitement des données 145
II.6.2.1 Annotation et nettoyage des pics UHPLC-MS 145
II.6.2.2 Formation de la liste de pic du CAD 146
II.6.2.3 Quantification avec le CAD 146
II.6.2.4 Combinaison des données UHPLC-MS et du CAD 148
II.6.3 Optimisation du processus sur des standards 148
II.6.4 Application aux espèces de violettes 151
II.6.4.1 Extraction hydroalcoolique exhaustive et rapide 152
II.6.4.2 Liste de pics UHPLC-MS et annotation des variables 153
II.6.4.3 Liste de pics quantitatifs du CAD 154
II.6.4.4 Combinaison des données UHPLC-MS et CAD 155
II.6.5 Conclusion 160

PARTIE III Conclusion générale 161

PARTIE IV Partie experimentale 166

IV.1 Conditionnement de la matière premiere 167
IV.1.1 Matière végétale à disposition 167
IV.1.2 Récolte et séchage des plantes 167

IV.2 Etude botanique 167
IV.2.1 Montage des coupes 167
IV.2.1.1 Solution d'hydrate de chloral 167
IV.2.1.2 Préparation des coupes 167
IV.2.1.3 Examen microscopique 168

IV.3 **Méthode d'étude génétique** 168
IV.3.1 Protocoles d'amplification, purification et séquençage de l'ADN 168
IV.3.1.1 Solution de préparation 168
IV.3.1.2 Cycle PCR et électrophorèse 168
IV.3.1.3 Purification 169
IV.3.1.4 Séquençage 169

IV.3.2 Protocole d'analyse phylogénétique 170
IV.3.2.1 Analyse de similarité de séquences 170
IV.3.2.2 Arbre phylogénétique 170
IV.4 Méthode **d'analyse métabolomique et réseau moléculaire** 170
IV.4.1 Conversion et alignement des données 171
IV.4.2 Identification des composés 171
IV.4.3 Analyse statistique 172
IV.4.3.1 ANOVA 172
IV.4.3.2 ACP et régression linéaire 172
IV.4.4 Réseau moléculaire 173
IV.5 **Méthodes d'études phytochimiques** 173
IV.5.1 Méthodes d'étude des composés volatils des fleurs 173
IV.5.1.1 Extraction par espace de tête et microextraction en phase solide 173
IV.5.1.1.1 Standardisation de la méthode d'extraction HS-SPME 174
IV.5.1.1.2 Optimisation de l'extraction HS-SPME par un plan d'expérience 175
IV.5.1.2 Méthode d'analyse sensorielle 175
IV.5.1.2.1 Le « Champ des odeurs »® 175
IV.5.1.2.2 Analyse des effluves odorantes par GC/O 176
IV.5.1.3 Méthode d'analyse par chromatographie en phase gazeuse 177
IV.5.1.3.1 Méthode analytique GC-O-MS 178
IV.5.1.3.2 Méthode d'analyse métabolomique des données GC-MS 178
IV.5.2 Méthodes d'étude des métabolites secondaires non volatils des feuilles 179
IV.5.2.1 Broyage des feuilles 179
IV.5.2.2 Méthodes d'extraction 179
IV.5.2.2.1 Extraction hydroalcoolique classique 179
IV.5.2.2.2 Extraction hydroalcoolique optimisée 180
IV.5.2.2.3 Extraction par électroporation 180
IV.5.2.2.4 Extraction par CO_2 supercritique 181
IV.5.2.2.5 Extraction assistée par micro-ondes 182
IV.5.2.2.6 Fractionnement par SPE 182
IV.5.2.3 Méthodes d'analyse chimique 183
IV.5.2.3.1 La chromatographie liquide couplée à la spectrométrie de masse 183
IV.5.2.3.2 Le spectre UV-visible 185
IV.5.2.3.3 La résonance magnétique nucléaire 185
IV.5.2.4 Méthodes d'analyse métabolomique des données UHPLC-HRMS 186
IV.5.2.4.1 Pour la violette de Toulouse 186
IV.5.2.4.2 Pour la collection de violettes 186
IV.5.2.4.3 Pour les essais d'extraction 187
IV.5.2.4.4 Identification des métabolites 187
IV.5.2.5 Méthodes de construction des réseaux moléculaires 187
IV.5.2.6 Méthodes de dosage des polyphénols et cyclotides 187
IV.5.2.6.1 Dosage UHPLC-UV des coumarines 188
IV.5.2.6.2 Dosage LC-UV des flavonoïdes 188
IV.5.2.6.3 Dosage Folin-Ciocalteu des polyphénols totaux 189
IV.5.2.6.4 Dosage $AlCl_3$ des flavonoïdes 189
IV.5.2.6.5 Dosage BCA des cyclotides 190
IV.6 Méthodes de détermination des potentiels biologiques des extraits 190
IV.6.1 Évaluation de l'activité antioxydante 190
IV.6.1.1 Méthode DPPH 191
IV.6.1.2 Méthode de résonance paramagnétique électronique 192

IV.6.2 Évaluation de la cytotoxicité sur cellules VERO 193
IV.6.2.1 Repiquage des cellules 194
IV.6.2.2 Dénombrement des cellules 194
IV.6.2.3 Test de cytotoxicité 195
IV.6.3 Évaluation des activités phytosanitaires 195
IV.6.3.1 Activité antifongique 195
IV.6.3.1.1 Préparation des souches 196
IV.6.3.1.2 Test antifongique en milieu liquide 196
IV.6.3.2 Activité de défense des plantes 197
REFERENCES 198
ANNEXES 210
PUBLICATIONS 240

Liste des abréviations

ACN : Acétonitrile
ACP : Analyse en composante principale (Principal component analysis)
ANOVA : Analyse de variance
APCI : Ionisation à pression atmosphérique
CA : Charbon actif
CAD : Détcteur Corona, détecteur d'aérosols chargés (Charged Aerosol Detector).
CAH : Classification ascendante hiérarchique
CAR : Carboxen
CI50 : Concentration inhibitrice à 50%
COSY : Correlation spectroscopy
CP : Composante Principale
d : Doublet
DAD : Détecteur à Barrettes de Diodes
dd : Doublet de doublet
DNP : Dictionnaire des Produits Naturels (Dictionary of Natural Products)
Doxo : Doxorubicin
DPPH : 2,2-diphényl-1-picryl-hydrazyl
DVB : Divinylbenzène
ESI : Ionisation par électronebulisation (Electrospray ionisation)
EtOH : Ethanol
FA : Acide formique
GC : Chromatographie en phase gazeuse (Gaz chromatography)
GC/O : Chromatographie en phase gazeuse couplée à un olfactomètre (Gaz chromatohraphy-olfactometry)
HMBC : Heteronuclear multiple bond correlation
HS : Technique de l'espace de tête (Headspace)
HSQC : Heteronuclear single quantum coherence
Hz : Hertz
ITS : Espaces internes transcrits (Internal transcribed spacers)
J : Constante de couplage
LC : Chromatographie en phase liquide (Liquid Chromatography)
LC-MS : Chromatographie en phase liquide couplée à un spectromètre de masse (Liquid chromatography-Mass spectrometry)
m : Multiplet
m/z : Rapport masse-sur-charge
MeOH ; Méthanol
MF : Formule brute (molecular formula)
MS : Spectrométrie de masse (Mass spectrometer)
MS/MS : Spectrométrie de masse en tandem
NCBI : National Center for Biotechnology Information
NOESY : Nuclear overhauser effect spectroscopy
OPLS : Régression orthogonale des Moindres Carrés (Orthogonal Partial Least Square Analysis)
OPLS-DA : Analyse discriminante selon la projection orthogonale des structures latentes (Orthogonal Projection to Latent Structure Discriminant Analysis)
PA : Polyacrylate
pb : Paire de base

PCR : Polymérisation par réaction en chaîne (Polymerase chain reaction)
PDMS : Polydiméthylsiloxane
PLS : Régression partielle des moindres carrés (Partial Least Square Analysis)
PLS-DA (Partial least square discriminant analysis)
ppm : Parties par million
RMN : Résonnance Magnétique Nucléaire
RPE : Résonnance Paramagnétique Electronique
RT : Temps de rétention (Retention Time)
s : Singulet
SI : Silice
SIMCA : Logiciel de traitement de données multivariées (Standard in multivariate data analysis)
SPE : Extraction en phase solide
SPME : Microextraction en Phase Solide (Solid Phase Microextraction)
t : Triplet
TIC : Courant ionique total (Total ion current)
ToFMS : Spectromètre de Masse à Temps de Vol (Time of Flight Mass Spectrometer)
UHPLC : Chromatographie en phase Liquide Ultra Haute Performance
δ : Déplacement chimique

Table des illustrations

Tables des figures

Figure 1. Répartition nationale de quelques espèces de pensées et de violettes d'après le réseau botanique francophone (Tela Botanica) ---- 28
Figure 2. Caractéristiques morphologiques des pensées (a) et violettes (b) (images issues de (Beille, 1909)) ---- 28
Figure 3. Structures chimiques des noyaux cyanidine (a), delphinidine (b), malvidine (c), et petunidine (d) ---- 29
Figure 4. Caryogrammes d'un échantillon de violette de Tourette (a) et de Toulouse (b) d'après Chauchard et al. ---- 29
Figure 5. Schématisation des espaces internes transcrits ---- 30
Figure 6. Répartition statistique des composés isolés chez le genre Viola ---- 36
Figure 7. Réduction de H_2O_2 par Fe^{2+} par la réaction de Fenton (a) et réactions associées (2, 3). Formation du radical hydroxyle par réaction de Haber-Weiss (4) ---- 40
Figure 8 : Réaction de piégeage des ERO ($R^{\bullet}$) par un noyau catéchol avec formation d'une o-quinone ---- 40
Figure 9. Groupes structuraux d'intérêt dans le potentiel antioxydant des flavonoïdes ---- 41
Figure 10 : Exemple de structure tridimensionnelle et de séquence d'un cyclotide (Ireland et al., 2006) ---- 41
Figure 11. Serre des violettes localisée aux serres municipales de Toulouse ---- 43
Figure 12. Carte des origines géographiques des violettes étudiées mettant en évidence la diversité de phénotypes. Seuls les noms encadrés en rouge sont scientifiques. ---- 43
Figure 13. Évolution des volumes traités (t) de violettes de Tourette et prix du kg (€) de 1996 à 2015 ---- 44
Figure 14 : Stratégie d'annotation des composés ---- 46
Figure 15. Partie de séquence de Luxonne (V91) illustrant la présence de doubles pics ---- 48
Figure 16. Alignement des séquences des huit échantillons distincts (a) et caractéristiques de chacune (b) ---- 48
Figure 17. Cladogramme de similarité de séquences par une méthode de distance avec indication des valeurs bootstrap supérieures à 80% ---- 49
Figure 18. Dendrogramme du maximum de parcimonie (a) et cladogramme circulaire du maximum de vraisemblance (b) de l'analyse phylogénétique des violettes avec indication des valeurs bootstrap supérieures à 60% ---- 51
Figure 19. Méthode de classification basée sur l'établissement de modèles statistiques de base et de prédiction. ---- 52
Figure 20. Principe du plan de Doehlert (a) et illustration d'une surface de réponse traduisant un maximum (b), un minimum (c) et un minimax (d) ---- 53
Figure 21. Surfaces de réponse dans le cas de l'analyse du nombre de pics (a) et dans le cas de l'analyse de l'intégration (b) ---- 55
Figure 22. Nouvelles courbes de réponses dans le cas de l'étude des pics (a) et dans le cas de l'étude de l'intégration (b) pour un temps d'équilibre fixé à 65 minutes ---- 55
Figure 23. Étude cinétique sur la violette de Toulouse illustrant la variabilité quantitative (a), analyse en composantes principales avec les données brutes (b) et les données binarisée (c) ---- 58
Figure 24. Modèle OPLS-DA des violettes identifiées (a), test de permutation (b), prédiction des violettes inconnues matérialisées par une étoile (c) et répercussions sur la classification de la collection (d) ---- 59
Figure 25. Identification in silico des composés volatils biomarqueurs de certaines espèces ---- 60
Figure 26. Mise en relation des profils volatils des violettes et de leur affiliation à une espèce particulière ---- 61

Figure 27. Types d'odeur détectés dans les échantillons de violette et nombre de violettes associées ---- 62
Figure 28. Superposition du chromatogramme GC et de l'olfactogramme pour la violette de Toulouse ---- 62
Figure 29. Comparaison du profil chimique de la violette de Toulouse avant et après traitement C_{18} (a) et couleur de l'extrait avant (b) et après traitement (c) ---- 63
Figure 30. ACP des profils UHPLC-HRMS des extraits de parties aériennes des violettes de la collection ---- 64
Figure 31. Modèle OPLS-DA des violettes identifiées (a), test de permutation (b), prédiction des violettes inconnues matérialisées par une étoile (c) et répercussions sur la classification de la collection (d) ---- 65
Figure 32. Profils UHPLC des différents extraits (a) et analyse en composantes principales relatifs à l'étude cinétique des métabolites secondaires de la violette de Toulouse ---- 66
Figure 33. Réseau de corrélation basé sur l'ensemble des données UHPLC-HRMS en modes NI et PI des extraits de violette (a) et identification de biomarqueurs caractéristiques (b). ---- 67
Figure 34. Arbre consensus combinant les séquences ITS et les données UHPLC-MS construites avec la méthode de parcimonie. ---- 71
Figure 35. Analyse CAH basée sur l'ACP des données UHPLC confirmant les cinq groupes distincts corrélés aux espèces et la scission des V. sp Hearn (H1 et H2) et V. odorata (O1 et O2) en deux groupes ---- 72
Figure 36 : Phénotypes de fleurs appartenant à V. alba (a), V. sp Hearn (b), V. subs Rostratae (c), V. suavis (d), et V. odorata (e) ---- 74
Figure 37. Grains de pollen de V. alba (a) ; V. tricolor (b); V. arvensis (c) ---- 74
Figure 38 : Description du pédoncule floral de V. alba (a), V. sp Hearn (b), V. suavis (c), V. odorata (d), V. labradorica (e) et V. grypoceras (f) ---- 75
Figure 39. Fleurs séchées (1ère photo) et fleurs fraiches (2nd photo) pour deux violettes de V. odorata (a), V. alba (b), V. suavis (c), V. subs Rostratae (d.1 : V. grypoceras et d.2 : V. labradorica) (d) et V. sp Hearn (e) ---- 76
Figure 40. Feuilles séchées (1ère photo) et feuilles fraiches (2nd photo) pour deux violettes V. odorata (a), V. alba (b), V. suavis (c), V. subs Rostratae (d.1 : V. grypoceras et d.2 : V. labradorica) (d) et V. sp Hearn (e) ---- 77
Figure 41. Autres formes de feuilles observées dans la collection de violettes ---- 77
Figure 42. Poils sécréteurs (a, b), hydathode (c) et poils tecteurs (d, e) de V. alba vus au microscope ---- 78
Figure 43. Observations microscopiques de nervures (a) et détection de la présence de mucilage au bord (b) et de cellules à mucilages dans le parenchyme (c) chez V. alba ---- 79
Figure 44. Observation microscopique de prismes d'oxalate de calcium au niveau de l'épiderme en lumière artificielle (a) et polarisée avec discernement de la présence de macles (b) chez V. alba - 79
Figure 45. Stomate anisocytique de V. alba vu au microscope optique ---- 79
Figure 46. Coupe transversale mettant en évidence les différentes strates (a) : la cuticule striée, l'épiderme supérieur [1], le parenchyme palissadique [2], le parenchyme lacuneux [3] et l'épiderme inférieur [4]. Cellule hypertrophiée de l'épiderme non striée (b) ---- 81
Figure 47. Identification et classification des violettes selon les méthodes employées ---- 85
Figure 48. Fractionnement de l'extrait éthanolique de la violette de Toulouse et profils chromatographiques UHPLC des fractions SPE collectées ---- 86
Figure 49. Bilan de masse moyen (a) et graphique associé (b) du fractionnement de l'extrait brut sur cartouche C_{18} ---- 87
Figure 50. ACP des fractions obtenues par SPE, des extraits bruts et des contrôles qualité ---- 87
Figure 51. Valeur des CI_{50} et intégration des signaux RPE de chaque fraction (a) et graphique associé (b) ---- 89
Figure 52. Modèle OPLS des données DPPH (a) et test de permutation associé (b) ---- 90

Figure 53. Classement des métabolites sur la base des coefficients obtenus par l'OPLS (a) et zoom sur les huit premiers hits (b) avec détails de leurs caractéristiques (c) ---------- 90
Figure 54. Réseau moléculaire des métabolites secondaires de la violette de Toulouse ---------- 91
Figure 55. Emphases des groupes des flavonoïdes et des coumarines aux forts potentiels antioxydants ---------- 92
Figure 56. Spectre RMN ^{1}H du composé 1 (500 MHz, DMSO) ---------- 93
Figure 57. Spectre RMN ^{13}C du composé 1 (500 MHz, DMSO) ---------- 94
Figure 58. Spectre RMN NOESY du composé 1 (500 MHz, DMSO) ---------- 94
Figure 59. Spectre RMN du ^{1}H des composés 2 et 23 (500 MHz, DMSO) ---------- 95
Figure 60. Spectre RMN du ^{13}C du composé 2 (500 MHz, DMSO) ---------- 96
Figure 61. Spectre RMN NOESY du mélange des composés 2 et 23 (500 MHz, DMSO) avec mise en évidence d'une interaction des protons du CH_3 et du glucose justifiant la substitution proposée -- 96
Figure 62. Extrait du spectre HMBC du composé 23 (en mélange avec 2) (500 MHz, DMSO) ---------- 97
Figure 63. Spectre MS/MS du composé 20, $C_{22}H_{28}H_{16}$ ---------- 98
Figure 64. Séparation sur couche mince des composés antioxydants et spectres de masse associés des fractions Fr2 (a) et Fr3 (b). ---------- 100
Figure 65. Capacité des composés 1 et 2 à piéger des radicaux ---------- 101
Figure 66. Graphique Q-Q de l'analyse de normalité des résidus centrés réduits des données DPPH ---------- 104
Figure 67. Polyphénols et flavonoïdes totaux déterminés par dosage colorimétrique pour certains échantillons ---------- 105
Figure 68. Moyenne de la teneur dans la poudre des coumarines et des flavonoïdes antioxydants déterminés par dosage UHPLC-UV pour chaque espèce ---------- 108
Figure 69. Structures chimiques des flavonoïdes dosés ---------- 109
Figure 70. Botrytis cinerea vu au microscope (a) et foyer de Botrytis sur une grappe (b) (images issues de la médiathèque de l'INRA) ---------- 110
Figure 71. Colletotrichum higginianum vu au microscope (a) et anthracnose sur une feuille de chou (image issue de University of Georgia Plant Pathology, University of Georgia, Bugwood.org) (b) -111
Figure 72. Fusarium graminearum vu au microscope (a) et divers niveaux d'attaque sur épis de maïs (b) (images issues de la médiathèque de l'INRA) ---------- 111
Figure 73. Phytophtora capsici vu au microscope (a) et manifestation de l'infection sur des concombres (b) (images issues de Hausbeck et Lamour, 2004) ---------- 111
Figure 74. Alternaria brassicicola vu au microscope (a) et illustration de la maladie de la tâche noire sur une feuille de colza (b) (images issues de la médiathèque de l'INRA) ---------- 111
Figure 75. Valeurs des densités optiques de l'eau et du méthanol 50% sur les souches Colletotrichum, Fusarium et Botrytis. ---------- 112
Figure 76. Photos du milieu au contact de l'antifongique, du méthanol 50% et de l'eau sur Fusarium ---------- 112
Figure 77 : Répartition des inhibitions mesurées pour Botrytis (a), Fusarium (b), Alternaria (c) et Phytophthora (d) ---------- 113
Figure 78. Relation espèce/activité pour les souches Botrytis (a), Fusarium (b), Alternaria (c) et Phytophthora (d) ---------- 114
Figure 79. Répartition des espèces actives sur au moins l'une des souches étudiées ---------- 115
Figure 80. Graphique Q-Q de l'analyse de normalité des résidus centrés réduits des données antifongiques ---------- 115
Figure 81. Moyenne des pourcentages d'inhibition avec standard de chaque espèce pour les quatre souches Fusarium, Botrytis, Alternaria et Phytophtora ---------- 117
Figure 82. ACP des V. sp Hearn à partir des données UHPLC-HRMS et coloration des plants actifs sur Fusarium ---------- 118
Figure 83. Modèle OPLS (a) validé par le test de permutation (b) et analyse des coefficients de la régression linéaire (c) avec emphase de 5 composés d'intérêt majoritaires chez V. sp Hearn (d) --119

Figure 84. ACP des données UHPLC-HRMS en mode PI des fractions SPE (a) et profils MS associés (b) ----120
Figure 85. Résultats des dosages BCA et flavonoïdes des fractions des extraits cyclotides ----122
Figure 86. Photos des tests antifongiques des extraits sur Botrytis (a) et moyennes des pourcentages d'inhibition avec déviation standard de l'antifongique (AF) et des différents extraits de violette (b) ----123
Figure 87. Photos des tests antifongiques des extraits sur Fusarium (a) et moyennes des pourcentages d'inhibition avec déviation standard de l'antifongique (AF) et des différents extraits de violette (b) ----125
Figure 88. Révélation des plantules par la solution de coloration GUS de la solution de Bion, du méthanol et de quelques extraits de violette ----126
Figure 89. Courbe de tendance des moyennes des intensités mesurées (a) et répartition selon la coloration des plantules (b) ----126
Figure 90. Relation espèce/activité pour l'étude de défense des plantes ----127
Figure 91. Graphique Q-Q des résidus centrés réduits des données logarithmées de coloration ----128
Figure 92. Moyennes des intensités de coloration avec intégration de l'erreur standard de chaque espèce ----128
Figure 93. Test de permutation de l'OPLS des données non loguées en pareto des données SDP ---129
Figure 94. Proportion des espèces impliquées dans les activités antifongique et de défense des plantes (a) et profils MS associés (b) ----130
Figure 95. Visualisation des cellules d'un puit comprenant uniquement du milieu MEM (a) et d'un puit avec de la doxorubicine à 0,5 µM (b) ----131
Figure 96. Courbe d'inhibition de la gamme de doxorubicine et équation de la courbe de tendance de la partie linéaire permettant le calcul de la CI_{50} ----131
Figure 97. Résultats moyens des tests de cytotoxicité pour chaque espèce et méthodes d'extraction ----132
Figure 98. Graphique Q-Q de l'analyse de normalité des résidus centrés réduits de l'ANOVA des données de cytotoxicité ----133
Figure 99. ACP (a) et OPLS (b) des données UHPLC-HRMS pour l'étude de la cytotoxicité ----134
Figure 100. Principales caractéristiques des méthodes d'extraction comparées ----135
Figure 101. ACP des profils chimiques des différents extraits de violettes (a) et CAH avec mise en évidence de deux principaux groupes (b) ----136
Figure 102. Rendements d'extraction des différentes méthodes employées ----137
Figure 103. Résultats des dosages UHPLC-UV et spectrophotométriques des polyphénols dont coumarines et flavonoïdes des différents extraits (a) et histogramme associé (b) ----138
Figure 104. CAH et ACP des échantillons issus de l'électroporation (a) et du CO_2 supercritique (b) avec sélection des extraits à traiter par « multivariate design » ----139
Figure 105. Résultats de la gamme d'étalonnage du Trolox dans le DMSO ----140
Figure 106. Valeurs moyennes des CI_{50} et intégration des signaux RPE pour chaque méthode d'extraction ----141
Figure 107. Système à reflux en série ----142
Figure 108. Rendements d'extraction aux différents ratios et temps d'extraction pour la violette de Toulouse et celle de Tourette ----142
Figure 109. CI_{50} par la méthode DPPH aux différents ratios et temps d'extraction pour la violette de Toulouse et celle de Tourette (a) et teneurs en polyphénols totaux par le dosage Folin-Ciocalteu (b) ----143
Figure 110. Comparaison des profils MS des violettes de Toulouse et de Tourette ----144
Figure 111. Explication schématique des différentes étapes de détection et regroupement des pics ----146
Figure 112 : Variabilité des réponses des détecteurs CAD et MS ----147
Figure 113. Courbes de calibration des signaux MS (a), des signaux UV (b), résultats d'intégration des signaux CAD (c) et droites de calibration associées (d) de la rutine et de l'umbelliferone ----148

Figure 114. Exemples de dégénérescence sur quelques standards en ESI du fraxetin (a), du tryptophane (b), de l'acide rosmarinique (c), de la phloridzine (d) et de l'ajmaline (e). ----------150
Figure 115. Sites ionisables sur de l'ajmaline (a) et du tryptophane (b) expliquant le phénomène d'ions multichargés ----------150
Figure 116. CCM révélée aux UV à 366 nm de différents extraits non traités (brut), traités par silice (Si) ou charbon actif (CA-quantité introduite) ----------152
Figure 117. Résultat de l'extraction biphasique et chromatogrammes MS associés aux deux phases ----------153
Figure 118. Signal CAD caractéristique d'un échantillon QC ----------154
Figure 119. Profils CAD caractéristiques d'une violette V. alba (a), V. suavis (b), V. odorata (c), V. subs Rostratae (d) et V. sp Hearn (e) ----------155
Figure 120. Teneurs en composés identifiés à partir des analyses semi-quantitatives par CAD-MS des parties aériennes des groupes génétiques de la collection ----------157
Figure 121. Schéma d'un montage pour analyse micrographique d'un matériel végétal ----------167
Figure 122. Principe de la méthode de Sanger ----------169
Figure 123. Cheminement de l'analyse métabolomique ----------171
Figure 124. Technique de piégeage pour l'analyse de la plante entière (a) et des fleurs coupées (b) ----------174
Figure 125. Construction du « Champ des odeurs »® ----------176
Figure 126 : Schéma du couplage GC/O (Fernandez et al., 2009) ----------177
Figure 127. Principe de l'extraction assistée par champ électrique ----------180
Figure 128. Caractéristique du signal enregistré ----------181
Figure 129. Droite de calibration UHPLC-UV de l'umbelliferone ----------188
Figure 130. Droite de calibration UHPLC-UV de la rutine ----------188
Figure 131. Spectres UV caractéristiques des coumarines (umbelliferone) et des flavonoïdes (rutine) ----------189
Figure 132. Droite de calibration de l'acide gallique ----------189
Figure 133. Droite de calibration de la rutine ----------190
Figure 134. Droite de calibration de la protéine BSA ----------190
Figure 135. Réduction du radical DPPH en présence d'un antioxydant donneur de proton ----------191
Figure 136. Exemple de plaque d'analyse DPPH d'extraits de violettes ----------192
Figure 137. Piégeage du radical superoxyde $O_2^{\bullet-}$ par le DMPO (a) et spectre caractéristique de l'adduit DMPO-OOH (b) ----------193
Figure 138. Signaux RPE du blanc solvant DMSO (a), du Trolox (b) et de la rutine (c) ----------193
Figure 139. Plaque et grille de dénombrement des cellules. La grille comporte 10 carrés subdivisés en 10 secteurs. ----------194
Figure 140. Clivage du X-Gluc par la β-glucuronidase conduisant à la production du chloro-bromoindigo se dimérisant en dichloro-dibromoindigo de couleur bleu ----------197

- Table des tableaux

Tableau 1. Positionnement taxonomique de la violette selon l'APG IV (2016) ---------- 27
Tableau 2. Nombre chromosomique d'espèces sauvages de violettes ---------- 29
Tableau 3. Ravageurs de la violette et leurs prédateurs ---------- 31
Tableau 4. Principaux composés volatils des extraits de fleur de violette (Saint-lary et al., 2015) 32
Tableau 5. Composés volatils majoritaires dans les absolues de feuilles de violette ---------- 33
Tableau 6. Classes et exemples de composés chimiques premièrement isolés chez Viola (DNP) ----- 34
Tableau 7. Résumé des activités biologiques retrouvées chez les violettes (Murray et al., 2015) -- 37
Tableau 8. Domaine d'étude des paramètres étudiés ---------- 53
Tableau 9. Expériences à réaliser et résultats obtenus : nombre de pics et intégration expérimentale (Yexp1 et Yexp2) et ceux prédits par le modèle quadratique (Ycalc1 et Ycalc2) ---- 54

Tableau 10. Test de Fischer dans le cadre de l'analyse sous cloche pour l'étude du nombre de pics (x/) et l'étude d'intégration des pics (/x) ---- 54
Tableau 11. Coefficient de l'équation de 2nd degré dans le cas de l'étude des pics (a) et de l'étude de l'intégration (b) ---- 54
Tableau 12. Nouveaux coefficients de l'équation dans le cas de l'étude des pics (a) et de l'étude de l'intégration (b) ---- 55
Tableau 13. Temps d'équilibre et d'exposition optimums pour l'analyse de la plante entière ---- 56
Tableau 14. Résultats des expériences pour l'optimum théorique sur plante entière ---- 56
Tableau 15. Résultats des expériences pour l'optimum technique sur plante entière ---- 56
Tableau 16. Résumé des composés dérépliqués à l'aide de MS-FINDER et du réseau moléculaire --- 68
Tableau 17 : Composés caractéristiques des deux chémotypes de V. odorata et V. sp Hearn ---- 70
Tableau 18. Résumé des synonymes et origine des espèces identifiées ---- 73
Tableau 19. Tableau des ressemblances-divergences des caractères morphologiques et anatomiques des six espèces de violettes majoritaires ---- 83
Tableau 20. Résumé des déplacements chimiques des trois composés identifiés en RMN ---- 97
Tableau 21. Composés identifiés et dérépliqués ayant un potentiel lien avec l'activité antioxydante ---- 98
Tableau 22. Violettes présentant un potentiel antioxydant d'intérêt ---- 103
Tableau 23. Classement des espèces sur la base du test de Tukey ---- 104
Tableau 24 Identifications des potentiels composés antioxydants déterminés par analyse statistique ---- 106
Tableau 25. Coumarines dosées dans les extraits de violette de la collection ---- 107
Tableau 26 : Flavonoïdes dosés dans les extraits de violette de la collection ---- 107
*Tableau 27. Résultats des activités phytosanitaires sur les quatre souches** ---- 113
Tableau 28. Résultats de l'analyse ANOVA pour chaque souche ---- 116
Tableau 29. Résultats des paramètres caractérisant le modèle OPLS des données antifongiques -- 118
Tableau 30 Identités des flavonoïdes et des cyclotides déterminés dans les fractions Fr2, Fr3 et Fr4 ---- 121
Tableau 31. Identification des violettes d'intérêt dans la recherche d'activité de défense ---- 127
Tableau 32. Résultats de normalisation et de l'ANOVA sur les données de cytotoxicité ---- 133
Tableau 33. Composés majoritaires des violettes de Toulouse et de Tourette ---- 144
Tableau 34. Relations mises en évidence par l'application du script MZ-unity ---- 151
Tableau 35. Relations mises en évidence par l'application du script MZ-unity ---- 154
Tableau 36 : Identification des pics UHPLC corrélés aux pics CAD ---- 158
Tableau 37. Fibres disponibles pour les analyses SPME ---- 174
Tableau 38 : Bases de données les plus utilisées pour l'identification en GC-MS ---- 177
Tableau 39. Conditions opératoires pour les extractions par électroporation ---- 181
Tableau 40. Constitution des solvants utilisés lors des fractionnements SPE sur cartouche C_{18} ---- 183
Tableau 41. Gradients de température mis en œuvre pour les analyses UHPLC-MS ---- 184
Tableau 42 : Gradients mis en œuvre dans la purification des coumarines et des flavonoïdes par UHPLC-UV ---- 185
Tableau 43. Principes des expériences de RMN bidimensionnelles ---- 186
Tableau 44. Plan de plaque pour les analyses DPPH ---- 191
Tableau 45 : Plan de plaque pour les analyses de cytotoxicité ---- 195

INTRODUCTION

La violette, très ancrée à l'image de Toulouse et de l'Occitanie, est un symbole de cette région au même titre que ses bâtiments aux briques roses ou encore son équipe de rugby. Ce lien particulier entre cette fleur et la ville rose démarre, d'après les dires, en 1850 lorsqu'un soldat de Napoléon III aurait rapporté une violette de Parme à sa bien-aimée résidant à Toulouse. À partir de cette époque, les Toulousains ont débuté la culture de cette fleur. Malgré sa fragilité et sa difficulté de culture, d'irréductibles producteurs et divers acteurs locaux, qu'ils soient industriels ou amateurs, essayent de continuer de faire vivre la violette de Toulouse, véritable image de marque de la ville. Cependant, le défi est difficile en raison des nombreuses contraintes liées à la culture de cette violette. Il s'ensuit que la majorité des produits y faisant usage proviennent principalement de la violette de Tourette cultivée plein champ à Tourrettes-sur-Loup.

Le projet multidisciplinaire « Viola Tolosa » financé par une bourse région avec un appui de la Mairie de Toulouse a pour principal but la valorisation industrielle d'extraits de violette de Toulouse, pouvant contribuer à la mise en place d'une filière Production-Extraction-Formulation. De nombreux acteurs locaux ont pris part à ce travail dont trois laboratoires universitaires, le LCA (Laboratoire de Chimie Agro-industrielle) porteur du projet, Pharma-Dev (Laboratoire de Pharmacochimie et Biologie pour le Développement) et le LRSV (Laboratoire de Recherche en Sciences Végétales) ; ainsi qu'un industriel, le Groupe Berdoues. En plus de la focalisation sur cette violette spécifique, la biodiversité de toute une collection de violettes enregistrée au conservatoire national de la violette a également été étudiée, en se basant en particulier sur l'analyse olfactive des composés volatils des fleurs ainsi que sur les propriétés biologiques des feuilles, dont la capacité antioxydante, pour une valorisation en dermo-cométique avec le développement d'une nouvelle gamme de produits. Des propriétés phytosanitaires ont été prospectées pour une éventuelle valorisation orientée vers des aspects de protection des cultures.

Un travail a donc été entrepris sur une centaine de cultivars de violette alliant les aspects phénotypiques, génétiques, phytochimiques et biologiques. Les méthodes adaptées à l'analyse d'un grand nombre d'échantillons ont été privilégiées et une stratégie basée sur l'utilisation de la métabolomique a été adoptée afin de cibler rapidement les extraits et composés d'intérêt, dans une matrice aussi complexe qu'un extrait végétal. Un volet industriel a été entrepris avec la valorisation d'un extrait brut de violette et la mise en place de premiers essais de formulation afin de rendre compte de la viabilité de la production de l'actif cosmétique recherché.

La thèse se décompose en trois grands chapitres détaillant l'état de l'art, les travaux accomplis et les perspectives. Une quatrième partie détaille les matériels et méthodes utilisés.

La **Partie I** est dédiée à l'état de l'art sur les violettes et porte notamment sur la taxonomie et la phylogénie de la plante, ses caractéristiques phytochimiques en particulier ses constituants volatils et non volatils, ses propriétés biologiques reconnues ainsi que son intérêt économique dans une optique de valorisation ;

La **Partie II** présente les résultats et est subdivisée en cinq parties. La première relate la classification de l'ensemble des violettes de la collection combinant aspects génétiques, phytochimiques et botaniques. La seconde se focalise plus particulièrement sur la capacité antioxydante de la violette de Toulouse avec l'isolement de trois molécules d'intérêt identifiées par RMN et rend compte des diverses optimisations de protocoles d'extraction effectuées sur cette dernière, avant application à l'ensemble de la collection pour l'étude des activités biologiques antioxydante et phytosanitaire dans un troisième temps. Une étude semi-quantitative des espèces de la collection a ensuite été proposée en quatrième partie en se basant sur la corrélation des données de plusieurs détecteurs assurant une spéciation des métabolites majoritaires des extraits. Une cinquième partie porte sur la valorisation industrielle d'un

extrait éthanolique de violette de Toulouse combinant extraction industrielle et premiers tests de formulation.

La **Partie III** présente les conclusions et perspectives de ce travail et clôture le corps de texte du manuscrit.

La **Partie IV** regroupe les matériels et méthodes de ce travail expliquant les principes de base de chaque technique mise en œuvre, notamment pour l'étude génétique, l'étude phytochimique des composés volatils et non volatils, l'étude du potentiel biologique des violettes, antioxydant et phytosanitaire, et le transfert industriel d'un extrait de violette. Elle clôture le corps de texte du manuscrit précédant les références et annexes.

PARTIE I
ETAT DE L'ART

Ce premier chapitre décrit les caractéristiques phénotypiques, chimiques et biologiques ainsi que leurs modes de culture et les intérêts économiques de la violette afin de rendre compte de l'intérêt industriel porté à cette espèce.

I.1 TAXONOMIE DE LA VIOLETTE

I.1.1 APPARTENANCE BOTANIQUE

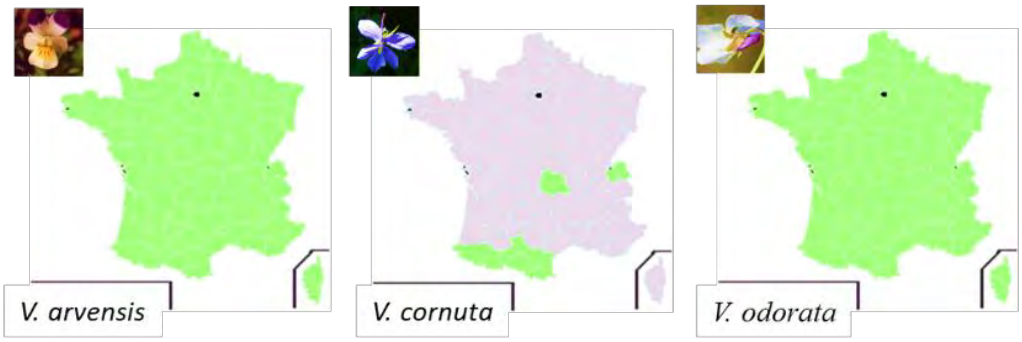
Les violettes appartiennent à la famille des Violacées (Violaceae), d'après la classification APG IV (**James et al., 2016**) (Tableau 1), qui comprend des arbres, des arbustes, des lianes tropicales et des plantes herbacées annuelles ou vivaces. Cette famille cosmopolite reste essentiellement répartie dans les régions tempérées et comprend 20 genres et environ 900 espèces (**Hekking, 1988**). La moitié de ces espèces appartient au genre *Viola* connu dans le bassin méditerranéen depuis l'Antiquité et qui rassemble environ 500 à 600 espèces réparties dans le monde entier, principalement dans l'hémisphère nord (**Ballard et al., 1999**).

Tableau 1. Positionnement taxonomique de la violette selon l'APG IV (2016)

Classification selon APG IV (2016)	
Groupe	Angiospermes
Groupe	Dicotylédones vraies
Groupe	Rosidées
Groupe	Fabidées
Ordre	Malpighiales
Famille	Violaceae
Genre	*Viola*

I.1.2 REPARTITION GEOGRAPHIQUE

Le genre *Viola* est réparti sur tout le territoire national avec une étendue plus ou moins importante selon les espèces. Par exemple, pour les pensées, *Viola arvensis* est répandue sur l'ensemble du territoire tandis que *Viola cornuta* est majoritairement présente sur les massifs pyrénéens. De même pour les violettes, *Viola odorata* est répandue sur l'ensemble du territoire tandis que *Viola pyrenaica* est présente dans les massifs montagneux pyrénéen et alpin et *Viola suavis* est plutôt spécifique du sud-est français (Figure 1).



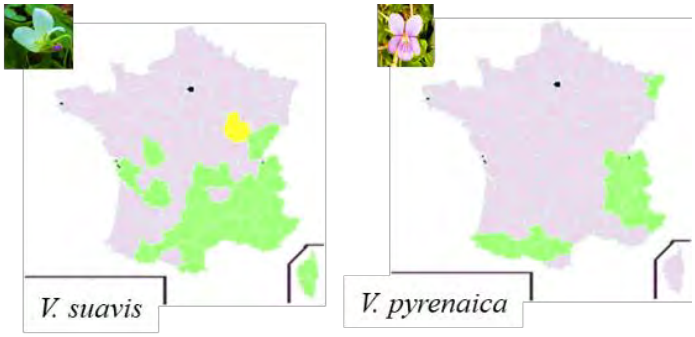


Figure 1. Répartition nationale de quelques espèces de pensées et de violettes d'après le réseau botanique francophone (Tela Botanica)

I.1.3 CARACTERISTIQUES MORPHOLOGIQUES

Les plantes appartenant au genre *Viola* possèdent une corolle à cinq pétales inégaux ainsi qu'un calice à cinq sépales inégaux. Ce genre peut être divisé en deux grands groupes selon le positionnement des pétales **(De Lamarck et De Candolle, 1805)** : les pensées qui présentent quatre pétales vers le haut et un vers le bas (Figure 2 a), et les violettes avec deux pétales vers le haut et trois vers le bas (Figure 2 b). Ce sont des plantes herbacées aux fleurs hivernales dont la floraison est précoce (**Willmer, 2011**), avec des feuilles rarement opposées et présentant de longues racines. Certaines espèces présentent des rejets, appelés stolons, assurant la multiplication végétative de la plante. Les violettes possèdent majoritairement des feuilles radicales réniformes et dentées sur les bords. Leurs fleurs sont solitaires et portées par un pédoncule comportant deux bractéoles.

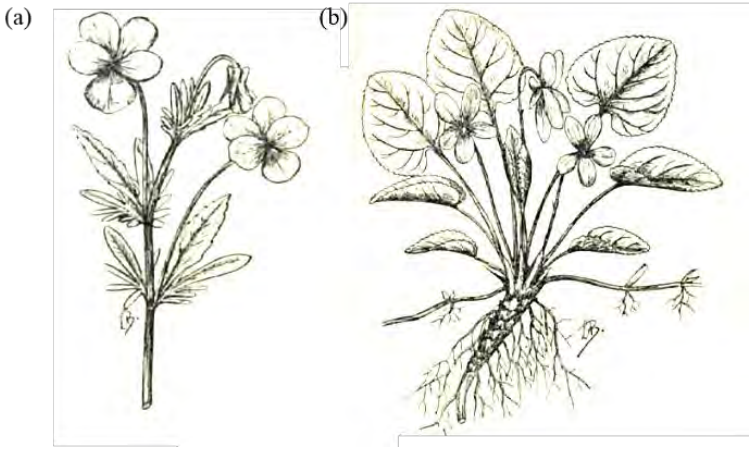


Figure 2. Caractéristiques morphologiques des pensées (a) et violettes (b) (images issues de (Beille, 1909))

La violette de Toulouse se distingue des autres violettes par des particularités morphologiques et anatomiques, dont ses fleurs chasmogames dites doubles constituées de plusieurs rangées de pétales. Cette particularité botanique s'explique par une transformation des étamines en pétales. La disparition des pièces fonctionnelles provoque alors la stérilité de l'appareil reproducteur.

Les *Viola* présentent différentes couleurs allant du blanc au violet en passant par le jaune et le rose. Ces différences de couleurs ont été étudiées chez *Viola x wittrockiana*. Il a été démontré que les fleurs rouges et violettes possèdent des molécules caractéristiques de la couleur, les anthocyanines et leurs dérivés glycosylés, en quantité importante. Les génines majoritaires (Figure 3) correspondent à la cyanidine, delphinine, petunidine et malvidine. En revanche, les fleurs jaunes sont moins riches en ces composés, justifiant ainsi leur couleur plus claire (**Skowyra et al., 2014**).

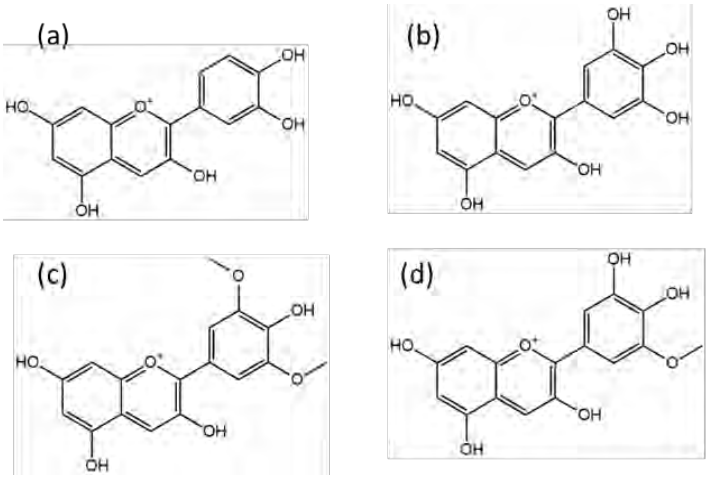


Figure 3. Structures chimiques des noyaux cyanidine (a), delphinidine (b), malvidine (c), et petunidine (d)

I.1.4 CARACTERISTIQUES CARYOLOGIQUES

L'appartenance taxonomique des cultivars de violette peut être obtenue par une analyse caryologique (étude du noyau cellulaire et de ses composants) : l'analyse de violettes sauvages par observation du nombre de paires de chromosomes (**Chauchard et al., 2013**) a permis de mettre en évidence qu'elles possédaient toutes soit 20, soit 40 chromosomes (Tableau 2).

Tableau 2. Nombre chromosomique d'espèces sauvages de violettes

2n = 20 chromosomes	2n = 40 chromosomes
Viola odorata	*Viola suavis*
Viola alba	*Viola ambigua*
Viola pyrenaica	
Viola hirta	

C'est une étude qui a été réalisée plus particulièrement sur deux cultivars très utilisés pour leur propriété olfactive (**Chauchard et al., 2013**): la violette de Tourette (aka *V. Victoria*) et la violette de Toulouse (aka *V. alba subs dehnhardtii*). Ces dernières possèdent respectivement 40 et 20 chromosomes comme observé sur le caryogramme ci-dessous (Figure 4).

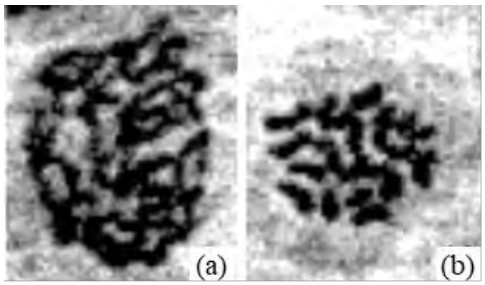


Figure 4. Caryogrammes d'un échantillon de violette de Tourette (a) et de Toulouse (b) d'après Chauchard et al.

Par comparaison, il a été déduit que les violettes de Tourette et de Toulouse appartiennent bien à deux taxons différents mais que contrairement à ce qui avait été décrit dans la littérature, la violette de Tourette se rapproche plus spécifiquement de *V. suavis* et la violette de Toulouse de *V. alba.*

I.2 PHYLOGENIE DES VIOLETTES

Quelques études phylogénétiques portant sur la famille des Violacées ont été réalisées et notamment sur le genre *Viola*, le plus étendu et constitué de près de 600 espèces réparties essentiellement dans l'hémisphère nord. Wihhelm Becker a été le premier taxonomiste à étudier en détails ce genre reconnaissant ainsi 14 sections dans les années 1920. D'autres spécialistes, dont Clausen (**Clausen, 1927**), ont ensuite apporté quelques révisions à la classification de leur prédécesseur (**Ballard et al., 1999**). Ces investigations reposent alors essentiellement sur les aspects phénotypiques et caryologiques.

La majorité des études plus récentes reposent sur la phylogénie moléculaire (ADN, ARN, protéines) et se basent sur la réaction de polymérisation en chaîne (PCR) (**Watson et al. 1994**), permettant de multiplier une région spécifique de l'ADN à l'aide d'amorces universelles (**White et al., 1990** ; **Malécot, et al., 2007**), assurant ainsi une meilleure lecture du code génétique (cf. bases théoriques en annexe 1, A, page 211). Le séquençage des acides nucléiques permet une estimation sensible des ressemblances et divergences entre espèces, puisqu'il s'appuie sur l'analyse de seulement quatre bases : adénine (A), thymine (T), cytosine (C) et guanine (G). L'enchainement de ces bases constitue le code génétique qui est propre à chaque individu. Les espaces internes transcrits (ITS) formés des ITS 1 et 2 séparés par la séquence 5,8S (Figure 5) forment la partie non codante l'ADN ribosomique. Ce sont des domaines d'intérêt pour les études phylogénétiques en raison de leur fort polymorphisme qui permet de mettre facilement en évidence des différences génétiques entre les espèces et les espèces proches grâce à la spécificité des séquences nucléotidiques **(Baldwin et al., 1995**).

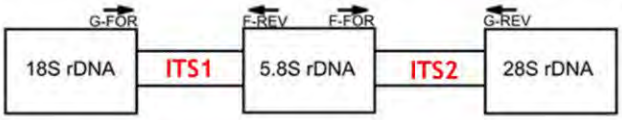


Figure 5. Schématisation des espaces internes transcrits

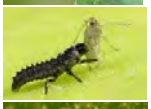
Ces études ont alors permis de confirmer certaines classifications obtenues par les méthodes traditionnelles et d'en proposer de nouvelles. Les analyses ont montré que la région ITS 1 est plus longue que celle d'ITS 2, avec une variabilité plus forte. L'ADN chloroplastique a également été utilisé en tant qu'outil complémentaire, en particulier pour détecter les lignées maternelles (**Cennamo et al., 2011**). Ceci a été le cas pour une violette identifiée comme *V. suavis* mais dont la position dans l'arbre phylogénétique des ITS et des chloroplastes était différente. Dans le premier cas, l'échantillon était lié au groupe des *V. hirta* tandis que dans le second il appartenait bien au groupe des *V. suavis*. Une analyse détaillée de la séquence ITS 1 a alors mis en relief qu'une partie était homologue à *V. hirta* suggérant une recombinaison entre les deux espèces. La comparaison des résultats des deux méthodes semble ainsi être un bon moyen de vérifier la présence d'éventuels hybrides : si les deux résultats sont identiques, l'échantillon appartient très certainement au taxon en question, mais si les résultats divergent, alors l'échantillon peut être la résultante d'une hybridation.

I.3 CULTURE ET REPRODUCTION

La culture traditionnelle de la violette est pratiquée en plein air et en pleine terre. Plusieurs sols sont retrouvés dans les zones de production de la violette : elle s'accommode relativement bien de sols peu profonds qui drainent rapidement l'excédent d'eau et se réchauffent vite. Mais à partir de 1975, la culture sous abri fait son apparition, suivie de la culture hors-sol qui permet d'éviter les contaminations systématiques par des champignons par exemple, améliorant ainsi l'état sanitaire et la productivité.

C'est une fleur relativement fragile, car sensible aux virus, tels que le virus du Haricot, du Concombre ou encore le Viola mottle virus transmis par piqures d'insectes ou bouturage, et les champignons qui entrainent une baisse des rendements. C'est pourquoi, dans les années 1990, la multiplication *in vitro* a été mise au point dans le but d'obtenir des plants sains, mais cette méthode reste limitée, car il a été observé une baisse de floraison et une réinfection des cultures chez les producteurs (**Grasse et al., 2009a**). Plusieurs ravageurs menacent la culture de la violette. Aujourd'hui, une lutte biologique naturelle, reposant sur le principe que tout ravageur a un prédateur naturel, est privilégiée par rapport à la lutte chimique (Tableau 3).

Tableau 3. Ravageurs de la violette et leurs prédateurs

Ravageurs		Prédateurs	
Acarien phytophage		Larve de *Feltiella* (cécidomyie) Phytoseïulus Persimilis (acarien)	
Aleurode		Macrolophus (punaise)	
Puceron		Coccinelle Larves d'Aphidoletes (cécidomyie)	
Thrips		*Amblyseius Cucumeris* (acarien)	

I.4 PHYTOCHIMIE DES VIOLETTES

Jusqu'à aujourd'hui, environ 200 composés chimiques de classes variées ont été identifiés pour le genre *Viola* (**Zhu et al., 2015**). Parmi les composés volatils, les ionones sont majoritaires et apportent l'odeur caractéristique de la violette ; pour les composés non volatils, les flavonoïdes et les cyclotides sont majoritairement retrouvés.

I.4.1 COMPOSES VOLATILS DE LA VIOLETTE

Les premières études analytiques sur les composés volatils, destinées à caractériser l'odeur de la violette, ont été réalisées à partir de l'huile essentielle ou d'extraits obtenus à l'aide de solvants organiques. Plus récemment, ce sont des techniques d'adsorption s'appuyant sur l'espace de tête qui ont été utilisées pour compléter les premières investigations. Elles sont adaptées à l'analyse des plantes fraîches et sont connues pour ne pas altérer la composition intrinsèque, permettant ainsi de retrouver une tonalité très proche de celle obtenue lorsque l'on sent la fleur.

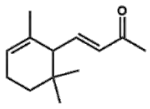
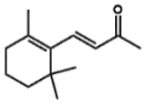
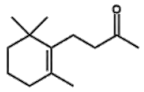
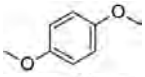
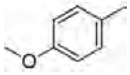
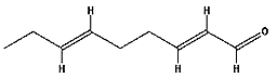
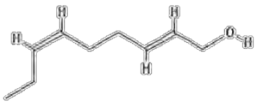
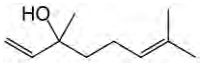
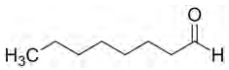
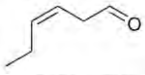
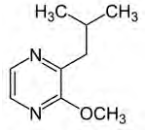
Les composés volatils sont des composés de différentes classes chimiques appartenant aux métabolismes primaires et secondaires (**Goff, 2006**). De manière générale, ce sont des composés de faibles masses moléculaires, de faibles températures d'ébullition et plutôt lipophiliques (**Pichersky, 2006**). Ils assurent la communication entre la plante et son environnement, par exemple dans une optique de défense et de reproduction. Jusqu'à aujourd'hui, plus de 1700 composés volatils provenant de plus de 90 familles de plantes ont été identifiés (**Dudareva et al., 2013**). La principale méthode analytique utilisée dans l'étude des composés volatils est la chromatographie gazeuse associée à un spectromètre de masse (GC-MS) (**Tholl et al., 2006**).

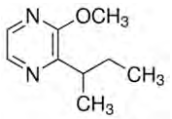
L'analyse des profils volatils des plantes reste une méthode pertinente par sa simplicité et l'apport considérable de données, qui permettent de différencier les espèces entre elles ainsi que les individus au sein d'une même espèce.

I.4.1.1 Les fleurs

La composition des substances volatiles, et plus particulièrement des substances odorantes caractéristiques de l'odeur de violette, a été mise en évidence au fur et à mesure de l'évolution des techniques analytiques (**Bedoukian, 1982**). Cependant, dès les années 1930, les substances les plus caractéristiques de cette odeur ont été révélées à savoir : l'α-ionone, un sesquiterpène isolé par Ruzicka sous le nom de parmone et finalement identifiée par Ohloff (**Uhde et Ohloff, 1972**), et les (2E,6Z)-nonadienal et (2E,6Z)-nonadienol. La présence d'un aldéhyde avait été mise en évidence par Soden en 1925 mais c'est finalement Ruzicka et ses collègues qui ont permis d'identifier, dans la fleur et la feuille, l'aldéhyde et son alcool (**Ruzicka et al., 1944**). Des études par espace de tête couplées à une détection olfactive, réalisées par Brünke et ses collègues, ont ensuite permis de confirmer et de compléter les principales substances caractéristiques de l'odeur de la violette (**Grasse et al., 2009b**) (Tableau 4).

Tableau 4. Principaux composés volatils des extraits de fleur de violette (Saint-lary et al., 2015)

Molécule	Formule brute	Masse moléculaire (g/mol)	Structure chimique	Notes olfactives**
*α-ionone	$C_{13}H_{20}O$	192,151		Florale Violette
*β-ionone	$C_{13}H_{20}O$	192,151		Florale Boisée
*Dihydro-β-ionone	$C_{13}H_{22}O$	194,167		Légèrement boisée Florale
*p- diméthoxybenzène	$C_8H_{10}O_2$	138,068		Verte
*4-méthylanisole	$C_8H_{10}O$	122,073		Florale
(2E,6Z)-nona-2,6-diénal	$C_9H_{14}O$	138,210		Verte Concombre Melon
(2E,6Z)-nona-2,6-diénol	$C_9H_{16}O$	140,120		Concombre Melon Plus fruité Moins vert
Linalool	$C_{10}H_{18}O$	154,136		NC
Octanal/nonanal/décanal	$C_8H_{16}O$ $C_9H_{18}O$ $C_{10}H_{20}O$	128,120 142,136 156,151		Grasse Fruitée (citrus)
(3Z)-hexénal	$C_6H_{10}O$	98,073		NC
3-isopropyl-2-méthoxy pyrazine	$C_8H_{12}N_2O$	152,095		NC

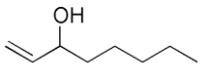
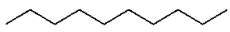
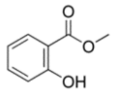
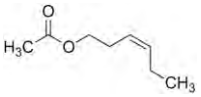
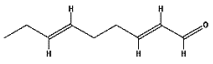
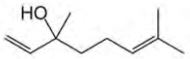
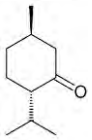
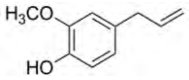
3-sec-butyl-2-méthoxy pyrazine	$C_9H_{14}N_2O$	166,111		NC

* Composés majoritaires ; NC : Non communiqué, ** Information issue du site The Good Scents Company (http://www.thegoodscentscompany.com)

I.4.1.2 Les feuilles

Les absolues de feuilles de violette sont également riches en composés volatils, dont une infime partie est commune avec les fleurs. Des analyses par espace de tête sur la violette de Tourette ont permis d'identifier 49 composés dont les majoritaires détectés par une étude de violettes française et égyptienne (**Saint-lary et al., 2015**) sont présentés ci-après (Tableau 5).

Tableau 5. Composés volatils majoritaires dans les absolues de feuilles de violette

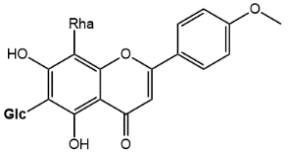
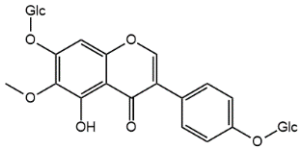
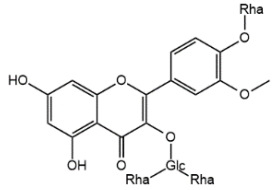
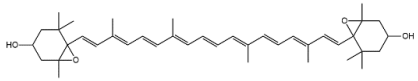
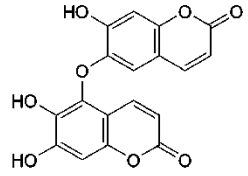
Molécule	Formule moléculaire	Masse moléculaire (g/mol)	Structure chimique	Note olfactive*
Oct-1-èn-3-ol	$C_8H_{16}O$	128,120		Terreuse Champignon
n-décane	$C_{10}H_{22}$	142,172		NC
Undécane	$C_{11}H_{24}$	156,188		NC
Dodécane	$C_{12}H_{26}$	170,203		NC
Salicylate de méthyle	$C_8H_8O_3$	152,047		Epicée Aromatique
Acétate de (Z)-hex-3-ényle	$C_8H_{14}O_2$	142,099		Fruitée Verte Fraiche
(E,Z)-nona-2,6-diénal	$C_9H_{14}O$	138,210		Verte Concombre
Eucalyptol	$C_{10}H_{18}O$	154,136		Verte Mentholée
Linalool	$C_{10}H_{18}O$	154,136		Florale Citronnée
Menthone	$C_{10}H_{18}O$	154,136		Mentholée Verte Fraiche
Eugénol	$C_{10}H_{12}O_2$	164,084		Epicée

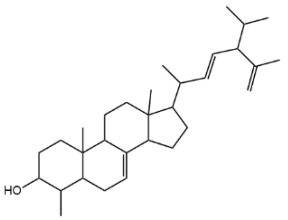
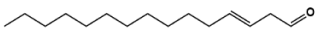
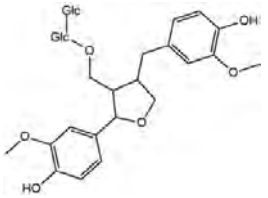
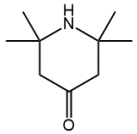
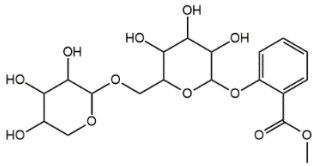
NC : Non communiqué ; * Information issue du site The Good Scents Company (http://www.thegoodscentscompany.com)

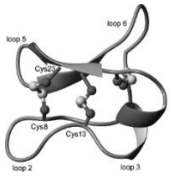
I.4.2 COMPOSES NON VOLATILS DE LA VIOLETTE

Le Tableau 6 résume les différentes classes de composés qui ont pu être retrouvées dans les extraits de *Viola*, avec des exemples de composés découverts chez ce genre, en particulier chez la violette, notre sujet d'étude.

Tableau 6. Classes et exemples de composés chimiques premièrement isolés chez *Viola* (DNP)

	Classes	Composés	Formule brute	Masse moléculaire (g/mol)	Structure chimique	Source
FLAVONOÏDES	Flavones	4'-O-Methylviolanthin	$C_{28}H_{32}O_{14}$	592,179		*V. etrusca*
	Isoflavones	Tectorigenin 4',7-diglucoside	$C_{28}H_{32}O_{16}$	624,169		*V. hondoensis* (**Muhammad et al., 2012**)
	Flavonols	Violetruscoside	$C_{40}H_{52}O_{24}$	916,285		*V. etrusca* (**Flamini, 2007**)
TERPENOÏDES	Tétraterpénoïdes	Di-cis-violaxanthin D	$C_{40}H_{56}O_4$	600,418		*V. tricolor* (**Molnár et Szabolcs, 1980**)
BENZOPYRANOÏDES	Coumarines	Dimeresculetine	$C_{18}H_{10}O_8$	354,038		*V. yedoensis* (**Zhou et al., 2009**)

STEROÏDES		Violasterol A	$C_{31}H_{50}O$	438,386		*V. formosana* (**Lee et al., 1993**)
ALIPHATIQUE		3-Pentadecenal	$C_{15}H_{28}O$	224,214		*V. odorata* (**Cu et al., 1992**)
LIGNANS		Tianshanoside A	$C_{32}H_{44}O_{16}$	684,263		*V. tianshanica* (**Qin et al., 2013**)
ALCALOIDES	Dérivé piperidine	Vincubine	$C_9H_{17}NO$	155,131		V. *yedoensis* (**Rodrigues et al., 2007**)
AROMATIQUES SIMPLES		Violutoside	$C_{19}H_{26}O_{12}$	446,142		*V. cornuta* (**Kanchanapoom, 2007**)

PEPTIDES	Cyclotides	Cycloviolacine O2	$C_{133}H_{213}N_{37}O_{39}S$	2984,554		*V. odorata* **(Irelandet al., 2006)** *V. hederaceae* **(Chen et al., 2005)** *V. yedoensis* **(Wang et al., 2007)**
		Cycloviolacine H3	$C_{129}H_{190}N_{36}O_{40}S_6$	3075,226		
		Cycloviolacine Y1	$C_{145}H_{195}N_{35}O_{48}S_6$	3386,222		

D'après la base de données du Dictionnaire des Produits Naturels (DNP), la classe des peptides est majoritaire chez les *Viola*, suivi des flavonoïdes, molécules ubiquitaires des plantes avec une forte diversité structurale. Les autres classes chimiques représentent des pourcentages inférieurs à 10% (Figure 6). Ces pourcentages correspondent aux molécules premièrement isolées chez les *Viola*. La présence d'autres dérivés non répertoriés en raison de leur premier isolement chez d'autres genres paraît évidente par exemple, au niveau de la catégorie des flavonoïdes, d'autres types de flavonoïdes *O*- et *C*-glycosylés, tels que des dérivés d'apigenine ou encore de kaempferol.

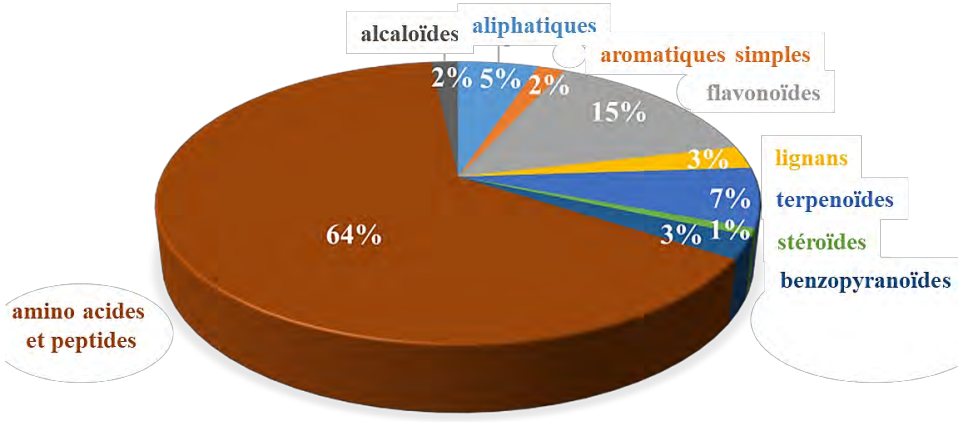


Figure 6. Répartition statistique des composés isolés chez le genre *Viola*

I.5 PROPRIETES BIOLOGIQUES DES *VIOLA*

Les violettes et certaines de leurs propriétés étaient connues dès l'Antiquité où on faisait des couronnes de fleurs pour combattre l'ivresse. C'est au Moyen-âge que les effets physiologiques de la violette ont été mis en lumière : à l'état frais, elles permettaient de favoriser le sommeil ou encore de guérir les angines, tandis qu'à l'état sec elles étaient utilisées pour les obstructions du foie et les inflammations. Différents usages ethnobotaniques ont été relatés dont des propriétés diurétiques (**Ahmad et al., 2009**), antipyrétiques et expectorantes. Les violettes étaient souvent utilisées dans le cas d'épilepsie ou de troubles nerveux mais aussi lors d'éruptions cutanées telles que l'eczéma (**Bhatt et Negi, 2006**).

Aujourd'hui, les activités biologiques autour de la violette sont relativement variées et sont présentées dans le Tableau 7. Jusqu'à ce jour, seule une trentaine d'espèces sur les 500 estimées ont été étudiées.

Tableau 7. Résumé des activités biologiques retrouvées chez les violettes (Murray et al., 2015)

Activité biologique	Extrait	Espèce	Organe végétal	Composés identifiés
Anti-inflammatoire	Aqueux Alcoolique	*V. odorata,* *V. diffusa,* *V. yedoensis*	Plante entière	Flavonoïdes **(Jeong et al., 2016)**
Immunomodulation	Aqueux Alcoolique	*V. yedoensis* *V. diffusa* *V. odorata*	Plante entière	Flavonoïdes Tannins **(Ibrahim, 2017)**
Antimalariale	Ether de pétrole	*V. verecunda* *V. canescens*	Plante entière	Dérivés acide salicylique Triterpenoïde **(Moon et al., 2007)**
Anti-cancer	Organique Aqueux	*V. odorata* *V. yedoensis*	Parties aériennes	β-ionone **(Ansari et Emami, 2016)** Cyclotides **(Gerlach et al., 2010)**
Anti-HIV	Organique Alcoolique	*V. yedoensis* *V. hederacea*	Parties aériennes	Cyclotides **(Wang et al., 2007)**
Insecticide	Organique Aqueux	*V. odorata*	Parties aériennes	Cyclotides **(Ireland et al., 2006)**
Antifongique	Huile essentielle	*V. odorata*	Fleur	NC **(Hammami et al., 2011)**
Antimicrobienne	Aqueux Alcoolique	*V. yedoensis* *V. diffusa* *V. thianschanica*	Parties aériennes	Coumarines **(Sun et al., 2011)** Alcaloïdes Flavonoïdes **(Shen et Xie, 2009)**
Antioxydante	Alcoolique	*V. odorata* *V. philippica* *V. diffusa* *V. yedoensis* *V. hondoensis*	Parties aériennes	Composés phénoliques Flavonoïdes **(Cao et al., 2013)**

NC : Non communiqué

La violette semble donc présenter de nombreuses propriétés biologiques pour la plupart en lien avec sa composition en polyphénols et cyclotides. Cependant, la majorité de ces activités restent aujourd'hui annexes, notamment les propriétés anti-HIV ou antimalariale par exemple, qui couvrent des secteurs suplantés par des drogues plus efficaces comme les quinolines et des plantes plus prometteuses. En revanche, les propriétés insecticide et antioxydante, qui peuvent respectivement intéresser les domaines phytosanitaire et cosmétique, sont les activités majoritairement étudiées, comme on le retrouve notamment en termes de publication. Néanmoins, aucun principe actif commercialisé n'a encore été proposé ouvrant donc la porte à une possible valorisation industrielle. Ces deux activités d'intérêt dans ce travail ont été développées ci-dessous en mettant en valeur deux catégories de composés, les cyclotides et les polyphénols.

I.5.1 PROPRIETES ANTIOXYDANTES

I.5.1.1 Stress oxydant

Un stress oxydant ou oxydatif est un état de déséquilibre entre la production d'espèces réactives et les défenses de l'organisme. Ces molécules sont principalement des espèces réactives de l'oxygène (ERO) incluant les radicaux libres oxygénés tels que l'oxygène singulet O, le peroxyde d'hydrogène H_2O_2, le radical superoxyde $O_2^{\bullet-}$ ou encore le radical hydroxyle $HO^{\bullet}$ (**Sies, 1985**). Toutes ces espèces peuvent présenter des effets néfastes pour la santé, engendrer des lésions et endommager les cellules. Elles sont quotidiennement produites par notre organisme au cours de réactions enzymatiques, par exemple sous l'effet des ultraviolets, mais sont maîtrisées par des systèmes de défense. A dose raisonnable, ces molécules sont utiles à l'organisme.

L'activité oxydante, favorisée par les pro-oxydants, est définie par une réaction avec une molécule d'oxygène conduisant à la perte d'un électron et la formation de radicaux. L'activité antioxydante est caractérisée par la capacité à piéger des radicaux libres, par l'apport d'un atome d'hydrogène ou d'un électron, en stabilisant les espèces formées. Dans l'organisme ou la cellule, la balance entre antioxydants/pro-oxydants peut parfois être déséquilibrée, par exemple lors d'une défaillance nutritionnelle engendrant une carence de l'approvisionnement en antioxydants (vitamines, oligo-éléments, …), ou encore d'anomalies génétiques provoquant un mauvais codage d'une enzyme synthétisant une molécule antioxydante (**Favier, 2003**). Le stress oxydant peut avoir diverses origines, aussi bien exogènes qu'endogènes (**Haleng et al., 2007**), et est mis en cause dans certaines maladies chroniques ou liées à l'âge par exemple les maladies cardiovasculaires, la dégénération neurologique type Alzheimer (**Kanti Das et al., 2014**) ainsi que la vieillesse et la détérioration de la peau. La théorie que le vieillissement pouvait être expliqué par l'action des radicaux libres a d'ailleurs été initialement développée par le Dr Denham Harman (**Harman, 1992**).

Pour se protéger des effets nocifs des radicaux libres et ERO, différentes défenses sont mises en œuvre au sein de l'organisme. Une première partie est endogène et se rapporte à l'activité de certaines enzymes et l'autre partie est apportée par l'alimentation, notamment par l'intermédiaire des fruits et légumes riches en polyphénols.

I.5.1.2 Antioxydants naturels

I.5.1.2.1 Défenses enzymatiques

Les défenses enzymatiques majoritaires de l'organisme reposent sur l'action des superoxydes dismutases qui sont des métalloprotéines assurant l'élimination du radical superoxyde par

l'intermédiaire d'une réaction de dismutation, le transformant alors en peroxyde d'hydrogène et en oxygène ; des glutathion peroxydases qui sont de sélénoprotéines réduisant principalement les peroxydes lipidiques formés par l'application d'un stress oxydant sur les acides gras polyinsaturés, et dont l'action est conditionnée par un apport adéquat en sélénium (céréales complètes, poissons, fruits de mer) ; ainsi que du système thiorédoxine qui intervient essentiellement dans la dégradation des peroxydases lipidiques ainsi que du peroxyde d'hydrogène (**Haleng et al., 2007**).

I.5.1.2.2 Défenses non enzymatiques

I.5.1.2.2.1 Vitamines C et E

Les vitamines C et E sont de très bons antioxydants. La vitamine C non synthétisée par l'Homme doit être apportée par une consommation journalière en fruits. C'est un excellent piégeur des radicaux hydroxyle et superoxyde. Elle assure également une action sur la péroxydase lipidique par régénération de la vitamine E, par réaction avec des radicaux lipidiques. Pour la vitamine E, antioxydant liposoluble, deux isomères sont particulièrement pertinents, l'α- et le γ-tocophérol qui luttent contre la propagation de la péroxydation lipidique en réagissant avec les radicaux peroxyles ($ROO^{\bullet}$). Le premier est particulièrement présent dans l'huile de sésame tandis que le second sera plutôt retrouvé dans les huiles, les noix et les noisettes (**Haleng et al., 2007**).

I.5.1.2.2.2 Les oligoéléments

Les oligoéléments ne sont pas des antioxydants mais jouent essentiellement un rôle de cofacteur des enzymes. Ils sont majoritairement au nombre de trois (**Haleng et al., 2007**):

- le sélénium, cofacteur de la glutathion peroxydase puisque sa métabolisation hépatique conduit à la formation d'intermédiaires essentiels à sa synthèse,
- le cuivre, cofacteur de la superoxyde dismutase. Une concentration trop élevée conduira à des effets inverses puisqu'il jouera alors le rôle d'un pro-oxydant.
- le zinc, cofacteur de diverses enzymes dont la superoxyde dismutase.

I.5.1.2.2.3 Les caroténoïdes

Les caroténoïdes sont capables de terminer les réactions en chaîne de la lipoperoxydation. Néanmoins, le caroténoïde jouant un rôle majoritaire dans le combat du stress oxydant reste le β-carotène, aussi nommé provitamine A. En effet, il est précurseur de la vitamine A et est capable de capter à faible pression l'oxygène singulet (**Goudable et Favier, 1997**).

I.5.1.2.2.4 Les polyphénols

Les polyphénols sont fréquemment cités dans la littérature pour leur propriété antioxydante. Ils peuvent agir en tant que chélateurs d'ions métalliques, tels que ceux du fer, du cuivre et de l'aluminium, impliqués dans la conversion de l'ion superoxyde et de l'eau oxygénée en radical hydroxyle, par les réactions de Haber-Weiss et Fenton comme expliqué Figure 7 (**Liochev et Fridovich, 2002**).

(1) $Fe(II) + H_2O_2 \rightarrow Fe(III) + HO^{\bullet} + OH^-$ (réaction de **Fenton**)

(2) $HO^{\bullet} + H_2O_2 \rightarrow H_2O + O_2^{\bullet-} + H^+$

(3) $Fe(III) + O_2^{\bullet-} \rightarrow Fe(II) + O_2$

(4 = 1 + 3) $O_2^{\bullet-} + H_2O_2 \xrightarrow{Fe} O_2 + HO^{\bullet} + OH^-$ (réaction de **Haber-Weiss**)

Figure 7. Réduction de H_2O_2 par Fe^{2+} par la réaction de Fenton (a) et réactions associées (2, 3). Formation du radical hydroxyle par réaction de Haber-Weiss (4)

Cependant, c'est par leur action directe sur les ERO et les radicaux que les polyphénols expriment le mieux leur propriété antioxydante selon deux mécanismes. La première réaction est basée sur la capacité des polyphénols (ArOH) de céder un atome d'hydrogène à un radical libre R•, se transformant alors lui-même en radical (ArO•). Cette réaction est dépendante de la stabilité de ce radical, notamment liée à la présence d'autres fonctions hydroxyles sur la structure impliquées dans la formation de ponts hydrogènes intramoléculaires et dans le phénomène de délocalisation électronique. Le deuxième mécanisme implique le transfert d'un électron vers le radical R• conduisant à la formation d'un anion R- et d'un radical cation $ArOH^{\bullet+}$. Deux principaux paramètres caractérisent ces réponses antioxydantes : l'énergie de dissociation de la liaison OH dans le premier cas et le potentiel d'ionisation de la molécule dans le second (**Quideau et al., 2011**). Les polyphénols présentant au niveau de leur structure la capacité de former des ponts hydrogènes, comme ceux avec des noyaux catéchol ou pyrogallol tel que l'epicatechine, réagiront sur le premier modèle de transfert d'hydrogène ; tandis que ceux dont cette faculté est limitée mais qui possèdent une forte délocalisation électronique, comme le kaempferol, réagiront plutôt sur le modèle de transfert d'électrons (**Leopoldini et al., 2004**).

Cependant, les polyphénols possédant des noyaux catéchols et/ou pyrogallol peuvent présenter une action pro-oxydante dans certaines conditions dont en forte concentration (**Wätjen et al., 2005**) et en présence de métaux . En effet, le radical formé peut réagir avec un autre radical et conduire à la formation d'une *o*-quinone (Figure 8). En outre, il peut réagir avec l'oxygène et conduire à la formation d'un anion superoxyde. Ils peuvent également réduire les ions métalliques Cu^{2+} et Fe^{3+} et favoriser la réaction de Fenton pour conduire à la production du radical hydroxyle.

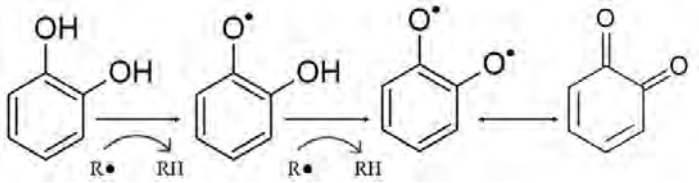


Figure 8 : Réaction de piégeage des ERO (R•) par un noyau catéchol avec formation d'une *o*-quinone

L'activité antioxydante des violettes est non seulement retrouvée au niveau des feuilles mais aussi des fleurs qui sont riches en composés phénoliques (**Skowyra et al., 2014**) dont les anthocyanines responsables de la couleur des fleurs. Ces molécules font partie de la classe des flavonoïdes types flavones, flavonols, flavanones et flavanols (**Malien-Aubert et al., 2001**). L'activité antioxydante est due à la présence de ces composés phénoliques, comme la rutine par exemple (**Vukics et al., 2008**). Trois principales structures conférant cette activité antioxydante ont été mises en évidence : le noyau catéchol (aromatique B) relié à la structure flavonoïde (Figure 9 a) entrainant une délocalisation des électrons et conférant donc une stabilité au radical formé ; la double liaison 2,3 conjuguée avec la fonction oxo en position 4 (Figure 9 b) qui entraine la délocalisation au niveau de l'aromatique B ; et la

présence de deux hydroxyles en position 3 et 5 (Figure 9 c) formant des ponts hydrogènes avec la fonction cétone en position 4 (**Bors, 1990**).

Figure 9. Groupes structuraux d'intérêt dans le potentiel antioxydant des flavonoïdes

Le principal but de ce travail est alors de trouver des antioxydants applicables en cosmétique et pouvant jouer le rôle de conservateur, tout en présentant des propriétés sur la peau.

I.5.2 PROPRIETES PHYTOSANITAIRES

Diverses maladies affectent les plantes par l'action de microorganismes tels que des virus, des bactéries, ou encore des champignons. Le contrôle de ces champignons, responsables des maladies les plus dommageables sur les plantes, est assuré en partie par des produits phytosanitaires chimiques (fongicides) agissant alors directement sur des fonctions essentielles des microorganismes, comme la respiration ou la division cellulaire. Cependant, cette lutte chimique est nettement entachée d'un caractère de toxicité et d'écotoxicité. L'alternative aujourd'hui est d'exploiter des biopesticides (**Hanem, 2012**) issus d'organismes vivants, de substances d'origine naturelle (**Villaverde et al., 2016**) ou encore d'exploiter des réactions de défense de la plante (réponse immunitaire) (**Burketova et al., 2015**), par l'intermédiaire de substances naturelles. Les bioessais utilisés reposent alors sur le criblage de champignons ainsi que sur l'expression du promoteur Pr1 (pour Pathogenesis related 1) couplé au gène GUS, bon marqueur des réponses de défense (**Ohashi et Ohshima, 1992**)

Les activités phytosanitaires de la violette semblent particulièrement liées aux cyclotides qui présentent des propriétés biologiques variées. Ce sont des peptides, comprenant généralement 30 acides aminés présentant un arrangement noué incluant six motifs cystéine avec trois ponts disulfures (Figure 10) dont le motif structural est appelé « cyclic cystine knot ».

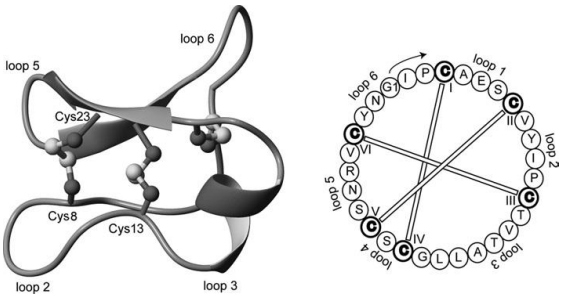


Figure 10 : Exemple de structure tridimensionnelle et de séquence d'un cyclotide (Ireland et al., 2006)

Aujourd'hui, plus de 300 cyclotides ont été décrits et classés en deux familles, les Möbius et les Bracelet respectivement caractérisées par la présence ou l'absence d'une cis proline. Ils sont retrouvés chez les Violacées, les Rubiacées et les Cucurbitacées.

Ils sont distribués dans différents organes des plantes dont les feuilles, les fleurs et les racines (**Trabi, 2004**). La première caractérisation structurale a été proposée par Schöpke et ses collaborateurs en 1993 à partir du violapeptide 1 isolé chez *Viola arvensis* (Violacées). Cette description a été complétée à partir du cyclotide kalata B1 isolé chez *Oldenlandia affinis* (Rubiacées), agent utérotonique utilisé en médecine traditionnelle africaine pour accélérer la naissance (**Gran et al., 2009**), en proposant une

structure tridimensionnelle. Ce sont des molécules très stables, en raison des contraintes structurales apportées par les ponts disulfures et la cyclisation, qui possèdent un large éventail de propriétés biologiques dont anti-HIV (**Chen et al., 2005**) due à leur efficacité contre les cellules infectées par un virus avec une concentration efficace médiane (CE_{50}) de l'ordre de 100 nM (**Gustafson et al., 2004**) et liée à leur structure cyclique. Ils présentent une activité antimicrobienne (**Pranting et al., 2010**) ou encore insecticide (**Gruber et al., 2007**). Ainsi, ils sont produits par les plantes pour leur rôle de défense contre les insectes, puisqu'ils semblent inhiber la croissance et le développement de certaines larves (**Jennings et al., 2001**). Ils présentent également une activité hémolytique (destruction des globules rouges), premièrement déterminée pour le cyclotide Viola peptide 1 et apportée par la structure cyclique (**Barry et al., 2003**), en lien avec leur rôle de défense. Ils offrent alors d'intéressantes perspectives en tant que biopesticides.

Malgré la disparité de ces activités, un mécanisme d'action commun semble prépondérant, impliquant l'interaction des cyclotides avec les membranes conduisant à leur rupture (**Craik, 2012**). Ces composés semblent présenter une certaine cytotoxicité d'après les CI_{50} sur cellules du lymphome humain de l'ordre de 1 à 5 µM (**Herrmann et al., 2008**). Des études réalisées sur cellules cancéreuses de cyclotides isolés chez *Viola tricolor*, ont montré que les cyclotides type Bracelet présentaient une cytotoxicité plus importante que les Möbius (**Tang et al., 2010**).

I.6 INTERETS ECONOMIQUES DE LA VIOLETTE

Les violettes étaient dans un premier temps exploitées par les fleuristes pour la qualité odorante des fleurs dans la confection de bouquets. C'est vers la fin du XIXe siècle que la valorisation des feuilles pour la parfumerie débute grâce à l'apparition de nouvelles technologies de fabrication et de l'engouement pour la recherche de nouvelles fragrances. La production de violettes atteint son apogée au XXe siècle avec la commercialisation de trois produits différents : la confiserie et la confection des bouquets pour la fleur et la parfumerie pour les feuilles. Aujourd'hui, la fleur récoltée d'octobre à mars est très peu utilisée en parfumerie de par sa rareté et sa difficulté de récolte pour *in fine* de très faibles rendements d'extraction. Les progrès de la chimie ont donc suppléé l'extraction, avec la synthèse des molécules odorantes aux caractéristiques olfactives identiques, appartenant essentiellement à la famille des ionones. Elle est également présente dans le domaine culinaire, par exemple en confiserie ou dans les salades, en combinaison avec la framboise ou encore en tant que liqueur. En revanche, les feuilles sont encore utilisées en parfumerie par l'intermédiaire de la concrète et de l'absolue qui présentent des notes vertes, grasses et cireuses (**Grasse et al., 2009b**).

En 1908, la coopérative des producteurs de violettes est créée à Toulouse, comptant jusqu'à 600 producteurs. Au plus fort de la production, jusqu'à 600 000 bouquets par an sont exportés en Europe et au-delà. Dans les années 50, en raison des hivers rigoureux et de la dégénérescence génétique de la violette de Toulouse à l'origine d'une plus grande sensibilité aux maladies, les producteurs abandonnent peu à peu la culture. La culture de la violette a été relancée dans les années 80 à l'initiative de la chambre d'agriculture de la Haute-Garonne, mais seulement quelques producteurs sont encore en activité aujourd'hui à la périphérie toulousaine comme à la ferme Viola 2000 située à Renneville (31). Néanmoins, plusieurs associations voient le jour dont en 1989, « l'Association des producteurs de violettes de Toulouse » qui regroupent producteurs de fleurs coupées et de plantes en pots ; en 1993, « l'Association Terre de Violettes » qui regroupe les industriels qui produisent des parfums, des liqueurs et des violettes cristallisées à partir des violettes doubles ; ou encore en 1997, la « Confrérie de la Violette ». Dans le but de sauvegarder ce patrimoine naturel, un conservatoire national a été créé en 1994 aux serres municipales de Toulouse (Figure 11). Cette collection a pour but de maintenir le

patrimoine culturel et génétique et de le mettre à la disposition des chercheurs, des producteurs, des industriels comme des amateurs.

Figure 11. Serre des violettes localisée aux serres municipales de Toulouse

Les serres conservent une centaine de plants de violettes de toutes origines : certaines ont été référencées en Amérique (15), d'autres en Asie (2), en Australie (9), en Nouvelle-Zélande (1), en Afrique (1) et principalement en Europe (55). Une grande diversité de phénotypes est observée, notamment au niveau de la couleur des fleurs, de leur forme avec la présence majoritaire de fleurs simples et de quelques fleurs doubles, ainsi que de la forme des feuilles pour l'essentielle, tout de même, en forme de cœur. Une grande partie de la collection est identifiée par des noms de cultivars ou vernaculaires et seulement un faible pourcentage, soit 21%, possède des noms scientifiques type *V. mirabilis* et *V. sulfurea* (Figure 12). La liste des violettes et leur appelation au sein des serres peuvent être retrouvées en annexe 10, page 232.

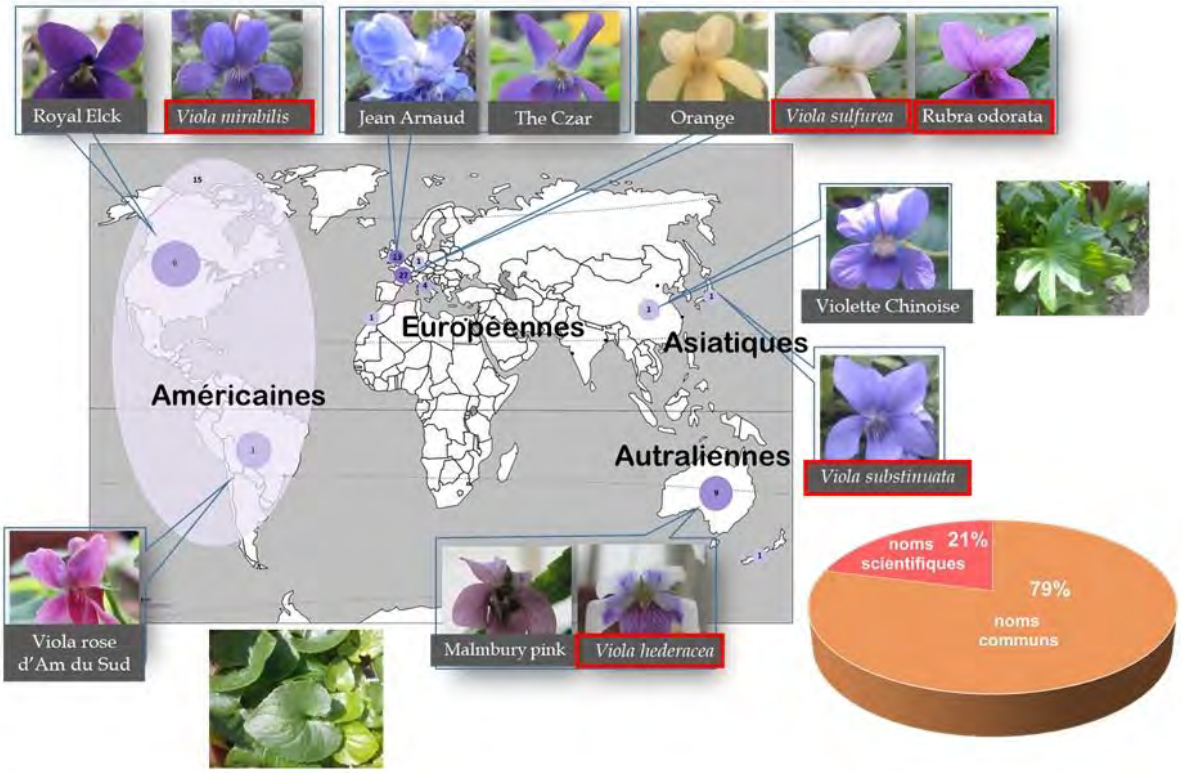


Figure 12. Carte des origines géographiques des violettes étudiées mettant en évidence la diversité de phénotypes. Seuls les noms encadrés en rouge sont scientifiques.

Même si la violette, symbole de la ville rose, est tombée progressivement en désuétude, elle reste déclinée en divers produits vendus dans les boutiques de souvenirs. Le groupe Berdoues fait par ailleurs vivre cette fleur depuis plus de cent ans à travers ses parfums et savons à l'odeur de violette. En effet,

en 1936 est commercialisé la première eau de toilette à la violette de Toulouse baptisée *Eau de Cologne Tradition 1902* présentant aujourd'hui onze fragrances. La violette est souvent associée à l'iris ou au musc blanc, comme c'est le cas pour le parfum Violette de Molinard et Arlésienne de L'Occitane en Provence. L'odeur de la feuille de violette est aujourd'hui encore très utilisée en notes de têtes de nombreux parfums, aussi bien masculins que féminins. La violette est également présente dans des produits cosmétiques tels que des savons déposés sous la marque « L'Originale – Violettes de Toulouse » par Berdoues. Cependant, l'ensemble des produits dérivés ne portent plus que son nom. En effet, la culture difficile de la violette de Toulouse a poussé les producteurs à abandonner leur activité et l'approvisionnement français provient essentiellement des Alpes-Maritimes à Tourettes sur Loup.

Ces dix dernières années, les récoltes de violette ont subi une forte diminution en 2009, 2012 et 2013 et une nette progression a été observée à partir de 2014 pour se stabiliser autour de 50 tonnes de violettes traitées en 2015. Une forte inflation du prix de la violette a été notée passant de 1,25 €/kg en 2008 à 1,64 €/kg en 2015 (http://www.franceagrimer.fr). Aujourd'hui, le prix du kilo tend à se stabiliser autour de cette valeur (Figure 13).

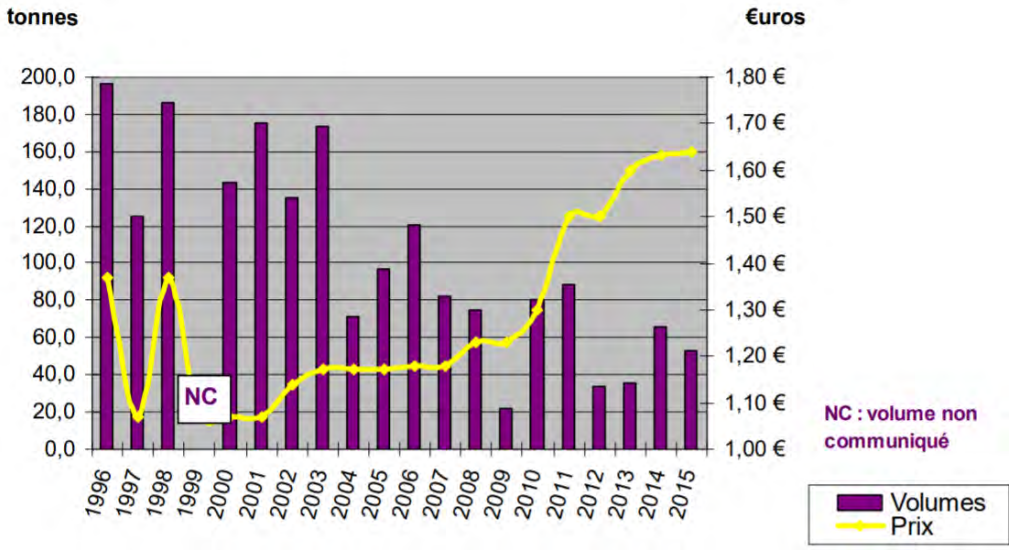


Figure 13. Évolution des volumes traités (t) de violettes de Tourette et prix du kg (€) de 1996 à 2015

La violette de Toulouse reste un symbole fort de la région Occitanie en général et de la ville de Toulouse en particulier. La Mairie de Toulouse la met en avant chaque année au cours de la fête de la violette regroupant tous les acteurs locaux, dont la célèbre Confrérie de la violette de Toulouse, présidée par Mme Calas.

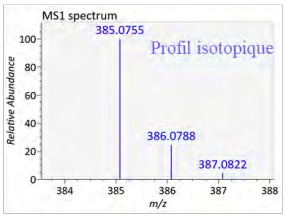
PARTIE II
RESULTATS & DISCUSSIONS

II.1 METHODOLOGIE GENERALE DE LA THESE

La méthodologie appliquée dans ces travaux repose sur les empreintes chimiques des violettes, c'est-à-dire les chromatogrammes LC-MS et GC-MS. En travaillant à partir de plusieurs acquisitions, des analyses statistiques peuvent ensuite être réalisées, par exemple avec le logiciel SIMCA-P, et qui peuvent être de deux ordres : les méthodes de classification telles que les arbres hiérarchiques (HCA) et les analyses en composantes principales (ACP) qui permettent d'ordonner les échantillons en groupes homogènes ; et les régressions type orthogonales sur structures latentes (OPLS) par exemple, généralement utilisées pour cibler les métabolites d'intérêt par rapport à une activité biologique étudiée. Dans ces travaux l'ACP et l'OPLS ont particulièrement été mises en œuvre puisque ce sont des méthodes adaptées aux matrices asymétriques, ce qui était le cas dans cette étude puisque beaucoup de variables étaient obtenues à partir de peu d'individus. Par ailleurs, ce sont des méthodes qui présentent l'avantage de proposer des réponses graphiques, facilement interprétables.

Tous les composés mis en évidence par ces analyses multivariées ont été identifiés essentiellement de manière *in silico* à l'aide du logiciel MS-FINDER (**Tsugawa et al., 2016**). La masse haute résolution obtenue à partir de l'Orbitrap permet de générer des formules brutes pertinentes avec une différence de masse inférieure à 5 ppm, en se basant notamment sur le profil isotopique. A l'aide de bases de données de produits naturels généralistes comme le Dictionnaire des Produits Naturels (DNP) ou plus centrées sur le genre *Viola*, et en se basant sur les résultats de fragmentation, le logiciel va proposer des candidats adaptés aux différentes données spectrales, pour remonter à une structure chimique avec l'identification des différents substituants (méthyle, glucose, …) et de la classe chimique du composé. Cette dernière est ensuite validée par une dernière donnée qui est le spectre UV, puisque celui-ci est spécifique d'un chromophore. La combinaison des données spectrales permet une identification de catégorie 2 (**Sumner et al., 2007**). Il est ainsi plus pertinenent de parler d'annotation puisque l'identification assurant la détermination structurale certaine est obtenue par comparaison des propriétés chimiques avec un standard. La stratégie d'annotation évoquée est résumée Figure 14.

1) Masse haute résolution (HRMS)



Détermination de la formule brute

2) Base de données (DNP,…)

3) Spectre MS/MS

Proposition d'identification

4) Spectre UV

Validation de la classe chimique

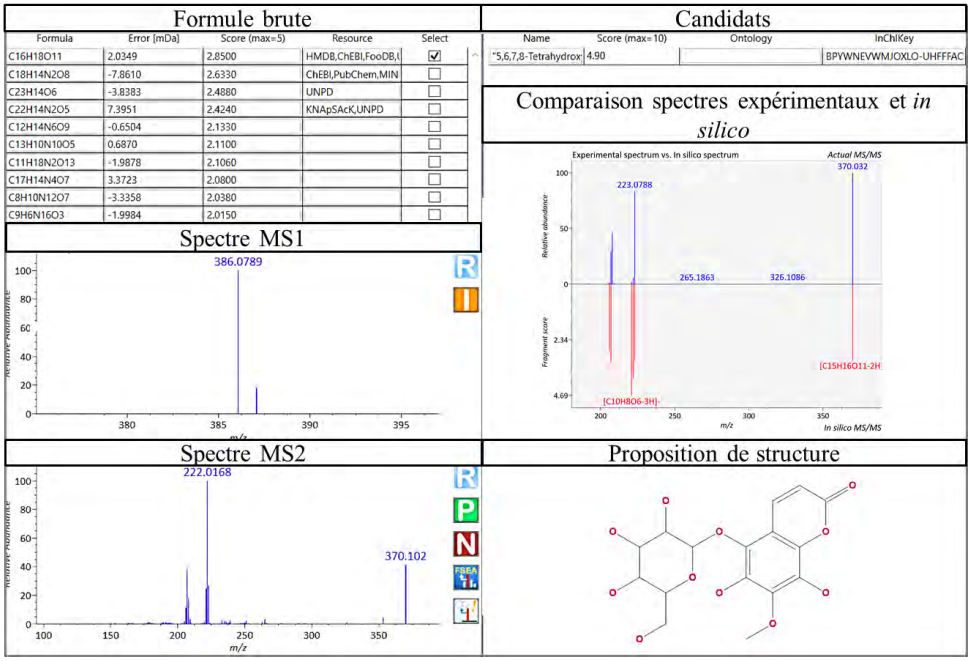


Figure 14 : Stratégie d'annotation des composés

II.2 CLASSIFICATION DES VIOLETTES DE LA COLLECTION

II.2.1 CONTEXTE

L'objectif de ce travail est d'obtenir une classification des violettes en divers groupes homogènes et de les identifier par un nom d'espèce. Cependant, la classification des espèces de *Viola* en sections et sous-sections peut être problématique, en particulier lors de l'utilisation de caractéristiques morphologiques. Ceci est dû à l'hybridation qui est relativement commune chez les plantes vasculaires et qui a un impact important sur la spéciation (**Erben, 1996**). Une telle hybridation peut contribuer à une diversification rapide de nombreuses lignées végétales, avec des changements caryotypiques et génomiques (**Arnold, 2006**), mais peut aussi conduire à la disparition d'espèces rares (**Todesco et al., 2016**).

Des analyses génétiques sur la collection (cf. bases théoriques en annexe 1.A, page 211), majoritairement identifiée par des noms de cultivars ou vernaculaires, ont été réalisées par l'intermédiaire des régions d'espaceurs internes transcrits (ITS) très populaires dans les études phylogénétiques (**Álvarez, 2003**) puisque les séquences nucléotidiques des régions ITS 1 et ITS 2 sont fortement polymorphes entre espèces, tout en étant relativement conservées. La méthode des ITS, adaptée pour les espèces voisines, est facile à mettre en œuvre et permet de s'affranchir de l'extraction préalable de l'ADN à l'aide d'un kit spécifique. Pour améliorer la résolution de la caractérisation génotypique de la collection, des études chimiotaxonomiques ont été menées en parallèle par une approche métabolomique (cf. bases théoriques en annexe 1.B page 215), par l'association de biomarqueurs potentiels associés à des génotypes spécifiques (**Cox et al., 2014**). Lors d'une première étude chimiotaxonomique les composés volatils des fleurs fraîches ont été analysés par espace de tête couplé à une micro-extraction en phase solide (HS-SPME) suivi d'une analyse chromatographique en phase gazeuse couplée à la spectrométrie de masse (GC-MS). Dans une seconde étude, les métabolites secondaires non volatils des extraits éthanoliques de feuilles ont été analysés par chromatographie liquide haute performance couplée à la spectrométrie de masse haute résolution (UHPLC-HRMS)[1].

II.2.2 CLASSIFICATION PAR ANALYSE GENETIQUE

II.2.2.1 Séquences des régions ITS

L'ensemble de la région ITS (~ 610 pb) a été amplifié pour 85 plants de la collection et séquencé (cf. IV.3.1 page 168). Il en est résulté 80 séquences bien définies. Certaines violettes n'ont pas été séquencées (N = 26) ou ont présenté des séquences non exploitables (N = 9), probablement en raison de la présence d'hétérozygotes ou d'un polymorphisme intra-individuel suggéré par l'observation de doubles pics pour certains sites nucléotidiques, comme c'était le cas pour la violette nommée Luxonne (V91) qui n'a par conséquent pas pu être clairement identifiée (Figure 15). Pour 31 violettes, seulement quelques doubles pics ont été observés le long de la séquence entière. Ainsi les deux séquences ont été séparées manuellement et une identification a été effectuée pour chacune. Dans chaque cas, les mêmes espèces, mais parfois des cultivars différents, ont été obtenues, révélant une hybridation significative entre les violettes de la même espèce.

[1] L'ensemble de cette étude a fait l'objet d'une publication dans le journal *Phytochemistry*, en cours de soumission

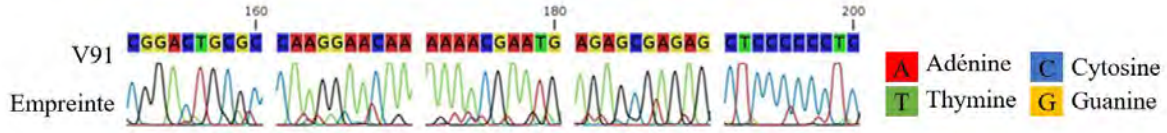
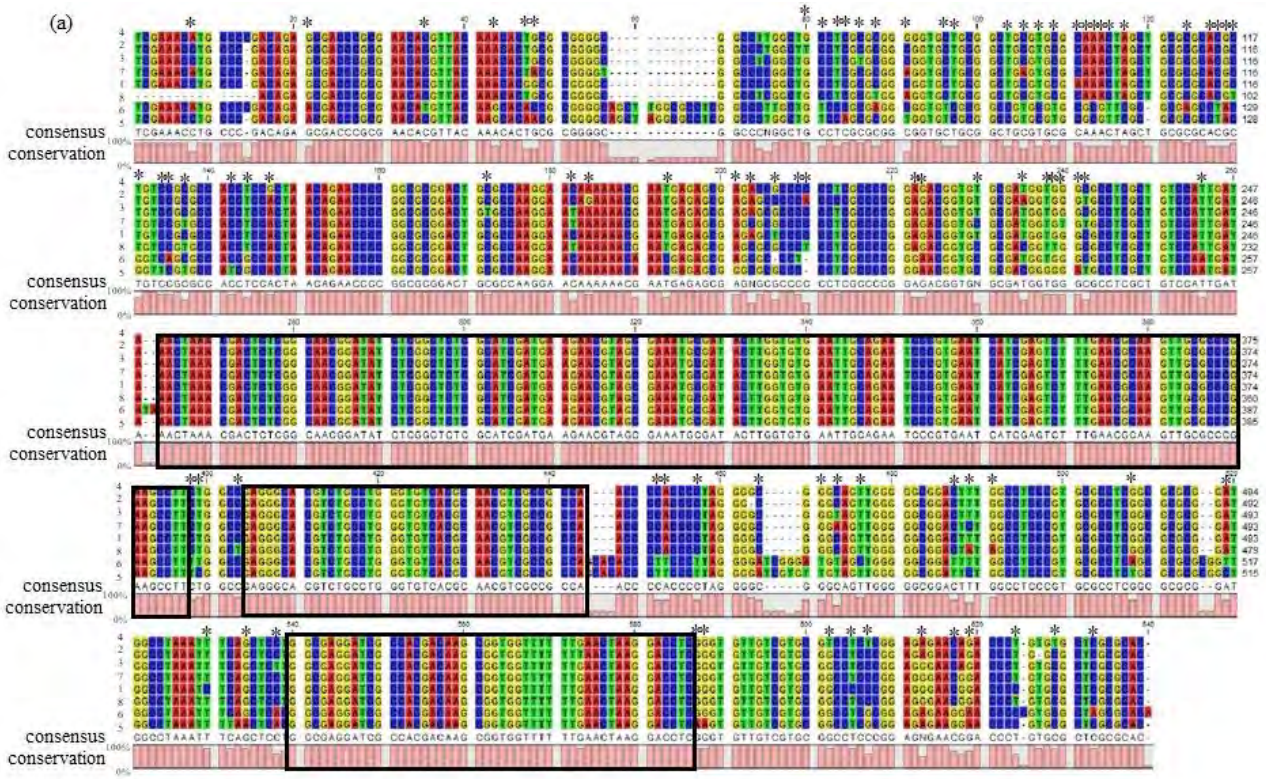


Figure 15. Partie de séquence de Luxonne (V91) illustrant la présence de doubles pics

L'alignement des séquences, obtenu en maximisant le nombre de sites communs (Figure 16 a) a mis en évidence huit principales séquences présentant des divergences et 91 sites polymorphes, dont 57 pour ITS 1 et 34 pour ITS 2. Quarante-sept intervalles (gaps) ont également été recensés, ce qui amène à 138 le nombre de sites présentant des divergences. Les séquences partagent donc 502 (71%) nucléotides sur les 640 positions de l'alignement. On observe trois principales zones de similarité : les positions 254-397, 404-443 et 540-443. En moyenne, les séquences sont constituées de 610 paires de bases (pb) et présentent 16 à 18% de base T, 20% de A, 31 à 33% de C et 30-31% de G (Figure 16 b).

(a)



* Positions variables parmi les huit séquences
\- Trou d'alignement
Zone de similarité
A Adénine C Cytosine T Thymine G Guanine

(b) Ordre des séquences (de haut en bas) : 4-2-3-7-1-8-6-5

Séquence	1	2	3	4	5	6	7	8
Nb pb	612	610	612	613	634	675	612	598
% T	16%	17%	18%	17%	18%	17%	17%	18%
% A	20%	20%	20%	20%	20%	20%	20%	20%
% C	33%	32%	31%	32%	31%	31%	32%	31%
% G	31%	31%	31%	31%	31%	30%	31%	31%
% N	-	-	-	-	-	2%	-	-

Figure 16. Alignement des séquences des huit échantillons distincts (a) et caractéristiques de chacune (b)

II.2.2.2 Construction des arbres phylogénétiques

II.2.2.2.1 Arbre de distance

Les séquences clairement définies (N = 80) ont été comparées les unes aux autres par un modèle de distance, avec la méthode « Neighbor-Joining » (NJ) et le modèle de substitution Kimura 80 (cf. bases théoriques en annexe 1.A page 211), pour obtenir une classification des violettes par l'identification de groupes homogènes (cf. IV.3.2.1, page 170). Une valeur seuil de 80% pour l'analyse bootstrap a été définie, ce qui signifie que les branches retranscrites ont été trouvées 800 fois sur les 1000 répliques de construction. Les branches sont évaluées en utilisant la technique de rééchantillonage pour étudier la fiabilité de la topologie de l'arbre phylogénétique. Six groupes de différentes tailles ont été mis en évidence et semblaient être corrélés à des espèces spécifiques (Figure 17). Cinquante-cinq pourcents des séquences sont regroupés dans un groupe, environ 15% pour deux autres groupes et environ 5% pour les trois derniers. Par conséquent, alors que la collection présente une variabilité phénotypique assez importante, la majorité des violettes possède des séquences d'ADN similaires. Sur la base de cette analyse des séquences ITS, une classification préliminaire de la collection de violettes a été obtenue. Le dendrogramme résultant représente une hypothèse de relations phylogénétiques entre taxons et relate une classification phylogénétique.

L'intégration de séquences de références trouvées dans la base de données de nucléotides du « National Center for Biotechnology Information » (NCBI), base de données publique utilisée au niveau international comme norme pour l'annotation du génome (**Pruitt et al., 2004**), permet d'apporter une hypothèse d'identification, mais ne permet pas de la confirmer. Ces références étaient : *Viola alba* isolat ALB184 V6 (GenBank: EU413916), *Viola sp* Hearn cult 33 (GenBank: DQ521291), *Viola suavis* isolat SUW170 V17 (GenBank: EU413930), *Viola mandshurica* isolat TKM201586 (GenBank: KX394615), *Viola alba subsp dehnhardtii* clone 200781 (GenBank: EU430656), *Viola odorata* ODO178 V11 (GenBank: EU413922), ODO214 V8 (GenBank: EU413919), ODO182 V3 (GenBank: EU413918).

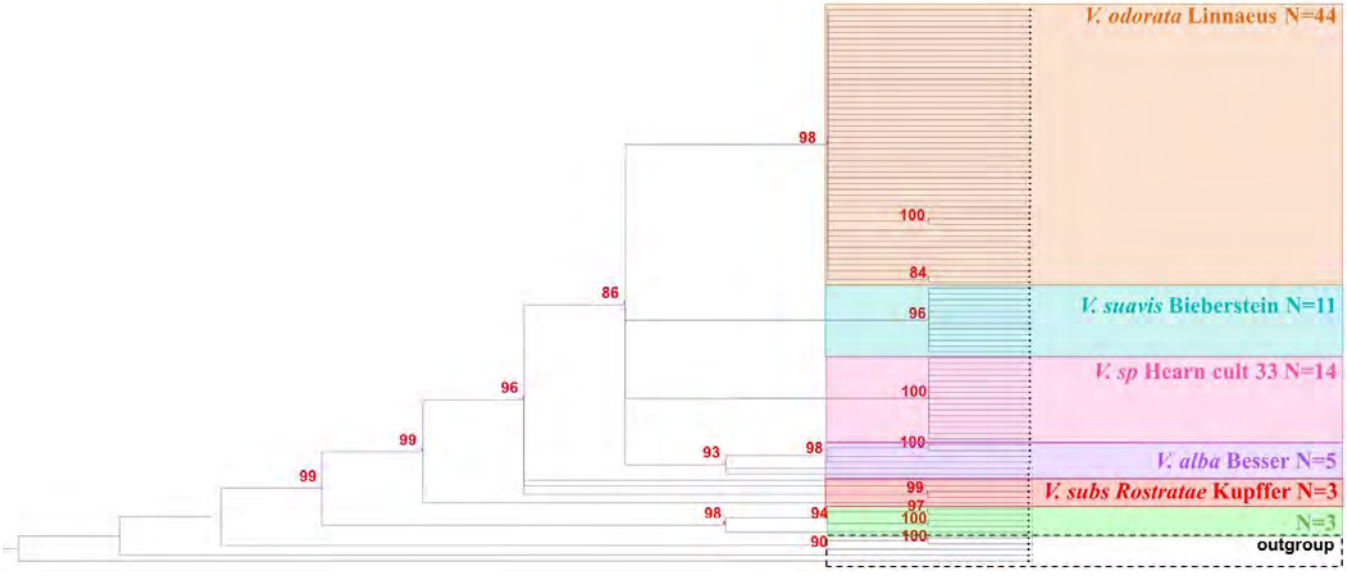


Figure 17. Cladogramme de similarité de séquences par une méthode de distance avec indication des valeurs bootstrap supérieures à 80%

II.2.2.2.2 Reconstruction phylogénétique

Les comparaisons de séquences faites avec des références trouvées dans NCBI ont servi à construire des arbres phylogénétiques (cf. IV.3.2.2, page 170) avec une méthode de maximum de parcimonie

(Figure 18 a) et une méthode de maximum de vraisemblance (Figure 18 b) (cf. bases théoriques en annexe 1.A page 211).

L'alignement des séquences expérimentales avec ces références a fourni un score de similarité, et une bonne identification a été validée pour des scores supérieurs à 98,5%. Ce fut le cas pour 64 violettes (58% de la collection totale): deux étaient référencées dans la sous-section *Rostratae* Kupffer (à savoir *V. labradorica* Schrank et *V. grypoceras* A. Gray), et les autres dans la sous-section *Viola* dont deux comme *V. alba* Besser, six comme *V. suavis* Bieberstein, douze comme *V. sp*. Hearn cult 33, une comme *V. verecunda* A Gray, une comme *V. mandshurica* Becker et quarante en tant que *V. odorata* Linnaeus. En combinant les résultats, il est ressorti que la séquence n°1 (Figure 16) correspondait à *V. odorata*, la n°2 à *V. alba*, la n°3 à *V. suavis*, la n°4 à *V.sp* Hearn, la n°5 à *V. verecunda*, la n°6 à *V. mandshurica*, la n°7 à *V. labradorica* et la n°8 à *V. grypoceras*. Il est à noter que le nom *V. sp.* Hearn cult 33 (abrégé par la suite *V. sp* Hearn) n'est pas un nom d'espèce reconnu. C'est vraisemblablement un nom issu de la personne ayant décrit cette violette pour la première fois, David J. Hearn, mais n'étant pas parvenu à l'identifier en tant qu'espèce. Par la suite, les violettes *V. labradorica* et *V. grypoceras* faisant partie d'un même groupe, elles ont été regroupées sous le nom de la sous-section *Rostratae* (**Marcussen et al., 2010**), aussi connue sous le nom *V. subs*. *Trigonocarpea* Godr., afin de ne pas alourdir le discours avec le nom des espèces et puisque c'est cet ensemble en général qui sera utilisé. Par ailleurs, il est à noter que les noms suivant les appellations scientifiques des espèces (type, Besser, Schrank, …) sont cités pour cette première apparition pour empêcher toute confusion avec d'autres entités pouvant porter ce nom. Par la suite, ces noms ne seront plus précisés, le cadre ayant été fixé dans ce paragraphe.

Afin de construire le cladogramme circulaire de maximum de vraisemblance (Figure 18 b), seules des séquences de quelques échantillons caractéristiques de chaque groupe identifié par la méthode de distance ont été sélectionnées. L'espèce *Rinorea ledermannii* (Violaceae) a été utilisée pour enraciner les arbres. Comme précédemment, une valeur seuil de bootstrap de 80% a été fixée et les identifications suivantes ont été faites : *V. odorata* pour le groupe orange avec 44 violettes, *V. alba* pour le groupe violet avec 5 violettes, *V. sp* Hearn cult 33 pour le groupe rose avec 14 violettes, *V. suavis* pour le groupe turquoise à 11 violettes et *V. subs Rostratae* pour le groupe rouge à 3 violettes. Une branche non identifiée (en vert) avec 4 violettes, dont *V. verecunda* et *V. mandshurica*, était bien séparée des autres groupes.

On peut remarquer des embranchements secondaires à l'intérieur des embranchements primaires délimitant les différents groupes. Cette observation pourrait être à mettre en relation avec la présence d'éventuelles sous-espèces non déterminées par la méthode phylogénétique appliquée ici. Finalement, 64 violettes ont été clairement indexées. Pour les autres violettes analysées, l'appartenance à un groupe génétique bien défini et identifié par un nom d'espèce apporte une hypothèse forte d'identification.

Cependant, la classification obtenue ici présente certaines limites, par exemple pour la violette de Toulouse identifiée comme *Viola alba* (groupe violet). En effet, dans la littérature, elle est connue sous le nom *V. alba* subs *dehnhardtii* (**Malécot et al., 2007**). De plus, des violettes appartenant à *V. odorata* ont été distribuées parmi les trois isolats précédemment cités (ODO178, ODO182 et ODO214). Cependant, leurs séquences étaient fortement analogues avec 99% voire même 100% de similarité. Ceci est indicatif de la limite d'usage de la séquence ITS pour une identification précise des variétés et donc de la nécessité de la complémentarité d'autres méthodes, telles que l'ADN chloroplastique, pour compléter et affiner les résultats. Néanmoins, le but principal de notre étude étant d'obtenir une classification des violettes en divers groupes génétiquement caractérisés, limiter leur identification au niveau de l'espèce paraissait suffisant à ce stade.

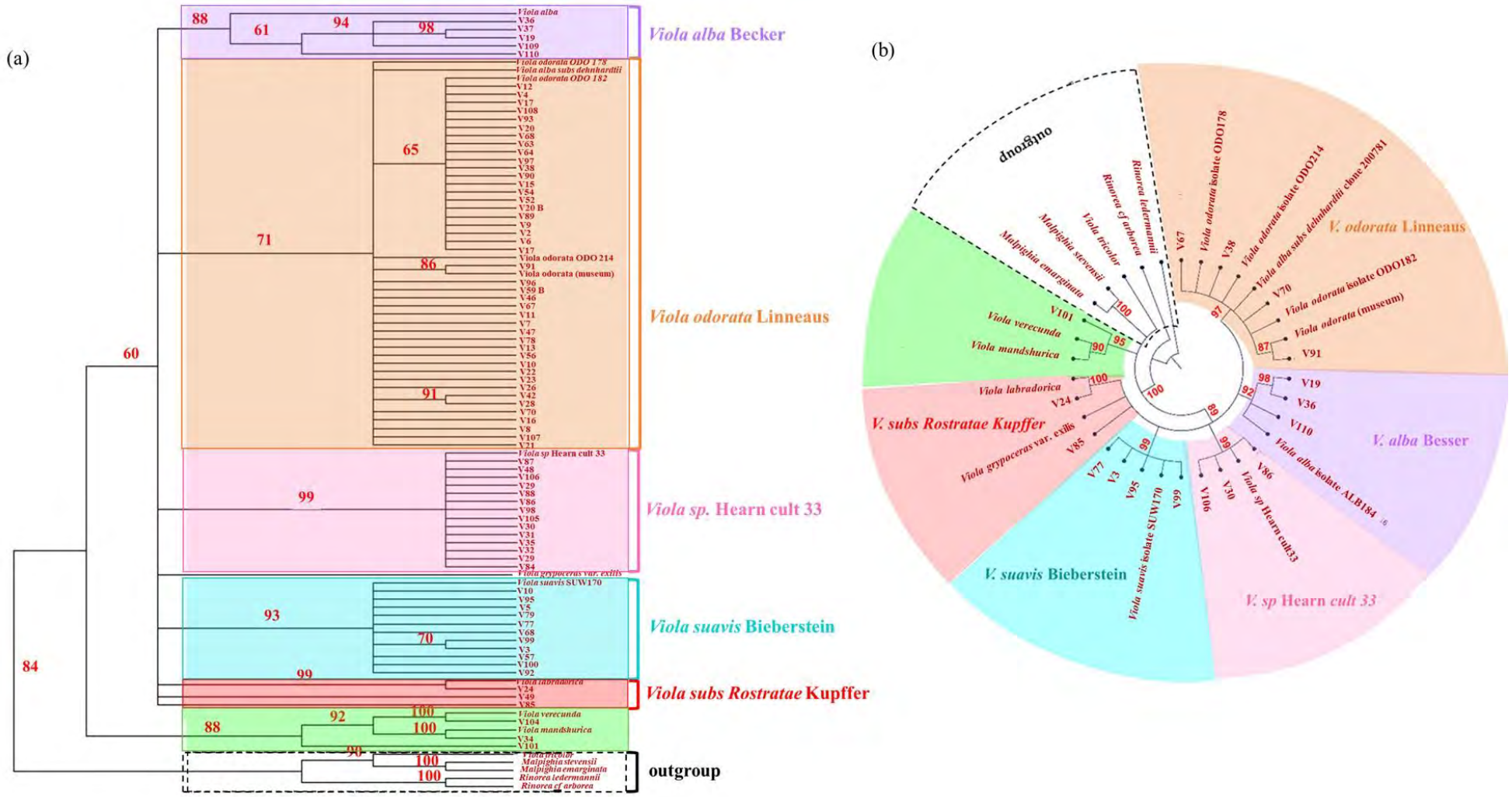


Figure 18. Dendrogramme du maximum de parcimonie (a) et cladogramme circulaire du maximum de vraisemblance (b) de l'analyse phylogénétique des violettes avec indication des valeurs bootstrap supérieures à 60%

II.2.2.3 Conclusion

L'analyse de distance a permis de mettre en évidence la présence de groupes bien définis, car caractérisés par des valeurs de bootstrap supérieures à 80%. Avec l'intégration de séquences de références issues de NCBI et par les méthodes de maximum de vraisemblance et de maximum de parcimonie, six groupes ont pu être distingués : 7 violettes identifiées comme *V. suavis* et 4 assimilées, 11 à *V. sp* Hearn et 3 assimilées, 2 à *V. alba* et 3 assimilées, 2 à *V. subs*. *Rostratae* et 1 assimilée, 35 à *V. odorata* et 8 assimilées et enfin 1 à *V. verecunda* et 1 à *V. mandshurica* intégrées dans un même groupe. Ainsi, cette analyse phylogénétique a permis de mettre en évidence une certaine disparité de séquences au sein de la collection. Cependant, près de la moitié des violettes reste non identifiée en raison de leur non-séquençage dû à un manque de matière végétale ou de mauvaises identifications avec des scores de similarité par rapport aux séquences de référence inférieurs à 98%. Ainsi, à ce stade la classification des violettes reste incomplète. Des analyses chimiotaxonomiques telles que les analyses des composés volatils des fleurs et des métabolites secondaires non volatils des feuilles ont donc été effectuées pour apporter un complément d'information, dans le but d'obtenir une classification plus étoffée de cette collection. Les approches métabolomiques non ciblées couplées à la LC-MS (**Martucci et al., 2014**) ainsi qu'aux données SPME GC-MS (**Khalil et al., 2017**) ont déjà démontré leur potentiel dans la classification taxonomique de différentes espèces. Comme expliqué avec le concept des espèces (**Smedsgaard et Nielsen, 2005**), la classification et la compréhension des espèces reposent sur trois axes principaux: les caractères phénotypiques, le comportement écologique et le génome. Cependant, ces aspects ne sont pas toujours suffisants dans le cas d'espèces voisines ou lors d'un processus de spéciation. Puisque le métabolome représente l'image finale résultant de l'interaction génomique et environnementale, le profilage des métabolites apparaît comme pertinent pour compléter les études taxonomiques basées sur la génétique. Pour cela une méthode statistique de classification a été mise en place (Figure 19) à l'aide d'un modèle construit sur la base du jeu de données des violettes clairement identifiées en génétique. La capacité prédictive du modèle est alors évaluée à l'aide des violettes non identifiées qui sont injectées pour tenter des les identifier par projection dans des groupes identifiés.

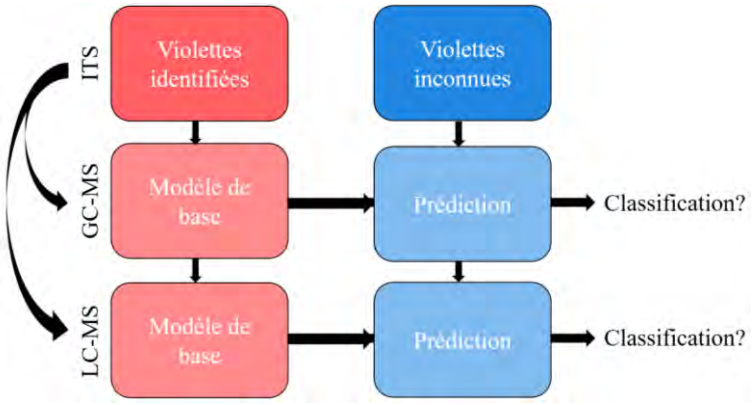


Figure 19. Méthode de classification basée sur l'établissement de modèles statistiques de base et de prédiction.

II.2.3 Extension de la classification par analyses phytochimiques

II.2.3.1 Intégration des profils volatils des fleurs

Pour des raisons de conservation et de non-dégradation des pots de la collection, il a été décidé de travailler uniquement dans la configuration de la plante entière sous cloche. L'analyse d'un plant dépourvu de fleurs a permis de valider l'approche sous cloche puisqu'une très faible émission des

feuilles a été constatée (annexe 2, page 218). La méthode d'analyse par HS-SPME et GC-MS couplée à l'olfactométrie (GC-O-MS) (cf. IV.5.1, page 173) a été premièrement optimisée avec la violette de Toulouse afin de définir les temps d'équilibre de la plante sous cloche et d'exposition de la fibre pour l'absorption des composés. Il a été décidé d'utiliser la fibre SPME DVB-CAR-PDMS pour son caractère mixte et puisqu'elle permet l'analyse des composés semi-volatils à volatils.

II.2.3.1.1 Optimisation de l'analyse HS-SPME par plan de Doehlert

II.2.3.1.1.1 Principe du plan de Doehlert

La particularité du plan de Doehlert (**Ferreira, 2004**) utilisé pour l'optimisation des paramètres d'extraction des composés volatils est de présenter des points expérimentaux uniformément répartis dans l'espace (Figure 20). L'expérimentateur peut facilement ajouter des points d'expériences supplémentaires et retrouver une disposition identique à celle de départ. Le plan de Doehlert est économique en nombre d'essais tout en conservant une qualité suffisante et facilitant la recherche d'un optimum ou d'une valeur cible. Les étapes à suivre pour sa réalisation sont les suivantes :

- Définition du domaine expérimental
- Réalisation des expériences avec au moins trois répétitions
- Validation du modèle par un test de Fisher
- Calcul des coefficients de l'équation de 2nd degré ($y = a0 + a1X1 + a2X2 + a1a2X1X2 + a11X1^2 + a22X2^2$)
  - a11 et a22 < 0 ➔ maximum (Figure 20 b)
  - a11 et a22 > 0 ➔ minimum (Figure 20 c)
  - a11 et a22 de signes différents ➔ minimax, illustré par une selle de cheval (Figure 20 d)
- Calcul du point optimal et vérification de l'optimum par l'expérience

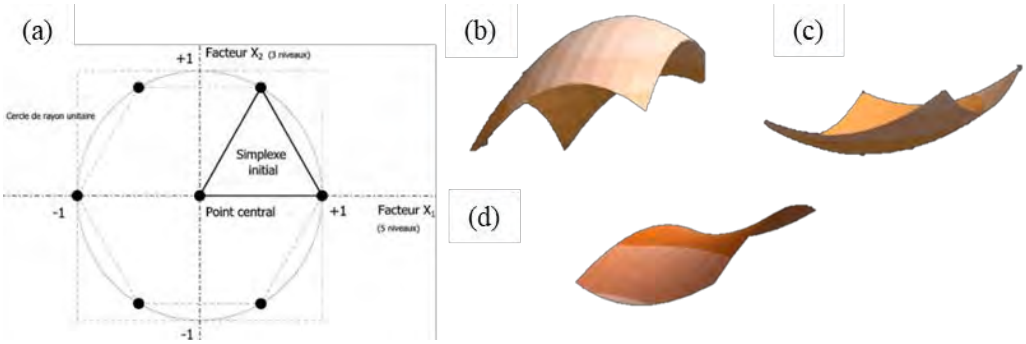


Figure 20. Principe du plan de Doehlert (a) et illustration d'une surface de réponse traduisant un maximum (b), un minimum (c) et un minimax (d)

II.2.3.1.1.2 Résultats de l'optimisation de l'analyse de la plante entière sous cloche

Le temps d'équilibre et d'exposition de la fibre étant à optimiser, le plan de Doehlert indique d'effectuer neuf expériences dont deux répétées au centre du domaine d'étude (Tableau 8), déterminé à partir de la littérature (**Vercammen et al., 2000**).

Tableau 8. Domaine d'étude des paramètres étudiés

	X1 Equilibre (min)	X2 Exposition (min)
Minimum	20	15
Maximum	80	45
Centre (X1max-X1min)/2	50	30,0
Pas (X1max – Centre(X1))/U	30	17,3

* U = valeur codée soit U=1 pour X1 et U=0,866 pour X2

Les résultats d'analyse des chromatogrammes des différentes expériences ont été les suivants (Tableau 9).

Tableau 9. Expériences à réaliser et résultats obtenus : nombre de pics et intégration expérimentale (Yexp1 et Yexp2) et ceux prédits par le modèle quadratique (Ycalc1 et Ycalc2)

Table de Doehlert Valeurs codées (U)		Valeurs réelles		Résultats			
X1	X2	Equilibre (min)	Exposition (min)	Yexp1	Ycalc1*	Yexp2	Ycalc2*
0	0	50	30	12	13	17,4	21,1
1	0	80	30	17	16	22,6	23,4
0.5	0.866	65	45	11	10	10,6	11,4
-0.5	0.866	35	45	5	6	10,1	9,3
-1	0	20	30	17	18	28,9	28,1
-0.5	-0.866	35	15	15	14	23,3	24,1
0.5	-0.866	65	15	14	13	23,0	21,1
0	0	50	30	15	16	22,4	21,6
0	0	50	30	13	13	22,8	21,1

* Réponses calculées par le modèle quadratique

Ces résultats ont permis de valider la répétabilité du plan puisque les répétitions au centre du domaine sont toutes du même ordre de grandeur et présentent respectivement des écart-types de 1 et 3 dans le cas de l'analyse des pics et de l'analyse des intégrations. Par ailleurs, à l'aide d'un test de Fisher ($F = \frac{\sigma_a^2}{\sigma_e^2} < F_{Fischer}$), le plan d'expérience a été validé dans les deux cas (Tableau 10) et les coefficients de l'équation de second degré, de type $y = a_{0+}\, a_1 X_1 + a_2 X_{2+} a_{12} X_1 X_2 + a_{11} X_1^2 + a_{22} {X_2}^2$, modélisant le phénomène ont été calculés[2] (Tableau 11).

Tableau 10. Test de Fischer dans le cadre de l'analyse sous cloche pour l'étude du nombre de pics (x/) et l'étude d'intégration des pics (/x)

Ddla[3] (n*-p**)	1 / 1	σ_a	6,0 / 4,0
Ddle[4] (N***-n)	2 / 2	σ_e	1,0 / 10,1
Ddlr[5] (N-p)	3 / 3	σ_r	2,7 / 8,1
Ft	18,5 / 18,5	*n : nombre de moyennes indépendantes = 7	
F	6,0 / 0,39	**p : nombre de coefficient calculés = 6	
		***N : nombre d'expériences = 9	

Tableau 11. Coefficient de l'équation de 2nd degré dans le cas de l'étude des pics (a) et de l'étude de l'intégration (b)

(a)

	$\hat{a}_{pic}$	±σ
a0	13,00	0,94
a1	1,00	0,94
a2	4,04	0,94
a12	-3,46	1,89
a11	4,00	1,49
a22	-3,33	1,49

(b)

	$\hat{a}_{int}$	±σ
a0	21,05	±σ
a1	2,34	1,64
a2	7,23	1,64
a12	0,23	3,28
a11	4,70	2,59
a22	-7,51	2,59

[2] $\hat{a} = ({}^tX \cdot X)^{-1} \cdot {}^tX \cdot Y$ avec Y : matrice de réponses expérimentales et X : matrice du modèle (table de Doehlert)

[3] ddla : degrés de liberté pour variance d'ajustement = nombre de moyennes indépendantes – nombre de coefficients calculés

[4] ddle : degrés de liberté pour variance expérimentale = nombre d'expériences – nombre de moyennes indépendantes

[5] ddlr : degrés de liberté pour variance des résidus = nombre d'expériences – nombre de coefficients calculés

Pour optimiser cette expérience, un optimum est recherché correspondant ici à un maximum puisque l'objectif est d'obtenir le plus grand nombre de pics chromatographiques et la plus grande intégration, ce qui se traduit par des termes carrés (a_{11} et a_{22}) négatifs. Or, les résultats indiquent des termes carrés de signes différents, traduisant la présence d'un minimax se matérialisant par une représentation des surfaces de réponse en forme de selle de cheval comme observé sur la Figure 21.

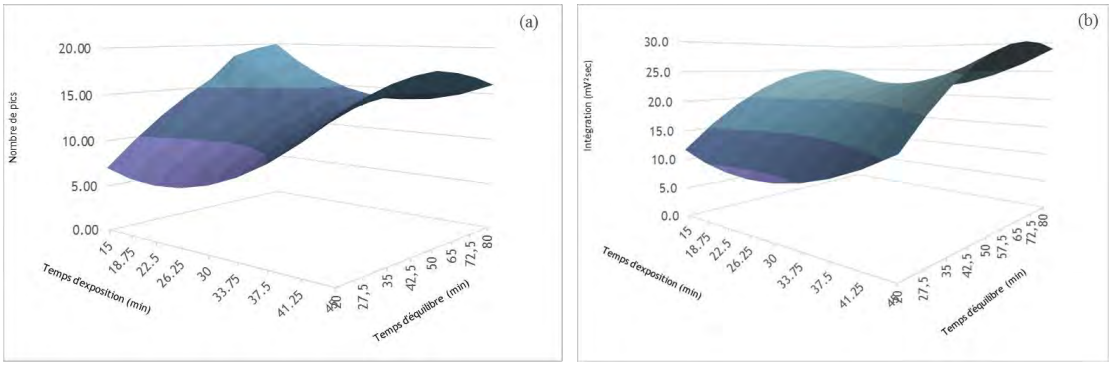


Figure 21. Surfaces de réponse dans le cas de l'analyse du nombre de pics (a) et dans le cas de l'analyse de l'intégration (b)

Par conséquent, une coupe a dû être réalisée au coefficient a_{11} qui est de signe inverse à celui recherché. Nous avons donc arbitrairement choisi de fixer le facteur X_1 (temps d'équilibre) lié à ce coefficient à la valeur codante 0,5 soit 65 minutes. A partir de là, de nouveaux coefficients a' ont été calculés et ces derniers sont résumés dans le Tableau 12.

Tableau 12. Nouveaux coefficients de l'équation dans le cas de l'étude des pics (a) et de l'étude de l'intégration (b)

(a)

	$\hat{a}'_{pic}$
a'0	14,50
a'2	2,31
a'22	-3,33

(b)

	$\hat{a}'_{int}$
a'0	23,39
a'2	7,34
a'22	-7,51

$a'2 = a2 + a12*0,5$
$a'22 = a22$
$a'0 = a0 + a1*0,5 + a11*0,5^2$

Ainsi, de nouvelles courbes de réponse ont été tracées (Figure 22) et traduisent bien le passage par un maximum (Xs) qui a pu être calculé[6] (Tableau 13).

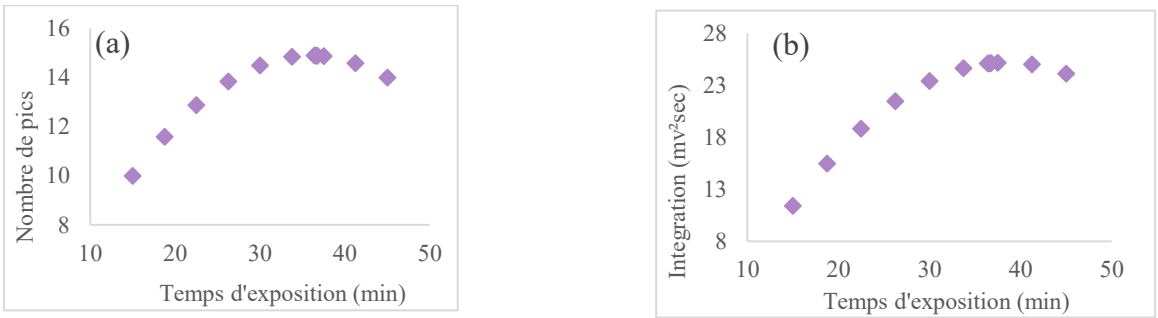


Figure 22. Nouvelles courbes de réponses dans le cas de l'étude des pics (a) et dans le cas de l'étude de l'intégration (b) pour un temps d'équilibre fixé à 65 minutes

[6] Xs (point stationnaire = optimum) = -½ A^{-1}.ak avec $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12}/2 & \cdots & a_{1k}/2 \\ a_{12}/2 & a_{22} & \cdots & a_{2k}/2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k}/2 & a_{2k}/2 & \cdots & a_{kk} \end{bmatrix}$ et $\mathbf{ak} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix}$

Tableau 13. Temps d'équilibre et d'exposition optimums pour l'analyse de la plante entière

	Optimum (pics)	Optimum (intégration)	Moyenne
Temps d'équilibre (min)	65	65	65
Temps d'exposition (min)	36	38	37

Cet optimum a dans un premier temps était vérifié par l'expérience. Pour cela l'analyse de trois pots différents a été réalisée sur trois jours. Les résultats sont regroupés dans le Tableau 14**.** Ils permettent de conclure à une reproductibilité satisfaisante de l'expérience. En effet, les écarts entre les différents résultats peuvent être expliqués par différents facteurs à savoir, l'incertitude sur l'adsorption au niveau de la fibre puisque des études ont montré que cette incertitude pouvait s'élever jusqu'à 10% ; mais aussi les fluctuations environnementales, les températures fraîches favorisant l'émission des composés volatils ; ou encore les fluctuations biologiques inhérentes à la plante.

Tableau 14. Résultats des expériences pour l'optimum théorique sur plante entière

	Pics	Intégration
Exp1	15	41,4
Exp2	17	46,8
Exp3	18	58,1
Moyenne	17	48,2
Ecart-type	2	9,3

Ainsi l'optimum théorique est ici validé. Cependant, sa mise en œuvre pratique n'est pas très satisfaisante, c'est pourquoi un optimum technique, défini à 60 minutes d'équilibre et 30 minutes d'exposition, a été choisi afin de faciliter la réalisation de l'expérience. Deux expériences ont été réalisées et les résultats sont présentés dans le Tableau 15.

Tableau 15. Résultats des expériences pour l'optimum technique sur plante entière

	Pics	Intégration
Exp1	16	55,4
Exp2	21	50,1
Exp3	18	47,6
Moyenne	18	51,0
Ecart-type	3	4,0

On observe que le résultat de l'optimum technique est comparable à celui de l'optimum théorique. Par conséquent, pour des raisons de mise en œuvre, l'optimum technique a été choisi pour la suite des expériences sur la collection de violettes.

II.2.3.1.1.3 Conclusion de l'optimisation

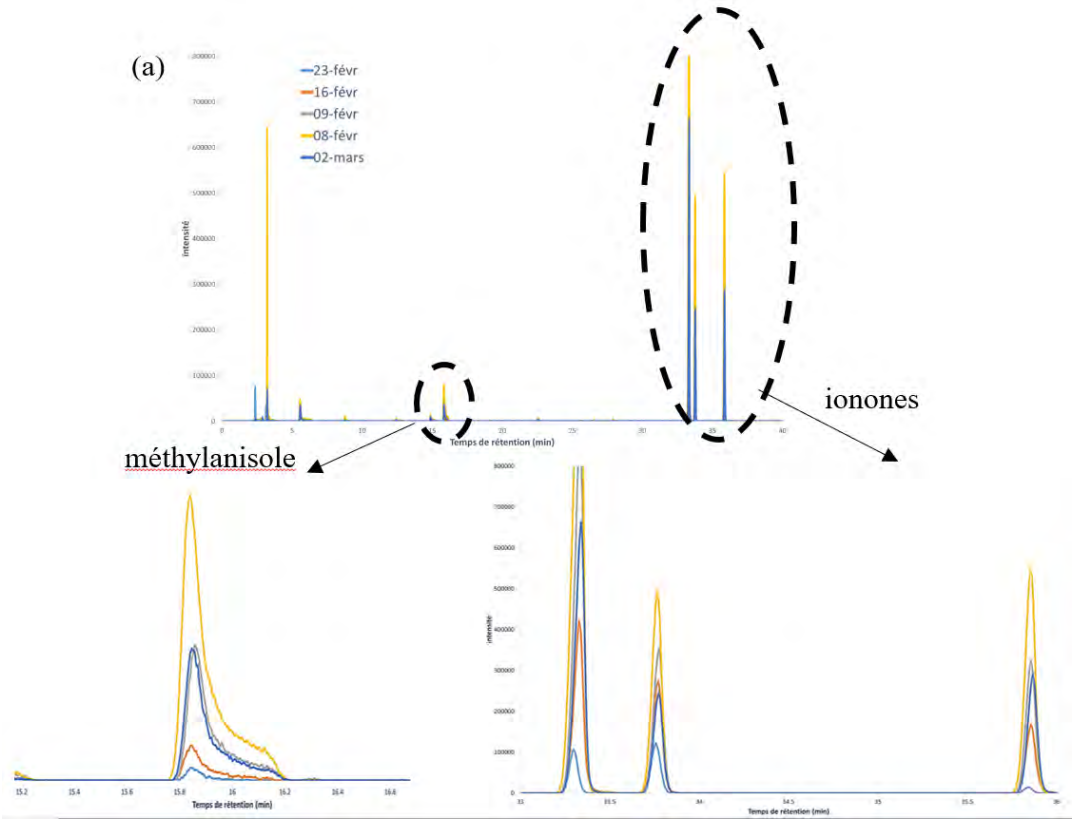
Les conditions de piégeage et d'adsorption ont donc été optimisées à l'aide de la mise en œuvre d'un plan d'expérience de Doehlert à partir d'échantillons de violettes de Toulouse. Les deux paramètres étudiés étaient le temps d'équilibre de la matière végétale sous la cloche et le temps d'adsorption des composés volatils émis sur la fibre SPME. Le plan d'expérience a permis de déterminer l'optimum théorique qui a été validé par l'expérience. Cependant, pour des raisons de mise en œuvre, un optimum technique a été proposé et validé expérimentalement puisque le résultat obtenu était comparable à celui de l'optimum théorique et a donc conduit à 60 minutes d'équilibre et 30 minutes d'adsorption. La nature de la fibre ou encore la température auraient également pu être optimisées pour plus de rigueur mais le temps imparti aux analyses sur environ deux mois de l'année ont conduit à sélectionner la fibre DVB-

CAR-PDMS puisqu'elle présente des propriétés intermédiaires et les expériences ont été réalisées à température ambiante pour limiter au maximum l'impact sur la floraison de la plante.

Les écarts observés au cours des expériences peuvent s'expliquer par la matrice végétale elle-même et son métabolisme, mais également la technique analytique choisie dont l'incertitude repose essentiellement sur les variabilités d'adsorption de la fibre. Néanmoins, les divergences observées ont été jugées acceptables et ainsi la répétabilité et la reproductibilité de l'expérience ont été confirmées. L'analyse d'un blanc, atmosphère sous la cloche, a également été réalisée afin d'éliminer les pics non pertinents. Les méthodes développées ont donc été appliquées à l'analyse d'une partie de la collection dans les mêmes conditions dans le but de mettre en évidence des critères de discrimination entre les différentes violettes. Dans cette étude comparative, des modèles statistiques ont été mis en œuvre en se basant sur l'alignement et la comparaison des profils GC.

II.2.3.1.2 Analyse métabolomique des données GC-MS

Les profils HS-SPME/GC-MS de 39 violettes de la collection ont été acquis et ont fourni 82 variables (paires m/z-RT) en mode d'ionisation positif (PI) (cf. IV.5.1.3.2.1, page 178). À titre d'étape préliminaire, l'analyse en composantes principales (ACP) a été utilisée comme outil exploratoire pour fournir une vue d'ensemble des empreintes volatiles. L'utilisation de données binaires basées sur la présence ou l'absence de pics dans chaque échantillon a été préférée pour limiter l'impact de la variabilité d'intensité. En effet, des analyses d'une même violette de Toulouse pendant une saison (février-mars) ont montré des variations quantitatives, mais non qualitatives, avec l'identification du même profil, mais des variations d'intensité de pic (Figure 23 a). Ainsi, les importants chevauchements observés sur l'ACP dans les deux cas, avec une petite amélioration pour les données binarisées due aux légers détachements des groupes *V. odorata, V. subs Rostratae* et *V. suavis* (Figure 23 b et c) ont conduit à conserver le mode binarisé dans la suite des traitements.



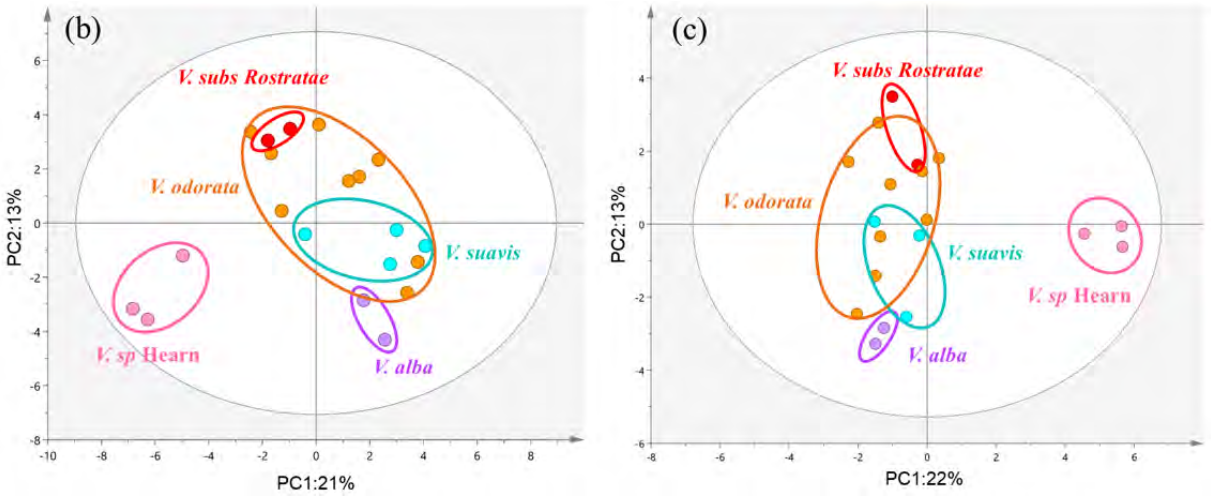


Figure 23. Étude cinétique sur la violette de Toulouse illustrant la variabilité quantitative (a), analyse en composantes principales avec les données brutes (b) et les données binarisée (c)

Cette ACP (Figure 23 c) a révélé un groupe, composé uniquement de violettes appartenant à *V.* sp Hearn, bien séparé des quatre autres groupes. Cette séparation est dépendante de la première composante PC1 tandis que les autres espèces sont plutôt séparées par la deuxième composante PC2. Cette observation suggère que *V.* sp Hearn présente un profil volatil qui diffère considérablement des autres espèces qui semblent, elles, partager des composés communs. Cependant, il faut prendre en compte le fait que les données ont été projetées sur un plan, mais qu'en réalité, la représentation est spatiale et la séparation pourrait subvenir sur d'autres axes non représentés.

Une analyse discriminante selon la projection orthogonale des structures latentes (OPLS-DA) a ensuite été appliquée, après une transformation logarithmique des données et une normalisation pareto, afin de tenter de classifier par projection les violettes non identifiées, et ainsi étendre la classification de la collection (cf. IV.5.2.4.2, page 186). Un modèle de base a donc été construit avec seulement les violettes génétiquement identifiées (Figure 24 a, N=19). La qualité de la prédiction du modèle était peu satisfaisante (R2Y = 0,93, Q2Y = 0,505, CV-ANOVA *p*-value = 0,59) et un test de permutation a permis d'évaluer une validité juste correcte (Figure 24 b). Cinq groupes caractérisés par les noms des espèces ont été distingués. La prédiction des violettes non identifiées (N = 20) a ensuite été tentée en prenant en compte 95% des données (Figure 24 c). Après prédiction, trois groupes, *V. alba*, *V. suavis* et *V. odorata* se chevauchent, indiquant la présence de profils probablement très similaires qui empêchent leur nette séparation. En revanche, les groupes *V. sp* Hearn et *V. subs Rostratae* sont restés bien distincts. Les résultats étaient plutôt peu satisfaisants car seulement dix violettes sur vingt, c'est-à-dire 50%, ont été projetées et classées dans les groupes définis. Ainsi pour 9% de plus de la collection des identifications potentielles ont été proposées en se basant sur les profils des composés volatils (Figure 24 d). Seulement, en raison de la qualité insuffisante du modèle, ces identifications ne peuvent être tenues pour acquises et doivent être validées par une autre méthode.

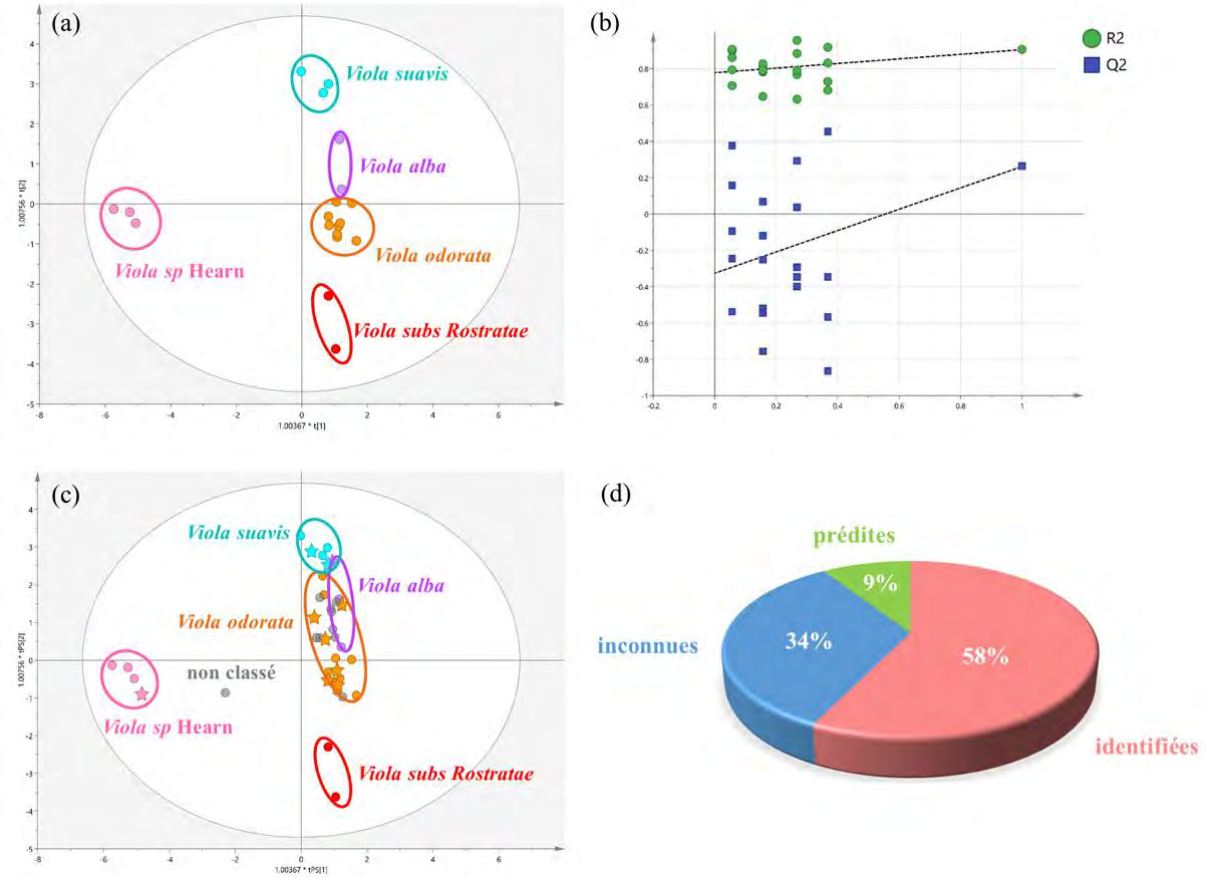
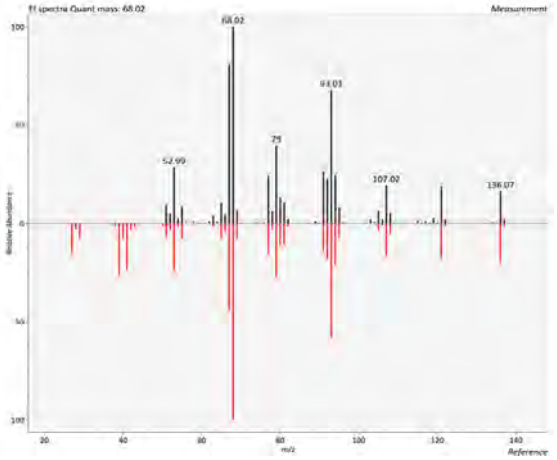


Figure 24. Modèle OPLS-DA des violettes identifiées (a), test de permutation (b), prédiction des violettes inconnues matérialisées par une étoile (c) et répercussions sur la classification de la collection (d)

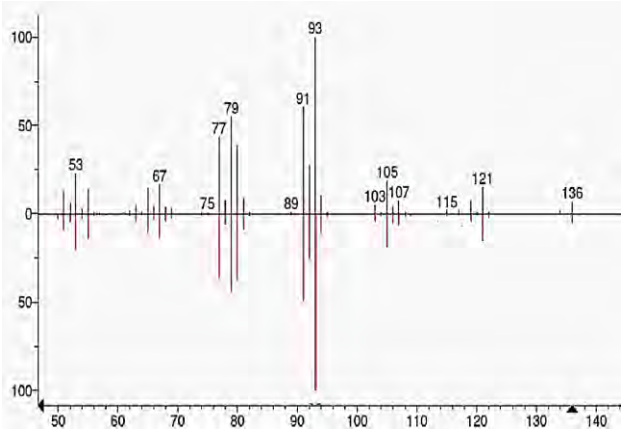
II.2.3.1.3 Identification de biomarqueurs des fleurs

Pour compléter cette classification des violettes basée sur les profils volatils des fleurs, certains biomarqueurs spécifiques aux espèces ont été recherchés par interrogation de la base de données NIST (cf. IV.5.1.3.2.2, page 178). Les comparaisons entre spectres de masse expérimentaux et *in silico* sont présentées Figure 25**.** Un profil limonène (a) a été révélé pour les violettes appartenant à *V. sp* Hearn avec la présence de certains terpènes dont l'un est semblable au β-cis-ocimène (b). Deux profils iononiques ont permis de distinguer les violettes appartenant à *V. alba* et *V. suavis*. En ce qui concerne le premier, les trois ionones classiquement trouvées chez les violettes ont été notées: α-ionone (c), β-ionone (d) et dihydro-β-ionone (e), ainsi qu'un méthoxybenzène, le méthylanisole (f). Ce dernier est absent pour *V. suavis* mais il est suspecté la présence d'autres dérivés près de la zone bien identifiée des ionones, avec l'apparition de deux nouveaux pics. Enfin, les violettes identifiées comme *V. subs Rostratae* ne présentent pas de composés volatils détectables, comme en témoigne l'absence de pics chromatographiques, ce qui corrèle avec la très faible odeur des fleurs par rapport aux autres (Figure 26). Ces profils sont bien corrélés aux résultats de l'OPLS-DA montrant des niveaux élevés de similarité entre *V. suavis* et *V. alba* en raison de la présence d'ionones, contrairement aux deux autres espèces *V.sp*. Hearn et *V. subs Rostratae* dépourvu de ces substances. En ce qui concerne les *V. odorata*, leurs profils sont soit similaires à *V. alba*, soit à *V. suavis*. Elles ne semblent donc pas posséder une signature spécifique, ce qui justifie le chevauchement avec ces deux autres groupes.

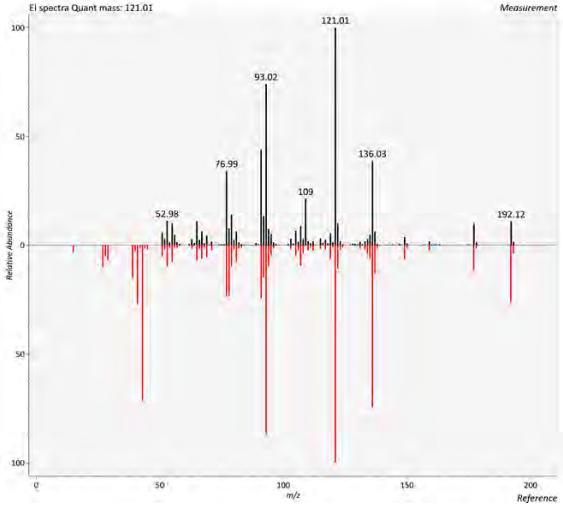
a) limonene (16,3 min)



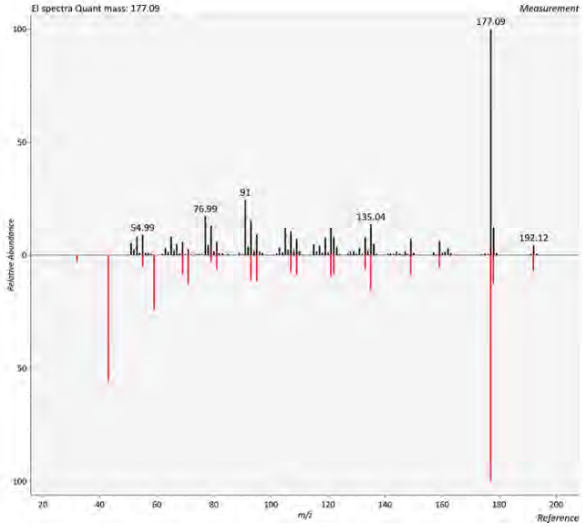
b) β-cis-ocimène (17,4 min)



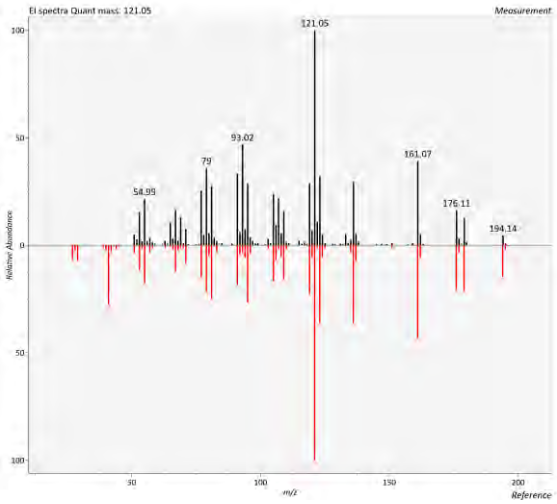
c) α-ionone (33,3 min)



d) β-ionone (35,8 min)



e) dihydro-β-ionone (33,7 min)



f) methylanisole (15,8 min)

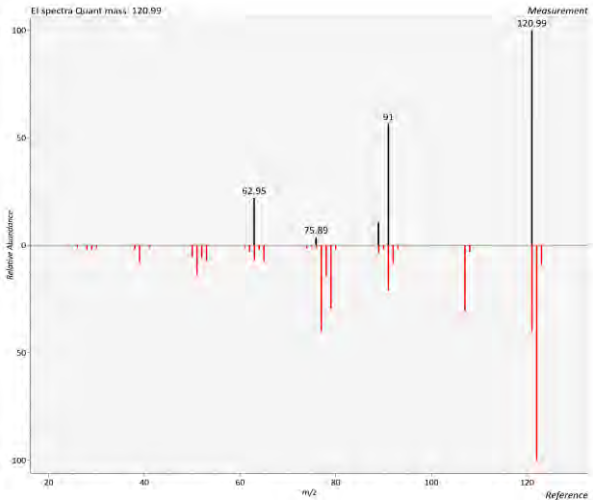


Figure 25. Identification *in silico* des composés volatils biomarqueurs de certaines espèces

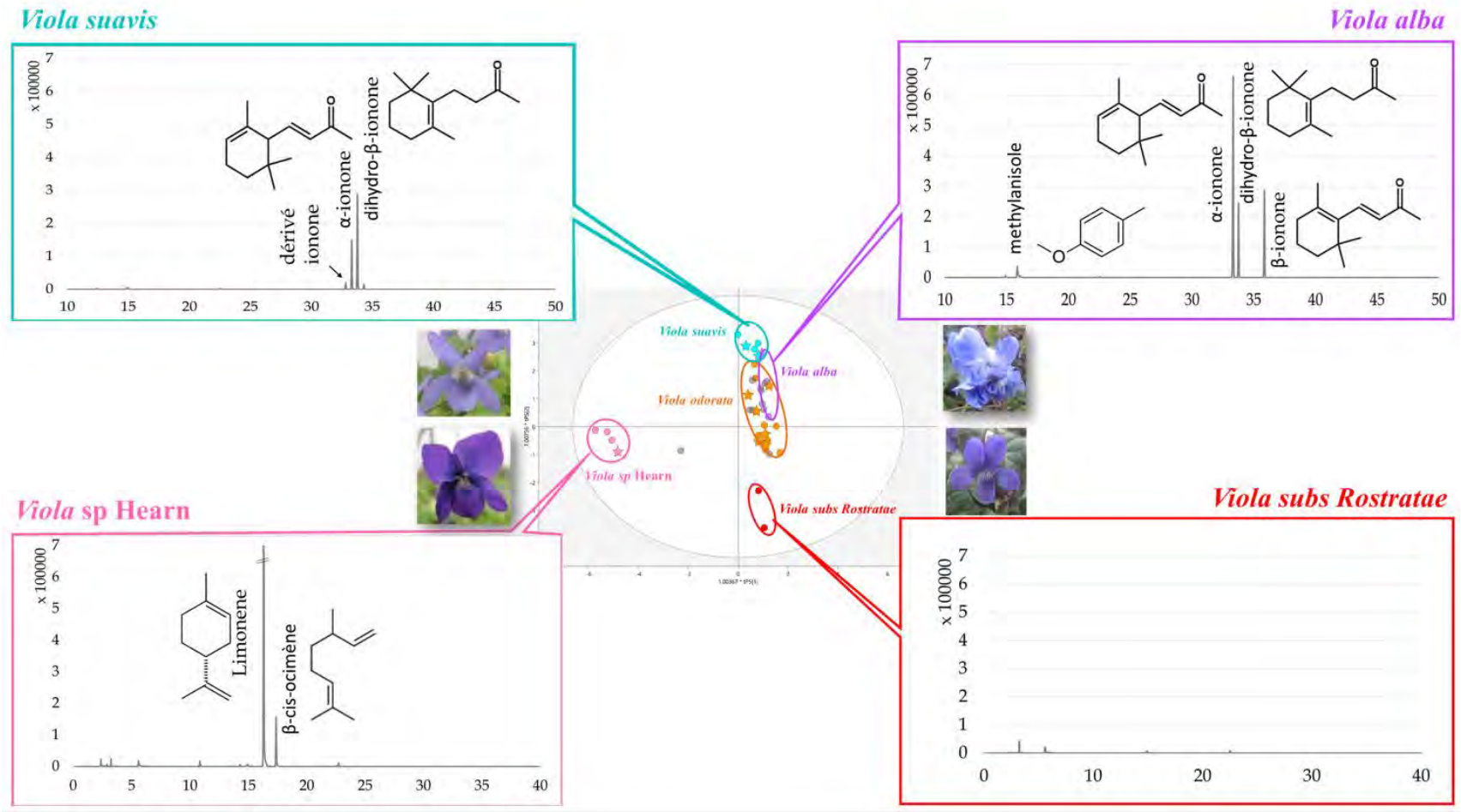


Figure 26. Mise en relation des profils volatils des violettes et de leur affiliation à une espèce particulière

L'injection de standards analysés avec la même méthode GC-MS a été entreprise afin de pouvoir valider ces résultats par comparaison des temps de rétention et des spectres de masse (annexe 3, page 219). L'écart de 7% entre les temps de rétention des standards et des composés issus des extraits de violette, ainsi que la forte similarité des spectres de masse expérimentaux permettent de valider la totalité des identifications proposées.

II.2.3.1.4 Étude des profils olfactifs

Les profils olfactifs des violettes ont été étudiés afin de mettre en évidence des critères complémentaires de différenciation entre espèces (cf. IV.5.1.2.2, page 176). En effet, une étude basée sur l'analyse des effluves odorantes d'espèces chinoises d'orchidées du genre *Cymbidium*, par l'utilisation d'un nez électronique (**Zhang et al., 2014**), a permis une bonne discrimination des espèces et cultivars. Dans cette étude, une analyse GC-MS couplée à un dispositif d'olfactométrie (GC-O/MS) a été mise en œuvre avec la participation de trois panélistes. Différents types d'odeur ont été détectés à partir des trente-six violettes analysées (Figure 27), dont cinq majoritaires : l'odeur florale et plus particulièrement de violette, une odeur plutôt sucrée rappelant le caramel, une odeur se rapprochant d'une sensation de brûlé, de cuit et enfin une odeur verte et boisée.

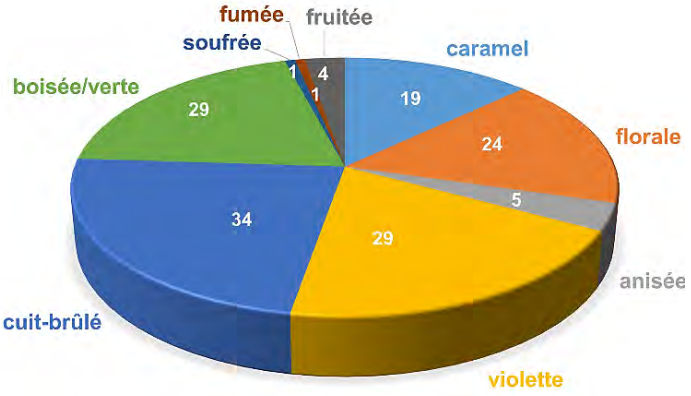


Figure 27. Types d'odeur détectés dans les échantillons de violette et nombre de violettes associées

Peu d'odeur correspondait à un pic chromatographique, mis à part l'odeur de violette (florale et verte) détectée pour les ionones (Figure 28). Ainsi, il semble que les fleurs de violettes soient olfactivement riches, mais les composés responsables de ces odeurs peu concentrés, ce qui n'a pas permis de les identifier.

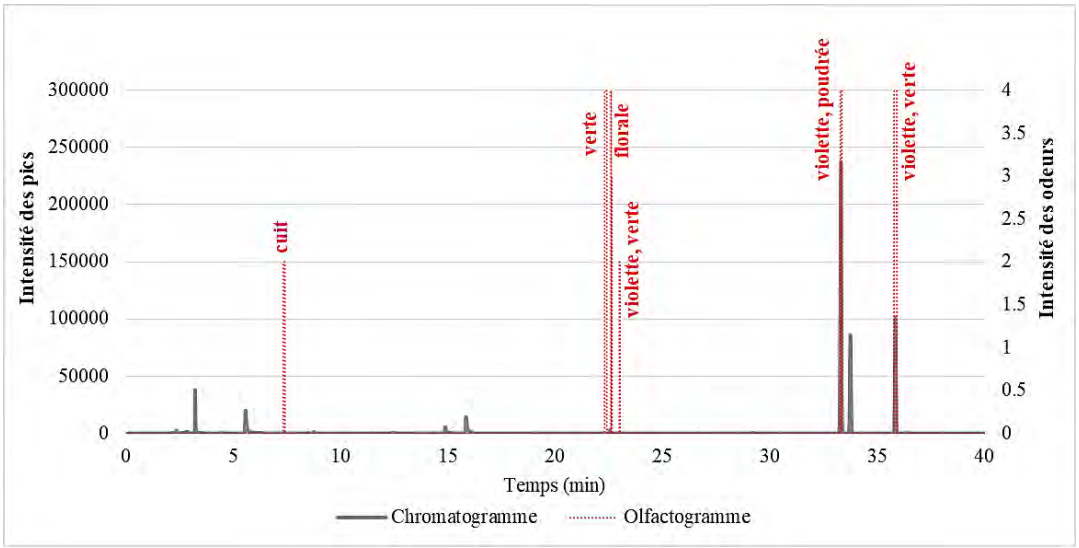


Figure 28. Superposition du chromatogramme GC et de l'olfactogramme pour la violette de Toulouse

Quelques différences mineures ont été constatées comme au niveau de l'odeur soufrée touchant une seule violette (n° 52) identifiée chez *V. odorata* et l'odeur fumée touchant une violette identifiée chez *V. suavis* (n° 57). Les odeurs anisée et fruitée sont retrouvées chez une violette de chaque espèce. Ces particularités ne sont pas généralisables à une espèce en particulier et de manière générale, aucune différence majeure n'a été observée entre les espèces. Par conséquent, les effluves odorantes semblent ici non pertinentes dans une optique de différenciation et classification des violettes.

II.2.3.2 Intégration des profils non volatils des parties aériennes

Les extraits bruts éthanoliques à 80% des parties aériennes de violettes, ont été réalisés sur la base d'un protocole commun pour limiter toutes variabilités de traitement (cf. IV.5.2.2.1, page 179). L'extraction hydroalcoolique a été choisie pour des raisons de commodités dans le traitement d'une collection. Un traitement à la silice C_{18} a été effectué pour adsorber les composés les plus apolaires et les profils avant et après traitement ont été analysés en UHPLC-MS (Figure 29 a). Il a été observé avant traitement un pic de très forte intensité en mode NI à RT = 15,95 min et m/z = 815,4990. L'identification de ce pic n'a pu être obtenue en raison de l'absence d'information de fragmentation MS/MS. Cependant, l'éclaircissement de la couleur verte de la solution de départ (Figure 29 b et c) indique que ce composé est très certainement un pigment, peut être un dérivé de la chlorophylle. Après traitement C_{18}, une diminution drastique de ce pic est constatée, sans pour autant observer d'impact sur les autres métabolites puisque les deux profils entre 0,5 et 15 min sont superposables. Quelques variabilités d'intensité sont parfois décelables et pourraient être expliquées par une légère rétention sur la silice, mais également par la préparation de l'échantillon en lui-même. Ainsi, afin de limiter les problèmes de rétention dans la colonne C_{18} pouvant endommager la colonne analytique et de se focaliser sur les composés du métabolisme secondaire, un traitement par la poudre de silice C_{18} a été envisagé dans une optique d'élimination des pigments et autres composés très apolaires.

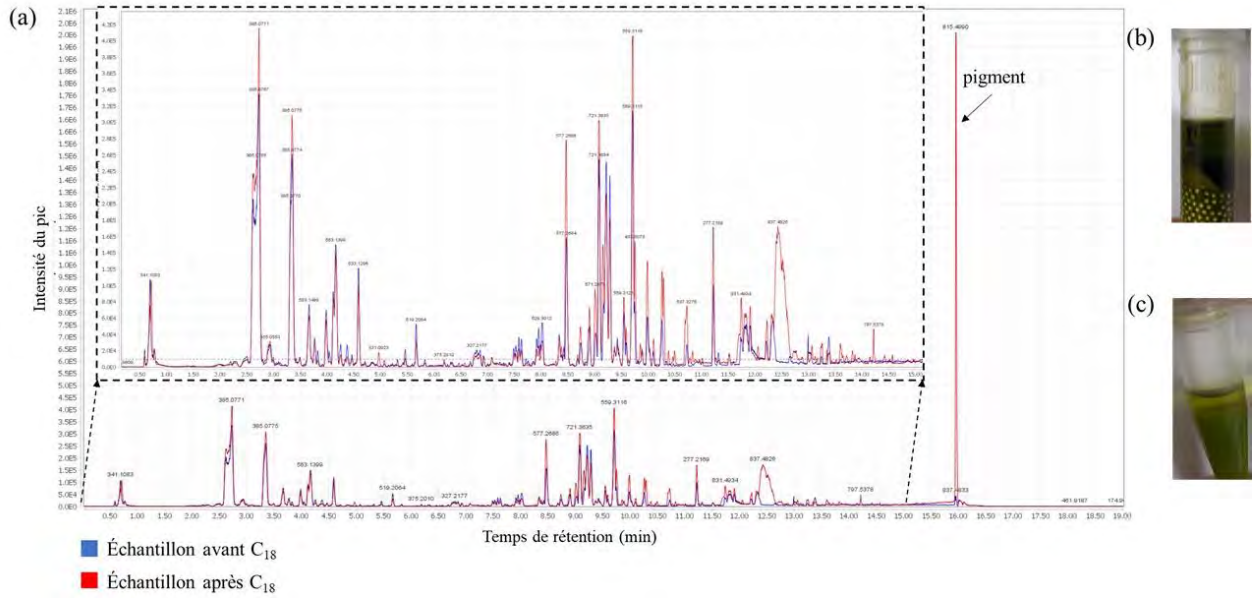


Figure 29. Comparaison du profil chimique de la violette de Toulouse avant et après traitement C_{18} (a) et couleur de l'extrait avant (b) et après traitement (c)

II.2.3.2.1 Analyse métabolomique des données UHPLC-HRMS/MS

La méthode analytique a été optimisée avant le début du projet de manière à obtenir le plus d'informations possible en un temps d'analyse satisfaisant (une vingtaine de minutes). Les quatres ions

les plus intenses de chaque scan ont été fragmentés pour apporter des informations structurales et faciliter l'annotation des composés. Les profils UHPLC-HRMS des 119 extraits (110 extraits bruts, 1 référence et 8 contrôles qualité (QC) : pool de tous les extraits) ont respectivement conduit à 208 et 198 variables en modes d'ionisation négatif (NI) et positif (PI) avec 120 variables communes (cf. IV.5.2.4.2, page 186). Un fichier combiné des données dans ces deux modes a été construit afin d'obtenir le profil complet des extraits pour les analyses statistiques. Comme attendu, l'ACP, avec les données logarithmées et l'échelle pareto, a groupé les QC près du centre de l'ACP avec des extraits de violette répartis autour (Figure 30). Des groupes corrélés avec les espèces d'après les résultats génétiques ont été mis en évidence. Ainsi, ces résultats indiquent une composition chimique stable dans les parties aériennes des violettes appartenant à la même espèce. En outre, on peut noter la formation de deux principaux groupes séparés par la composante PC1 avec d'une part, *V. odorata* et d'autre part, *V. suavis*, *V. sp.* Hearn, *V. alba* et *V. subs Rostratae*. Dans ce second groupe, *V. alba* et *V. subs Rostratae* semblent présenter des profils similaires, tandis que *V. sp* Hearn et *V. suavis,* qui sont tous deux bien séparés des autres groupes par la composante PC2, semblent posséder des métabolites discriminants.

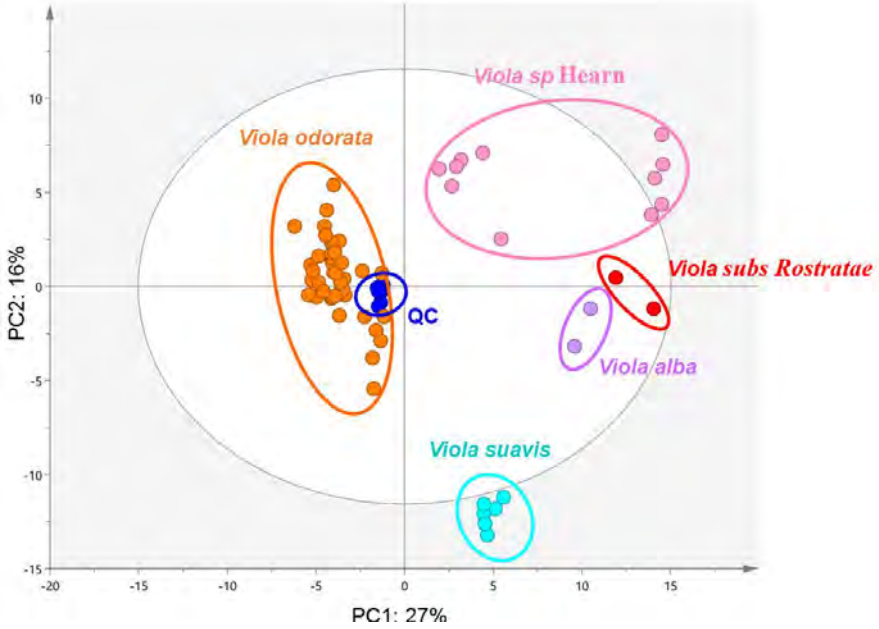


Figure 30. ACP des profils UHPLC-HRMS des extraits de parties aériennes des violettes de la collection

Après l'ACP, une OPLS-DA a été appliquée, afin d'essayer de classer les violettes non identifiées par projection et d'étendre la classification de la collection, comme effectué précédemment avec les fleurs. Le modèle de base construit avec les violettes génétiquement identifiées (Figure 31 a, N = 62) était bien défini (R2Y = 0,962, Q2Y = 0,822, $p < 0,001$) et validé par le test de permutation (Figure 31 b). La prédiction des violettes non identifiées (N = 46) a ensuite été effectuée. Quarante-deux des quarante-six violettes inconnues ont été affiliées à des groupes définis par des profils chimiques spécifiques, et ainsi statistiquement corrélées à une espèce identifiée (Figure 31 c). Trois violettes ont été associées à *V. alba*, deux à *V. subs Rostratae*, neuf à *V. sp* Hearn, deux à *V. suavis* et vingt-six à *V. odorata*. Quatre violettes sont restées totalement séparées de ces groupes et n'ont donc pu être ni classées, ni identifiées.
Dans les précédents résultats génétiques, certaines violettes n'avaient pas été clairement identifiées en raison de scores de similarité insuffisants, mais ont tout de même été affiliées à un groupe par l'intermédiaire du modèle de distance. Par comparaison avec les résultats UHPLC-HRMS, les prédictions par profil chimique recroisent les affiliations génétiques, renforçant ainsi l'identification hypothétique de ces violettes, qui a permis d'atteindre la classification de 96% de la collection (Figure 31 d). Concernant les quatre violettes non définies, deux n'ont pas été séquencées et les autres ont révélé

une séquence claire avec de faibles scores d'identification. Ainsi 4% de la collection de violettes restent non identifiés et pourraient correspondre à des espèces non décrites.
Par ailleurs, les neuf violettes présentant un fort polymorphisme génétique ont toutes été affiliées à un groupe : 2 chez les *V. alba* (V19 et V37), 1 chez les *V. suavis* (V71) et 5 chez les *V. odorata* (V12, V39, V57, V71 et V91). Les profils chimiques ont ainsi permis de définir une classification là où la génétique a échoué, et ces résultats pourraient alors indiquer que ces violettes sont des hybrides de ces espèces.

La stabilité du métabolome des feuilles au cours du temps a été évaluée par une étude cinétique. Tous les deux mois, dix feuilles de violette de Toulouse ont été collectées et extraites sur le même principe que précédemment. La superposition des profils UHPLC a permis de constater que les profils étaient très similaires, avec parfois quelques variations d'intensité. Une ACP a été réalisée en intégrant quelques échantillons de chaque espèce de la collection (Figure 32). Tous les extraits cinétiques de la violette de Toulouse (*V. alba*) ont été distinctement projetés dans ce groupe, quelle que soit la période de récolte. Ainsi, la validité d'un modèle statistique obtenue à un moment précis de la saison peut être ici plus généralement attestée par ces résultats cinétiques, qui reflètent une similarité qualitative des extraits avec tout de même quelques variations quantitatives.

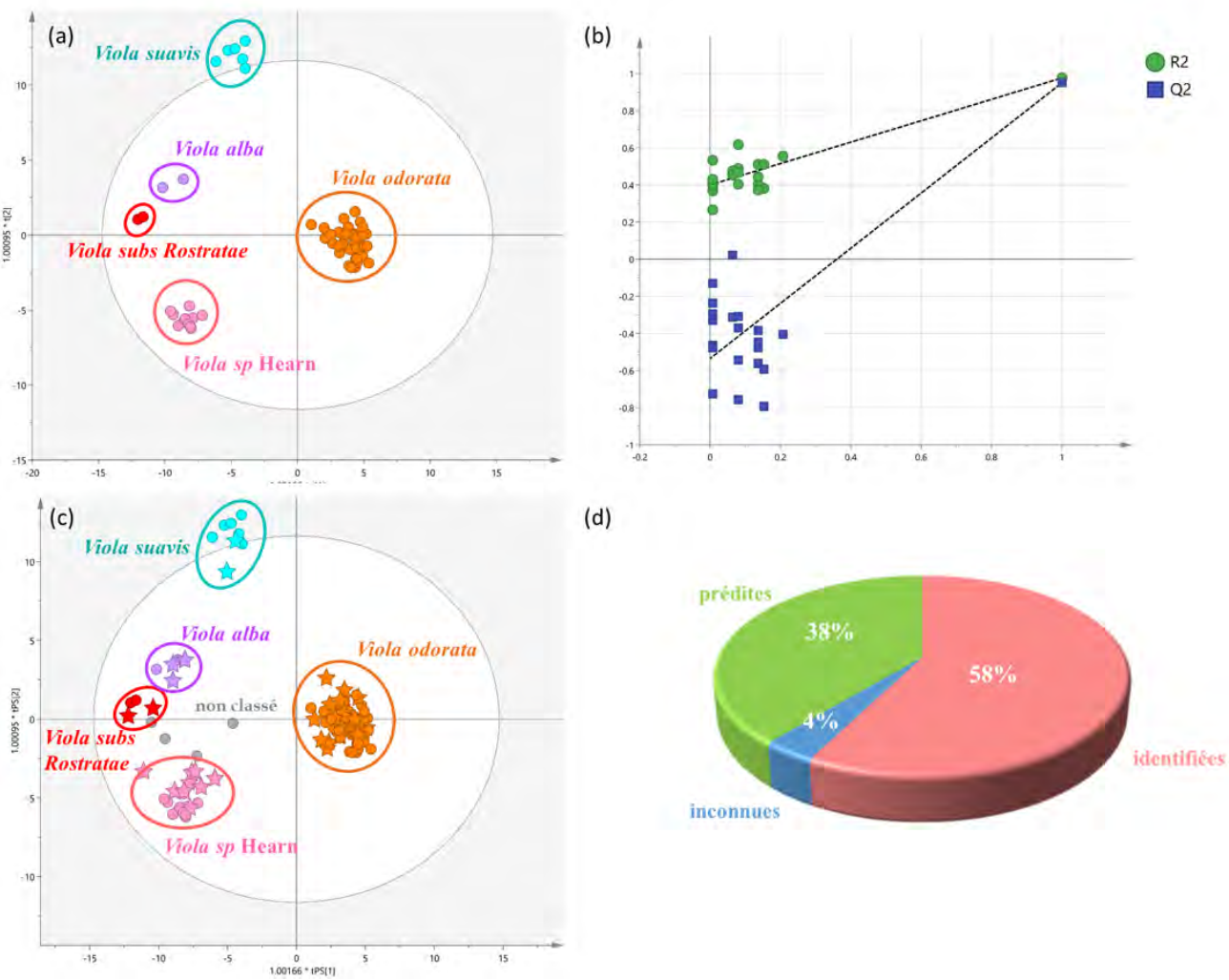


Figure 31. Modèle OPLS-DA des violettes identifiées (a), test de permutation (b), prédiction des violettes inconnues matérialisées par une étoile (c) et répercussions sur la classification de la collection (d)

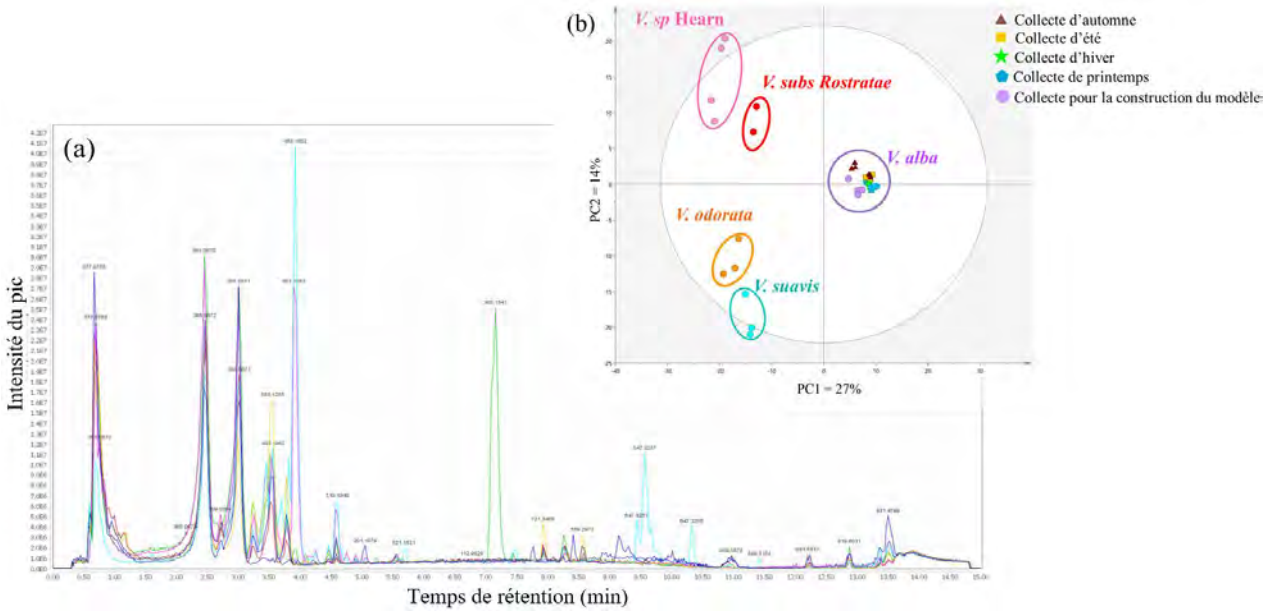
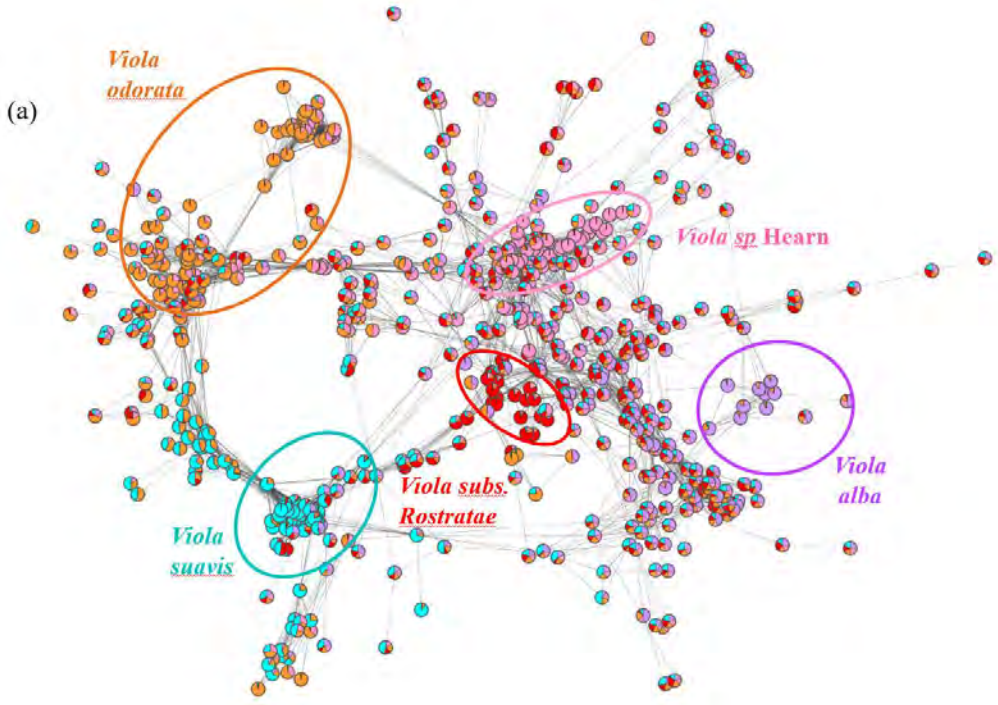


Figure 32. Profils UHPLC des différents extraits (a) et analyse en composantes principales relatifs à l'étude cinétique des métabolites secondaires de la violette de Toulouse

II.2.3.2.2 Identification de biomarqueurs des parties aériennes

Dans le but d'approfondir ces résultats et d'identifier de potentiels biomarqueurs chimiotaxonomiques, un réseau de corrélation a été construit (cf. IV.4.4, page 173). Celui-ci est composé d'un ensemble de nœuds correspondant aux pics UHPLC-HRMS et corrélés par des arêtes (Figure 33 a). Cette corrélation est basée sur l'intensité relative de chaque pic dans chaque extrait (corrélation de pearson > 0,9). Les composés ayant la même tendance sont ainsi corrélés ensemble. Le réseau ainsi obtenu est composé d'environ 250 nœuds et a permis d'observer, des amas clairs de métabolites spécifiques à des espèces identifiées. Des métabolites non discriminants, partagés par différentes espèces et parfois toutes, ont été identifiés au centre du réseau.



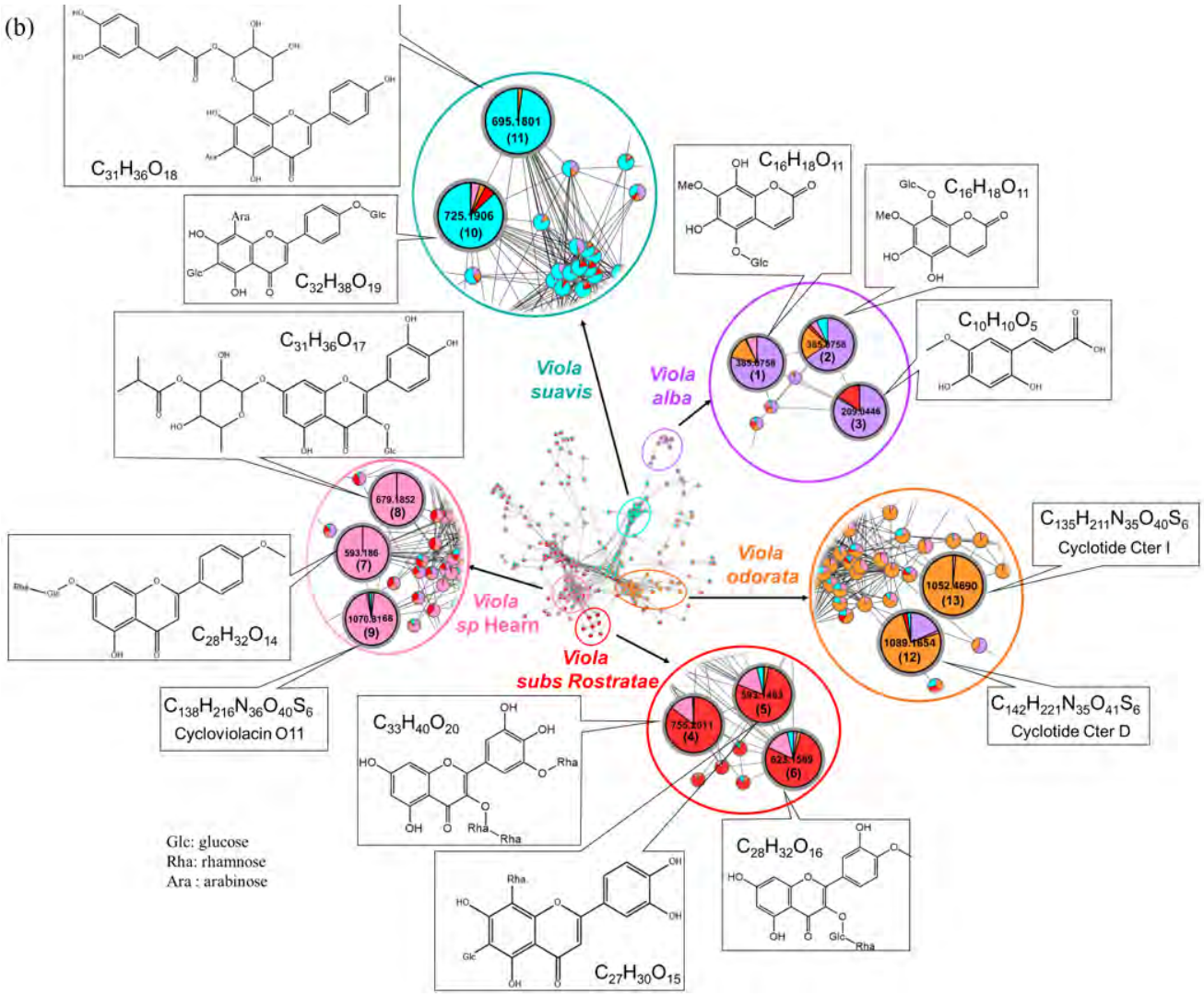


Figure 33. Réseau de corrélation basé sur l'ensemble des données UHPLC-HRMS en modes NI et PI des extraits de violette (a) et identification de biomarqueurs caractéristiques (b).

L'identification des composés caractéristiques de chaque espèce (Figure 33 b) a été effectuée en utilisant la fragmentation *in silico* avec nos propres bases de données *Viola,* ainsi quedes bases de données locales du logiciel MS-FINDER (cf. IV.5.2.4.4, page 187). Les résultats sont présentés dans le Tableau 16. L'acquisition de profils UHPLC-HRMS a fourni des rapports m/z précis indispensables pour la détermination de la formule brute et les fragments MS/MS ont été utilisés pour la détermination de la structure. La base de données *Viola* fournit des identifications plus pertinentes que celles des bases de données de produits naturels de MS-FINDER, qui conduisent à une identification plus générale et pour lesquelles essentiellement la classe chimique reste pertinente. Toutes les correspondances entre les fragments expérimentaux et *in silico* sont présentées en annexe 4 page 221.

Les coefficients OPLS-DA ont été déterminés pour chaque variable et pour chaque espèce. Les composés caractéristiques ont été corrélés aux scores les plus élevés des coefficients. *V. odorata* présente principalement des cyclotides en tant que biomarqueurs comme en témoignent l'identification des ions pseudomoléculaires $[M+3H]^{3+}$ et $[M+4H]^{4+}$, ainsi que la forme isotopique caractéristique (**Hashempour et al., 2013**). *V. alba*, *V. subs Rostratae*, *V. sp* Hearn et *V. suavis* semblent plutôt caractérisés par des polyphénols, en particulier des flavonoïdes *C*- et *O*-glycosylés mais aussi des coumarines spécifiques à *V. alba*. Comme dans l'approche métabolomique par UHPLC-HRMS mentionnée ci-dessus, l'identification de deux groupes a été renforcée par les biomarqueurs spécifiques.

Tableau 16. Résumé des composés dérépliqués à l'aide de MS-FINDER et du réseau moléculaire

N°	m/z	RT (min)	Mode	Formule brute	Erreur (ppm)	Identification potentielle[a]	Classe chimique	coefficient OPLS	Affiliation
1	385,0758	2,95	$[M-H]^-$	$C_{16}H_{18}O_{11}$	4,6744	7-Methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside[b]*	Coumarine	0,0486768	*V. alba*
2	385,0758	3,47	$[M-H]^-$	$C_{16}H_{18}O_{11}$	4, 6744	7-Methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[b]*	Coumarine	0,040651	*V. alba*
3	209,0446	3,32	$[M-H]^-$	$C_{10}H_{10}O_5$	4,3053	3-(2,4,5-Trihydroxyphenyl)-2-propenoic acid; 5-Me ether*	Aromatique simple	0,0474672	*V. alba*
4	755,2011	3,92	$[M-H]^-$	$C_{33}H_{40}O_{20}$	3,8625	3,3',4',5,5',7-Hexahydroxyflavone; 3-*O*-[α -L-Rhamnopyranosyl-(1,2)-α -L-rhamnopyranoside], 3'-*O*-α -L-rhamnopyranoside*	Flavonoïde	0,0938658	*V. subs Rostratae*
5	593,1483	4,46	$[M-H]^-$	$C_{27}H_{30}O_{15}$	4,8891	6-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxyflavone; 2''-*O*-α-L-Rhamnopyranosyl*	Flavonoïde	0,0619192	*V. subs Rostratae*
6	623,1589	4,63	$[M-H]^-$	$C_{28}H_{32}O_{16}$	4,4932	Isorhamnetin 3-neohesperidoside*	Flavonoïde	0,0909114	*V. subs Rostratae*
7	593,1860	5,42	$[M+H]^+$	$C_{28}H_{32}O_{14}$	0,8429	5,7-Dihydroxy-4'-methoxyflavone; 7-*O*-[α-L-Rhamnopyranosyl-(1→6)-β-D-glucopyranoside]	Flavonoïde	0,0825076	*V. sp* Hearn
8	679,1852	5,71	$[M-H]^-$	$C_{31}H_{36}O_{17}$	3,8281	Quercetin 3,7-diglycosides; 3-*O*-β-D-Glucopyranoside, 7-*O*-[2-methylpropanoyl-(→3)-α-L-rhamnopyranoside]*	Flavonoïde	0,0587476	*V. sp* Hearn
9	1070,8168 803,3678	5,94	$[M+3H]^{3+}$ $[M+4H]^{4+}$	$C_{138}H_{216}N_{36}O_{40}S_6$	1,0273	Cycloviolacine O11*	Cyclotide	0,0574069	*V. sp* Hearn
10	725,1906	3,31	$[M-H]^-$	$C_{32}H_{38}O_{19}$	3,8610	6-*C*-α-L-Arabinopyranosyl-8-C-β-D-glucopyranosyl-4',5,7-trihydroxyflavone; 4'-O-β-D-Glucopyranoside*	Flavonoïde	0,0430546	*V. suavis*

11	695,1801	3,52	$[M-H]^-$	$C_{31}H_{36}O_{18}$	4,0277	3,4',5,7-Tetrahydroxy-3',8-dimethoxyflavone; 7-*O*-[α-L-Rhamnopyranosyl-(1→2)-6-*O*-acetyl-β-D-glucopyranoside*	Flavonoïde	0,0571469	*V. suavis*
12	1089,1653 817,1260	5,9	$[M+3H]^{3+}$ $[M+4H]^{4+}$	$C_{142}H_{221}N_{35}O_{41}S_6$	3,4889	Cter D**	Cyclotide	0,0403016	*V. odorata*
13	1052,4690	5,42	$[M+3H]^{3+}$	$C_{135}H_{211}N_{35}O_{40}S_6$	1,4252	Cter I**	Cyclotide	0,0533443	*V. odorata*

[a] Premier composé déterminé par fragmentation *in silico* avec MS-FINDER

[b] Composés confirmés par RMN

* Composés trouvés en interrogeant la base de données *Viola* extraite à partir du DNP

** Composés trouvés en interrogeant les bases de données généralistes de MS-FINDER

II.2.3.3 Combinaison des données ITS et UHPLC-MS

Un arbre consensus basé sur les séquences ITS et la matrice UHPLC-MS binarisée a validé la majorité des groupes précédemment obtenus (Figure 34). Cependant, pour *V. odorata*, une scission a été constatée conduisant à deux groupes au lieu du seul observé avec la séquence ITS. La première hypothèse pouvant expliquer ce phénomène concerne la perte potentielle d'informations due à la binarisation des données UHPLC-MS. En effet, ce sont des données continues avec des informations semi-quantitatives à l'échelle d'une variable. Cependant, la classification ascendante hiérarchique (CAH) des données UHPLC-MS (Figure 35) a mis en évidence une scission au sein des *V. odorata* ainsi que des *V. sp* Hearn, qui explique l'étendue de ce groupe sur l'ACP (cf. Figure 30, page 64**)** avec la visualisation de deux groupes. Ces observations indiquent ainsi la présence de deux chémotypes différents au sein de ces espèces pouvant révéler la présence potentielle de plusieurs variétés.

La comparaison des deux groupes de *V. odorata* a mis en évidence une concentration plus élevée de la cycloviolacine O9 (14) et d'un cyclotide non identifié (15) dans le groupe majoritaire des *V. odorata* O2. Celui minoritaire O1 présente quant à lui une teneur plus forte en diterpène Reveromycin L. La comparaison des deux groupes de *V. sp* Hearn a mis en évidence une teneur supérieure en flavonoïdes 17 et 18 dans le groupe H2 et en flavonoïdes 4 et 19 dans le groupe H1 (Tableau 17).

Tableau 17 : Composés caractéristiques des deux chémotypes de *V. odorata* et *V. sp* Hearn

N°	m/z	RT (min)	Formule brute	Identification*	Classe chimique	Affiliation
14	$[M+H]^{2+}$ 1047,1316	5,17	$C_{133}H_{207}N_{37}O_{39}S_6$	Cycloviolacine O9	Cyclotide	*V. odorata* O1
15	$[M+H]^{2+}$ 1057,7927	5,31	ND	ND	Cyclotide	*V. odorata* O1
16	$[M-H]^-$ 660,351	8,47	$C_{36}H_{52}O_{11}$	Reveromycin L	Diterpène	*V. odorata* O2
17	$[M-H]^-$ 739,2076	3,51	$C_{33}H_{40}O_{19}$	Kaempferol 3-glycosides; 3-*O*-[α-L-Rhamnopyranosyl-(1→4)-α-L-rhamnopyranosyl-(1→6)-β-D-galactopyranoside]	Flavonoïde	*V. sp* Hearn H2
18	$[M-H]^-$ 625,1381	3,91	$C_{27}H_{30}O_{14}$	Quercetin 3-glycosides; 3-*O*-[β-D-Glucopyranosyl-(1→6)-β-D-glucopyranoside]	Flavonoïde	*V. sp* Hearn H2
4	$[M-H]^-$ 755,2006	3,92	$C_{33}H_{40}O_{20}$	3,3',4',5,5',7-Hexahydroxyflavone; 3-*O*-[α-L-Rhamnopyranosyl-(1,2)-α-L-rhamnopyranoside], 3'-*O*-α-L-rhamnopyranoside	Flavonoïde	*V. sp* Hearn H1
19	$[M-H]^-$ 533,1274	4,14	$C_{25}H_{26}O_{13}$	6,8-Di-*C*-α-L-arabinopyranosylapigenin	Flavonoïde	*V. sp* Hearn H1

* Composés trouvés en interrogeant la base de données *Viola* extraite à partir du DNP

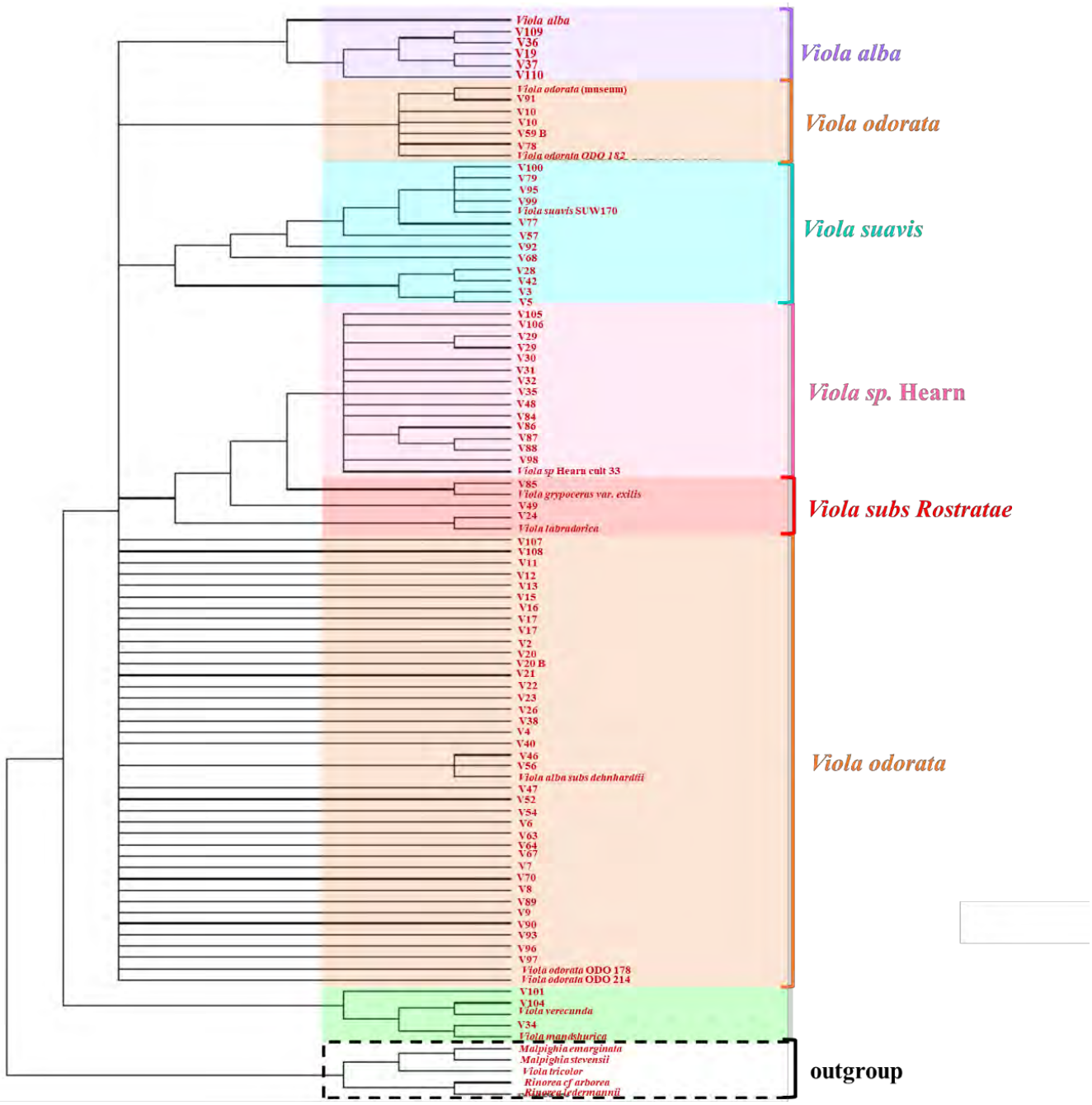


Figure 34. Arbre consensus combinant les séquences ITS et les données UHPLC-MS construites avec la méthode de parcimonie.

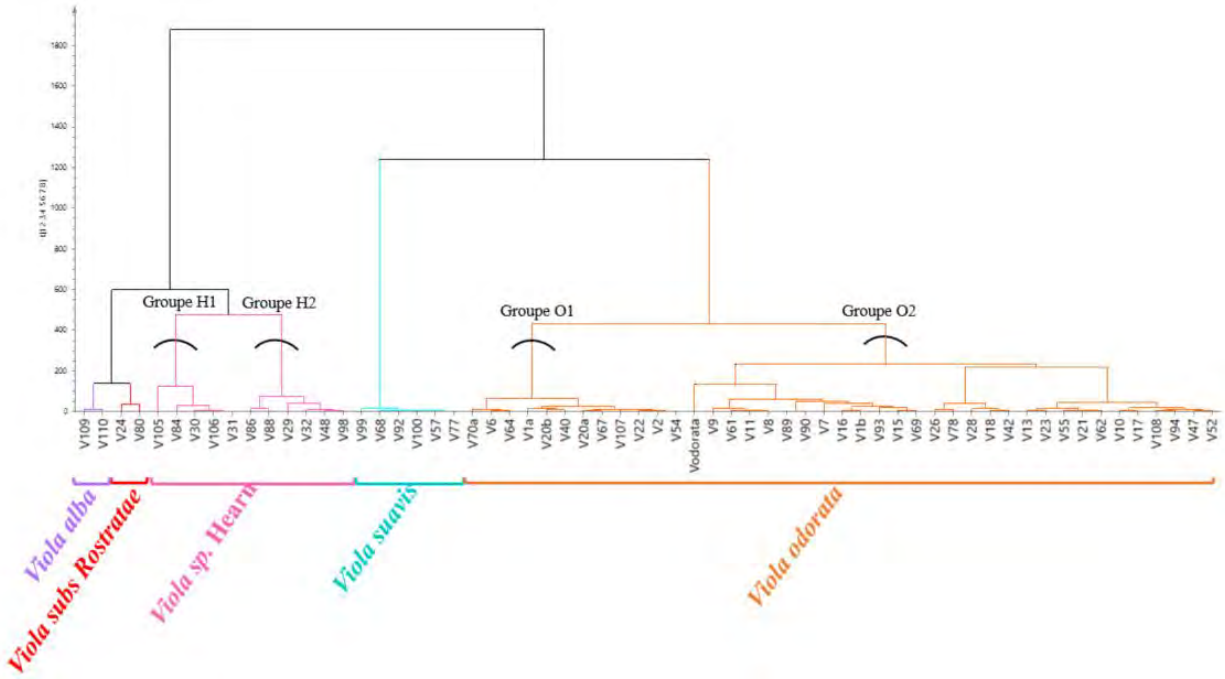


Figure 35. Analyse CAH basée sur l'ACP des données UHPLC confirmant les cinq groupes distincts corrélés aux espèces et la scission des *V. sp* Hearn (H1 et H2) et *V. odorata* (O1 et O2) en deux groupes

L'analyse indépendante des données ITS n'a pas permis d'identifier les variétés, mais les données chimiotaxonomiques ont permis de mettre en évidence la présence de chémotypes, renforçant ainsi la complémentarité des méthodes et la pertinence des profils métabolomiques dans la classification. Des analyses botaniques ont été entreprises sur quelques échantillons de chaque groupe afin de vérifier et valider les identifications obtenues par analyse phylogénétique en comparant les critères morphologiques et anatomiques étudiés à ceux de la littérature.

II.2.4 MONOGRAPHIE VEGETALE

II.2.4.1 Description des espèces

Les Violacées se reconnaissent à leurs fleurs irrégulières disposées en verticilles autour du pistil et possédant cinq étamines libres. Elles présentent cinq sépales inégaux, libres ou pratiquement libres, disposés symétriquement par rapport à un plan vertical, et dont l'un est prolongé en un éperon plus ou moins allongé en forme de cornet. Il y a très souvent chez les violettes deux types de fleurs : celles aux pétales développés qui ne produiront pas de fruits excepté les dernières formées et celles à fleurs plus tardives restant fermées, sans pétales et produisant des fruits. Ce sont des plantes à feuilles pétiolées en forme de cœur renversé munies de stipules dont les fleurs isolées sont généralement portées par de longs pédoncules comportant deux bractéoles (**Beille, 1909**).

D'après les résultats obtenus précédemment par l'étude génétique, huit espèces ont été identifiées (Tableau 18). On a pu constater que l'intégralité des noms d'espèces définis par l'étude génétique et la base de données Nucléotides de NCBI sont des noms acceptés excepté deux : *V. verecunda* qui est en réalité synonyme de *V. arcuata* Blume et *V. sp* Hearn, espèce non identifiée, mais nommée comme telle car étudiée par David J. Hearn. Cinq de ces espèces sont répertoriées dans la flore européenne, trois dans la flore chinoise et trois dans celle américaine.

Tableau 18. Résumé des synonymes et origine des espèces identifiées

Nom accepté	Synonymes	Origine
Viola alba Besser	*Viola alba subsp.cretica* (Boiss. & Heldr.) Marcussen *Viola alba subsp.dehnhardtii* (Ten.) W.Becker *Viola alba subsp.scotophylla* (Jord.) Nyman *Viola pontica* W.Becker	Européenne*
Viola suavis Bieberstein	*Viola segobricensis* Pau *Viola sepincola* Jord. *Viola suavis var.segobricensis* (Pau) O.Bolòs & Vigo	Européenne
Viola odorata Linnaeus	*Viola odora* Neck. *Viola odorata* f.odorata *Viola wiedemannii* Boiss.	Européennne Chinoise ** Américaine***
Viola arcuata Blume	*Viola alata* Burgersd. *Viola alata subsp. verecunda* (A. Gray) W. Becker *Viola amurica* W. Becker *Viola arcuata f. radicans* (Makino) Nakai *Viola arcuata var. verecunda* (A. Gray) Nakai *Viola distans* Wall. *Viola excisa* Hance *Viola hupeiana* W. Becker *Viola verecunda* A. Gray *Viola verecunda f. hensoaensis* Kudô & Sasaki *Viola verecunda f. radicans* Makino *Viola verecunda var. semilunaris* Maxim.	Européenne
Viola mandshurica Becker	*Viola alisoviana f. intermedia* (Kitag.) Taken. *Viola hsinganensis* Taken. *Viola mandshurica f. albiflora* P.Y. Fu & Y.C. Teng *Viola mandshurica var. ciliata* Nakai *Viola mandshurica f. ciliata* (Nakai) F. Maek. *Viola mandshurica var. gla*bra Nakai *Viola mandshurica f. glabra* (Nakai) Hiyama ex Maek. *Viola mandshurica f. macrantha* (Maxim.) Nakai & Kitag. *Viola oblongosagittata f. ishizakii* Yamam. *Viola patrinii f. glabra* (Nakai) F. Maek. *Viola patrinii var. macrantha* Maxim. *Viola philippica f. intermedia* (Kitag.) Kitag. *Viola rhodosepala* Kitag. *Viola yedoensis f. intermedia* Kitag.	Européenne Chinoise
Viola grypoceras A Gray	*Viola grayi var.candida* H. Boissieu *Viola leveillei* H. Boissieu *Viola sylvestris var.candida* (H. Boissieu) H. Lév. *Viola sylvestris var.grypoceras* (A. Gray) Maxim.	Chinoise
Viola labradorica Schrank	-	Américaine
Viola sp Hearn	-	Américaine

* Identifiée dans The Plant List ; ** Identifiée dans Flora of China ; *** Identifiée dans Flora of North America

La reconnaissance d'une plante est basée sur la caractérisation morphologique des organes reproducteurs, et à un degré moindre de la partie végétative. Des descriptions morphologique et anatomique ont été réalisées ici en s'appuyant sur les données de la littérature, de manière à pouvoir mettre en évidence d'éventuels critères de différenciation, renforçant les précédentes séparations en groupes par analyses génétique et chimiotaxonomique. Les fleurs et les feuilles des espèces de violettes ont été morphologiquement comparées les unes aux autres. Il a été constaté au niveau des fleurs des

différences phénotypiques au sein d'une même espèce (Figure 36), attestant de la probable présence de variétés. L'étude des grains de pollen n'a pas révélé de divergences significatives entre les espèces de violettes (annexe 5, page 224), mais permet seulement de différencier les violettes des pensées. Une étude microscopique réalisée à partir de poudre de plante fleurie des espèces de violette et de pensées (*V. arvensis* et *V. tricolor*), a permis de mettre en évidence que les violettes présentent des grains de pollen de l'ordre de 50 µm à trois pores (triangulaires), contrairement aux pensées qui présentent des grains de l'ordre de 50 µm à quatre (quadrangulaires), ou cinq (polygonaux) pores (Figure 37). Ce phénomène de polymorphisme des grains de pollen des espèces du genre *Viola* a notamment été mis en évidence par Wodehouse (**Wodehouse, 1935**).

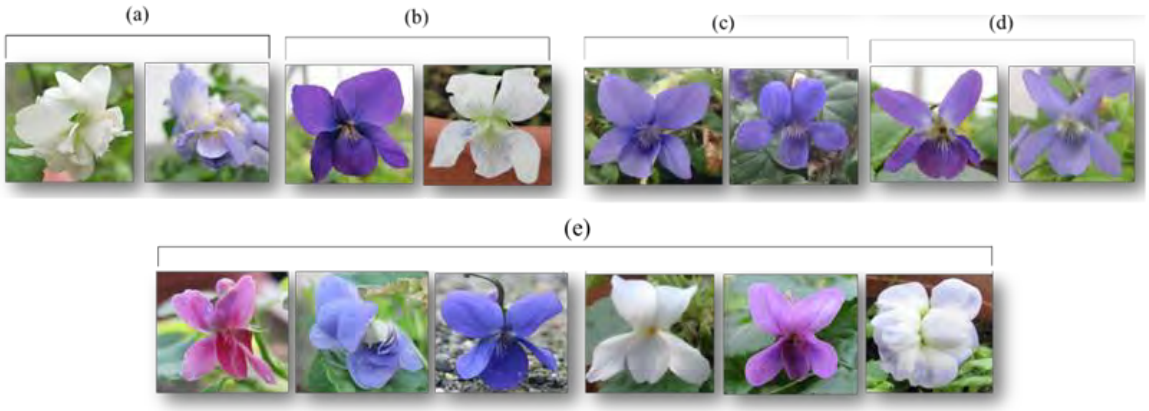


Figure 36 : Phénotypes de fleurs appartenant à *V. alba* (a), *V. sp* Hearn (b), *V. subs Rostratae* (c), *V. suavis* (d), et *V. odorata* (e)

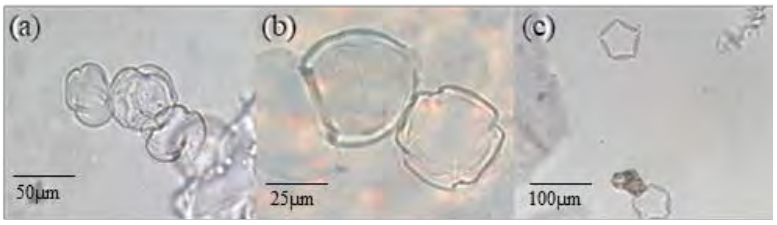


Figure 37. Grains de pollen de *V. alba* (a) ; *V. tricolor* (b); *V. arvensis* (c)

C'est pourquoi il a été décidé d'étudier au niveau microscopique deux principaux organes (cf. IV.2, page 167), à savoir les bractéoles situées au niveau du pédoncule floral ainsi qu'un échantillon d'épiderme prélevé au niveau de la feuille, afin de valider les identifications d'espèce précédemment obtenues par analyse phylogénétique.

II.2.4.2 Analyse morphologique

II.2.4.2.1 L'inflorescence

II.2.4.2.1.1 Pédoncule floral

Chez les violettes ayant été identifiées comme *V. alba*, les bractéoles présentent une taille moyenne comprise entre 0,6 et 0,8 cm et sont relativement éloignées de la fleur avec un pédoncule de longueur comprise entre 6 et 7 cm (Figure 38 a). Celles-ci sont de forme lancéolée (fer de lance), large et dentée à la base. Ces caractéristiques sont globalement retrouvées chez les violettes identifiées comme *V. sp* Hearn (Figure 38 b). En revanche, chez les violettes ayant été identifiées comme *V. suavis* et *V. odorata*, les bractéoles présentent une taille moyenne plus petite atteignant au maximum 0,5 à 0,6 cm et sont plus

proches de la fleur avec un pédoncule de 2-3 cm pour le premier (Figure 38 c) et 4-5 cm pour le second (Figure 38 d). Celles-ci sont de même forme lancéolée, étroite et non dentée. Par ailleurs, elles présentent une couleur plus pâle que celle de *V. alba* de couleur bien verte. Quant aux violettes de la sous-section *Rostratae*, ce sont celles qui possèdent les bractéoles les plus petites atteignant au maximum 0,5 cm et situées très proche de la fleur avec un pédoncule de 1-2 cm. Elles sont cependant toujours de forme lancéolée, étroite et dentée pour *V. grypoceras* (Figure 38 e et f).

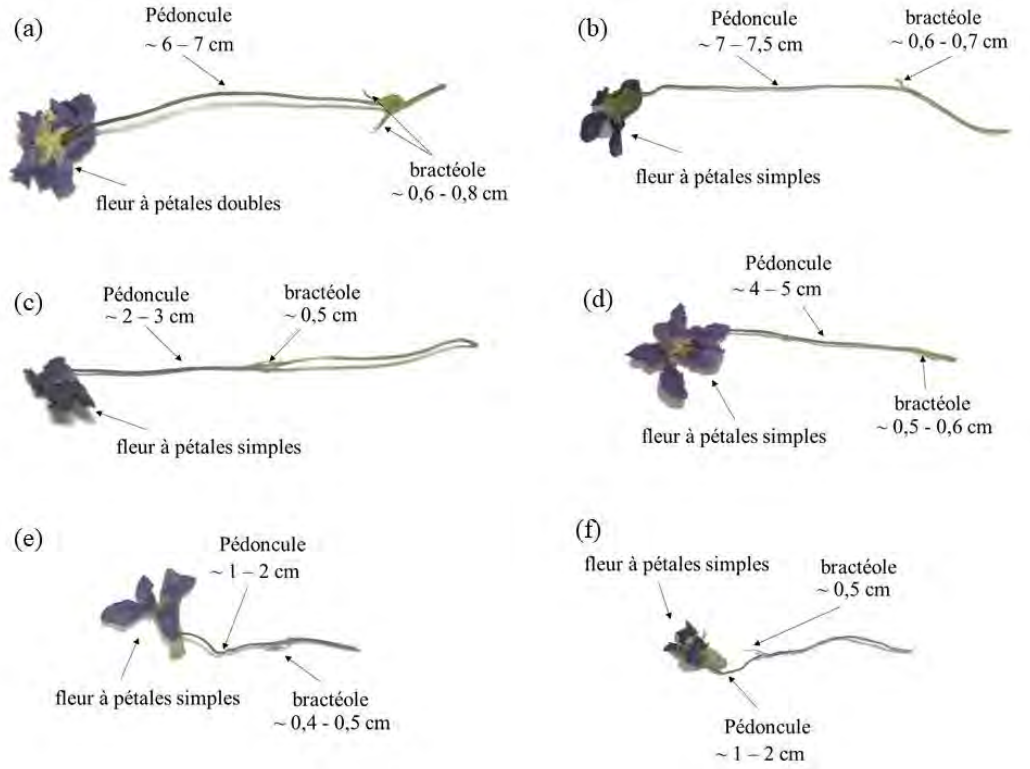


Figure 38 : Description du pédoncule floral de *V. alba* (a), *V. sp* Hearn (b), *V. suavis* (c), *V. odorata* (d), *V. labradorica* (e) et *V. grypoceras* (f)

II.2.4.2.1.2 La fleur

D'après Bonnier (**Bonnier et Douin, 1990**), les violettes appartenant à *V. odorata* se reconnaissent généralement à leurs pétales dont seul l'inférieur est très échancré, tandis que les quatre autres le sont beaucoup moins au sommet. On constate que c'est bien le cas pour les deux violettes identifiées comme telles (Figure 39 a) ainsi que pour *V. alba* (Figure 39 b) et *V. suavis* (Figure 39 c). Ces observations permettent de comprendre pourquoi Bonnier les avait considérées comme des dérivés de *V. odorata*. Leurs fleurs sont plus ou moins odorantes de couleur violette, lilas, violacée ou même blanche. *V. alba*, quant à elle, possède des fleurs un peu odorante à pétales blancs, jaunes vers la base ou bien bleus violacé voire violets avec du blanc vers la base. Les quatre pétales supérieurs sont inclinés en dehors et celui inférieur très largement échancré. *V. suavis* présente des fleurs odorantes d'un bleu violet avec du blanc dans la moitié inférieure. Leurs fleurs mesurent en moyenne entre 1 et 2 cm (**Tutin et al., 1968**). La présence de pétales doubles pour *V. alba* dans le cas présent (Figure 39 b) indique probablement la présence d'un mutant.

D'après la flore de Chine (**Cai Ke, 2007**), *V. grypoceras* présente des fleurs violacées à blanches de forme oblongue (plus longue que large à bords parallèles) à obovale (de forme ovale mais élargie vers l'apex), de l'ordre de 1-1,5 cm (Figure 39 d.1).

D'après la Flore d'Amérique du Nord **(Little et McKinney, 2016**), *V. labradorica* dispose de petites fleurs de l'ordre de 1 cm aux pétales violet lavande voire violet sur toute la surface (Figure 39 d.2).

Pour les violettes identifiées comme *V. sp* Hearn, des fleurs de grande taille sont observées, entre 2 et 3 cm, de couleur violet foncé avec un cœur blanc (Figure 39 e).

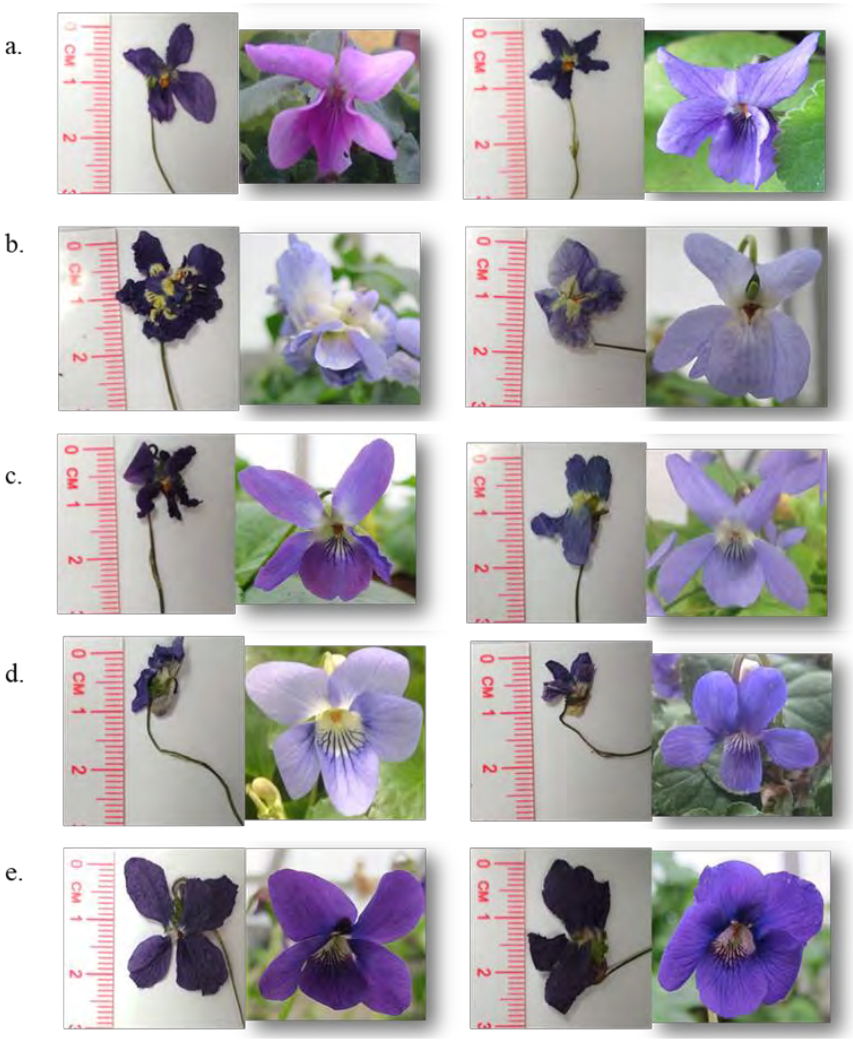


Figure 39. Fleurs séchées (1ère photo) et fleurs fraiches (2nd photo) pour deux violettes de *V. odorata* (a), *V. alba* (b), *V. suavis* (c), *V. subs Rostratae* (d.1 : *V. grypoceras* et d.2 : *V. labradorica*) (d) et *V. sp* Hearn (e)

II.2.4.2.2 La feuille

Les violettes présentent majoritairement des feuilles simples pétiolées ovales à rondes avec une base cordée, c'est-à-dire en forme de cœur. Elles présentent un limbe réniforme avec une extrémité supérieure plutôt aigüe et un bord dentelé. On peut constater que pour *V. odorata* et les espèces apparentées, *V. alba* et *V. suavis*, cette extrémité est plus arrondie que pour les espèces de la sous-section *Rostratae* et les *V. sp* Hearn dont la pointe est bien distincte. Leur taille est en moyenne comprise entre 3 et 4,5 cm mais il a été remarqué des feuilles de plus petite taille pour les espèces issues de la sous-section *Rostratae* mesurant en moyenne 2 cm ; ou à l'inverse de plus grande taille pour les *V. sp* Hearn avec une moyenne de 5-6 cm. La couleur verte semble plus intense pour *V. alba* et il a été observé une base foncée pour *V. labradorica* (Figure 40).

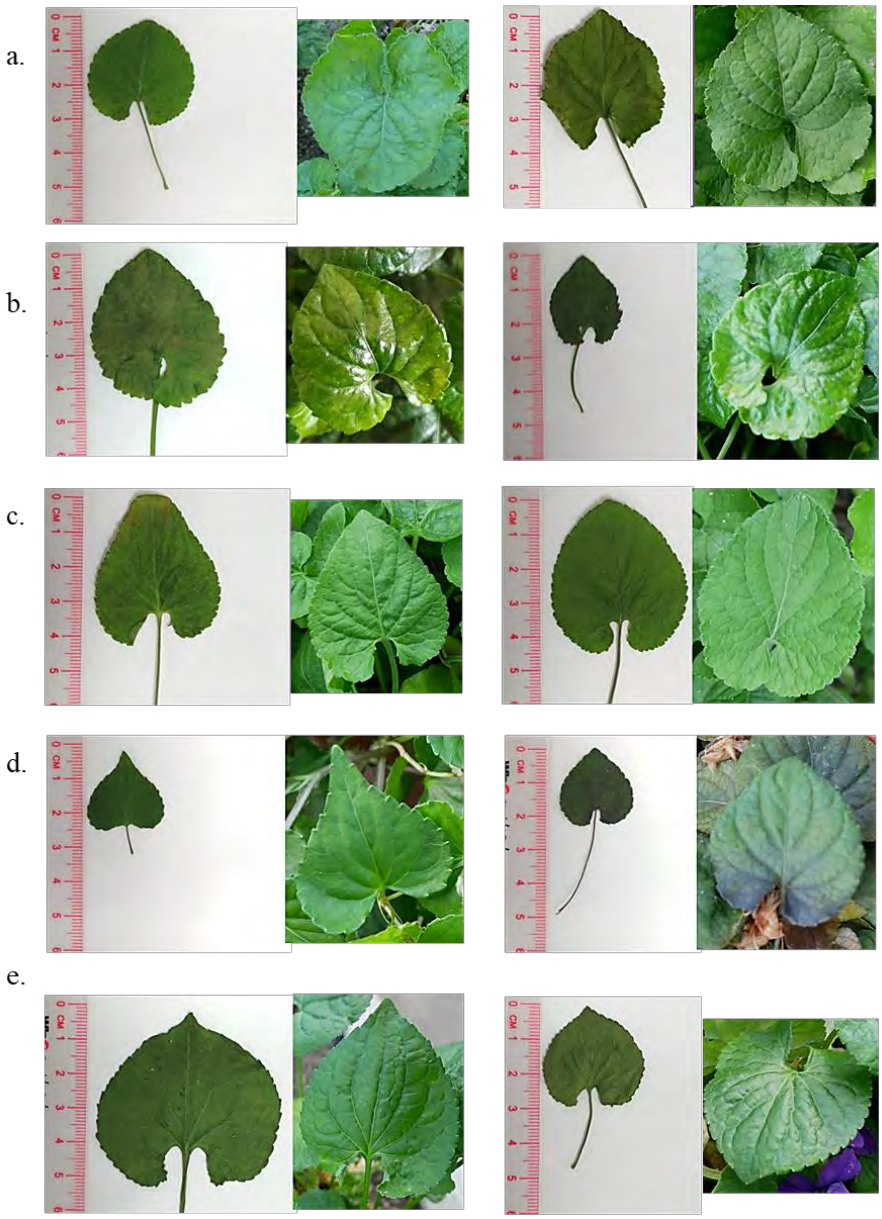


Figure 40. Feuilles séchées (1ère photo) et feuilles fraiches (2nd photo) pour deux violettes *V. odorata* (a), *V. alba* (b), *V. suavis* (c), *V. subs Rostrata*e (d.1 : *V. grypoceras* et d.2 : *V. labradorica*) (d) et *V. sp* Hearn (e)

Néanmoins deux autres formes ont été identifiées dans la collection. La première est une forme sagittée, feuille découpée d'environ 4,5 cm de long, dont certaines nervures secondaires partent en éventail depuis la base (Figure 41 a). Elle a été observée sur une violette non identifiée en génétique, mais d'origine asiatique. La deuxième forme est plutôt lancéolée avec une base cordée et correspond à une violette identifiée *V. mandshurica* décrite dans le Flore de Chine. La taille de la feuille est de l'ordre de 3 cm (Figure 41 b).

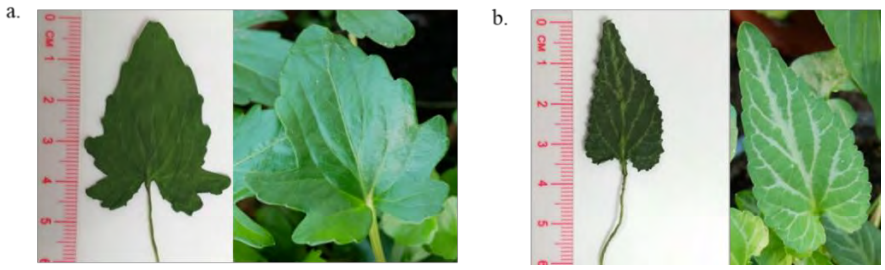


Figure 41. Autres formes de feuilles observées dans la collection de violettes

II.2.4.3 Analyse anatomique

II.2.4.3.1 Les bractéoles

Les fleurs portent deux bractéoles sur le pédoncule. Une analyse microscopique a donc été réalisée en premier lieu sur ces dernières. Une description fine de la première violette analysée, la référente violette de Toulouse (*Viola alba*), a été faite et les autres descriptions n'ont été que comparatives pour mettre en évidence le partage de mêmes critères et d'éventuelles divergences.

II.2.4.3.1.1 Viola alba

L'analyse microscopique a révélé la présence de poils sécréteurs sur le tiers inférieur de la bractéole (Figure 42 a). Ils sont localisés sur les extrémités des dents ou sur la marge. Deux parties le constituent à savoir un pied pluricellulaire et plurisérié, ainsi qu'une tête sécrétrice globuleuse pluricellulaire de diamètre moyen 75 µm (Figure 42 b). On suspecte que ces dernières sont constituées de poches d'huile essentielle. Une hydathode, orifice par lequel s'élimine l'eau liquide, semble former la partie effilée de la bractéole (Figure 42 c) et des poils tecteurs unicellulaires ont également étaient identifiés (Figure 42 d). Ils sont de l'ordre de 90 µm de long, présentent une paroi lisse épaisse et leur extrémité est pointue (Figure 42 e).

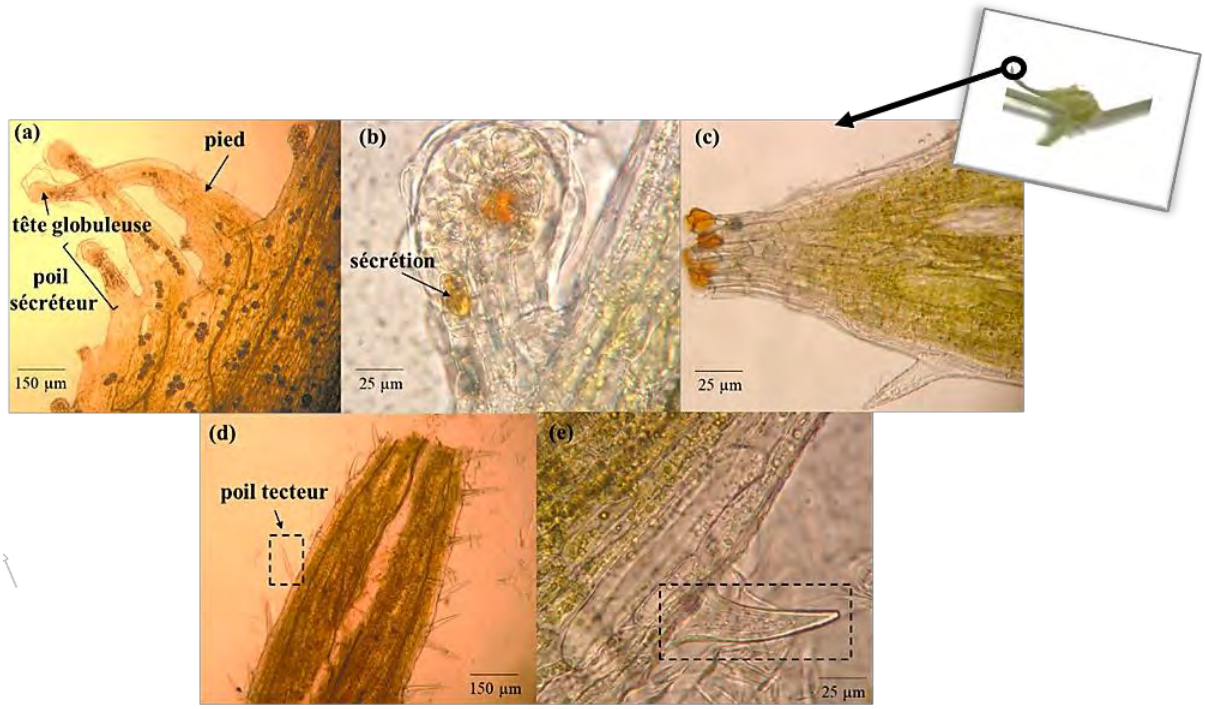


Figure 42. Poils sécréteurs (a, b), hydathode (c) et poils tecteurs (d, e) de *V. alba* vus au microscope

Cinq nervures ont été identifiées, une centrale et deux secondaires de chaque côté (Figure 43 a). La présence de mucilage a été détectée tout d'abord au niveau des bords : l'éclatement de cellules probablement à mucilage se traduit par l'observation d'un débordement, comme observé au niveau de la tête d'un poil sécréteur (Figure 43 b). Des cellules à mucilage sont également présentes au niveau de l'épiderme et du parenchyme chlorophyllien (Figure 43 c). Elles sont particulièrement longues et mesurent environ 30 µm de largeur et 0,2 mm de longueur.

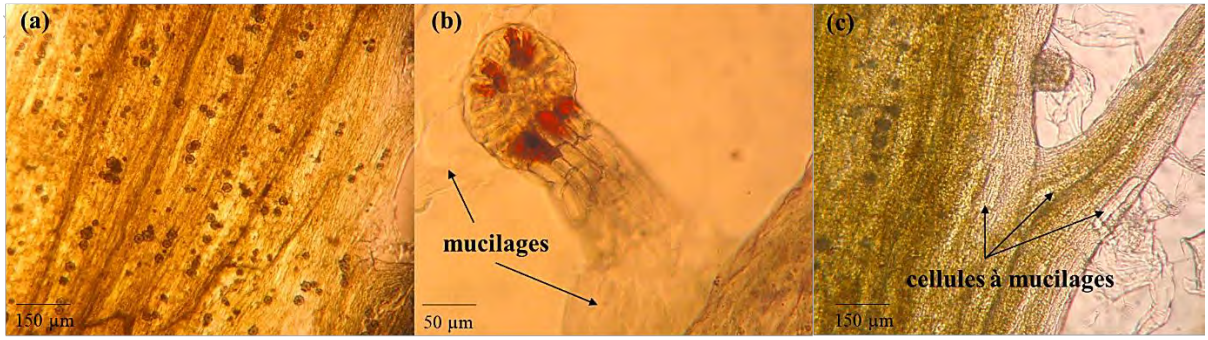


Figure 43. Observations microscopiques de nervures (a) et détection de la présence de mucilage au bord (b) et de cellules à mucilages dans le parenchyme (c) chez *V. alba*

Des prismes et des macles d'oxalate de calcium ont été observés (Figure 44 a). Les premiers sont localisés au niveau de l'épiderme, tandis que les seconds sont plutôt situés au niveau du parenchyme chlorophyllien. Ils ont été confirmés par lumière polarisée (Figure 44 b). Le parenchyme étant situé en dessous de l'épiderme, il en ressort que les macles apparaissent floues en lumière polarisée, la focalisation ayant été faite au niveau de l'épiderme. Tous deux sont de l'ordre de 25 µm.

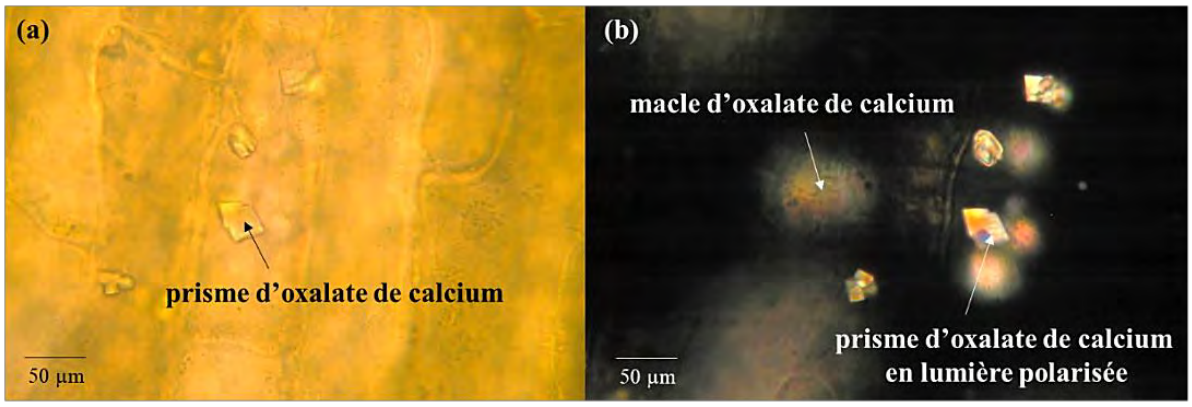


Figure 44. Observation microscopique de prismes d'oxalate de calcium au niveau de l'épiderme en lumière artificielle (a) et polarisée avec discernement de la présence de macles (b) chez *V. alba*

Enfin, quelques stomates ont été observés (Figure 45). Ils sont constitués par deux cellules stomatiques délimitant l'ostiole. Elles sont accompagnées de cellules annexes de type anisocytique, c'est-à-dire composées de trois cellules dont une plus petite. Il est à noter qu'une fois le montage chauffé, les trois cellules ne sont plus discernables. La littérature relate bien ce type de conformation pour le genre *Viola* (**Beille, 1909**). Ces derniers sont observés sur les deux faces.

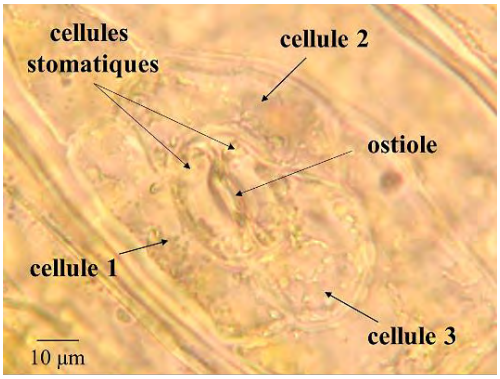


Figure 45. Stomate anisocytique de *V. alba* vu au microscope optique

II.2.4.3.1.2 Viola suavis

L'analyse microscopique de *V. suavis* a révélé la présence de poils sécréteurs, d'une hydathode et de poils tecteurs. Néanmoins, des différences sont constatées notamment au niveau des pieds des poils sécréteurs qui sont très courts et de l'extrémité plus arrondie (dit en doigt de gant) des poils tecteurs de l'ordre de 75 µm. Seulement trois nervures ont été identifiées ainsi qu'une cuticule striée. De nombreux

stomates de type anisocytique à trois cellules ont été observés. La présence de mucilage est avérée au niveau des bords et quelques cellules à mucilage, d'environ 0,3 mm, ont été observées, mais en nombre plus restreint. Uniquement des macles d'oxalate de calcium entre 25 et 35 µm ont été observées au niveau du parenchyme chlorophyllien.

II.2.4.3.1.3 Viola odorata

L'analyse microscopique de *V. odorata* a révélé la présence de poils sécréteurs à pied court avec indication de la présence de mucilage à la marge ainsi que d'une hydathode à l'extrémité supérieure et de poils tecteurs entre 200 et 400 µm en moyenne et à l'extrémité arrondie. Seules des macles de l'ordre de 25 µm ont été observées au niveau du parenchyme chlorophyllien, ainsi que des stomates de l'ordre de 25 µm de type anisocytique et une cuticule striée.

II.2.4.3.1.4 *Viola sp* Hearn

L'analyse microscopique de *V. sp* Hearn a révélé la présence de poils sécréteurs à pied court avec d'importantes sécrétions et la présence de mucilage à la marge et d'une hydathode à l'extrémité supérieure. Seules des macles de l'ordre de 25 µm ont été observées au niveau du parenchyme chlorophyllien, ainsi que des stomates de type anisocytique de l'ordre de 30 µm et une cuticule striée.

II.2.4.3.1.5 Viola labradorica

L'analyse microscopique de *V. labradorica,* de la sous-section *Rostratae*, a révélé la présence d'une hydathode à l'extrémité de la bractéole avec indication de la présence de mucilage et de sécrétions, ainsi que de cellules à mucilage de l'ordre de 0,1 mm de longueur. Seules des macles de l'ordre de 30 µm ont été observées au niveau du parenchyme chlorophyllien, ainsi que des stomates de l'ordre de 25 µm de type anisocytique et une cuticule striée. Il est à noter que pour cette espèce, une coloration rose est obtenue dans l'hydrate de chloral.

II.2.4.3.1.6 Viola grypoceras

L'analyse microscopique de *V. grypoceras* de la sous-section *Rostratae* a révélé la présence de poils sécréteurs à pied long avec indication de la présence de mucilage et d'une hydathode à l'extrémité. Seules des macles en moyenne de l'ordre de 30 µm ont été observées au niveau du parenchyme chlorophyllien, ainsi que des stomates de l'ordre de 25 µm de type anisocytique et une cuticule striée.

II.2.4.3.2 L'épiderme des feuilles

Les feuilles, en particulier l'épiderme, ont été analysées pour toutes ces espèces afin de mettre en évidence d'éventuels autres critères de différenciation que ceux obtenus avec les bractéoles.

II.2.4.3.2.1 Viola alba

L'analyse microscopique de l'épiderme de deux violettes identifiées chez *V. alba* a permis de définir les critères d'identification suivants : des stomates de type anisocytique de 23 µm dont les cellules présentent une forme lobée, de fins poils tecteurs à l'extrémité pointue au niveau des nervures, ainsi que de la marge de l'ordre de 0,16 mm et des macles confirmées en lumière polarisée et en moyenne de 20 µm. Il a été observé de nombreux espaces blancs au niveau de l'épiderme. Une coupe transversale (Figure 46 a) a permis de définir que ces espaces correspondaient à des cellules hypertrophiées de l'épiderme laissant passer plus facilement la lumière (Figure 46 b).

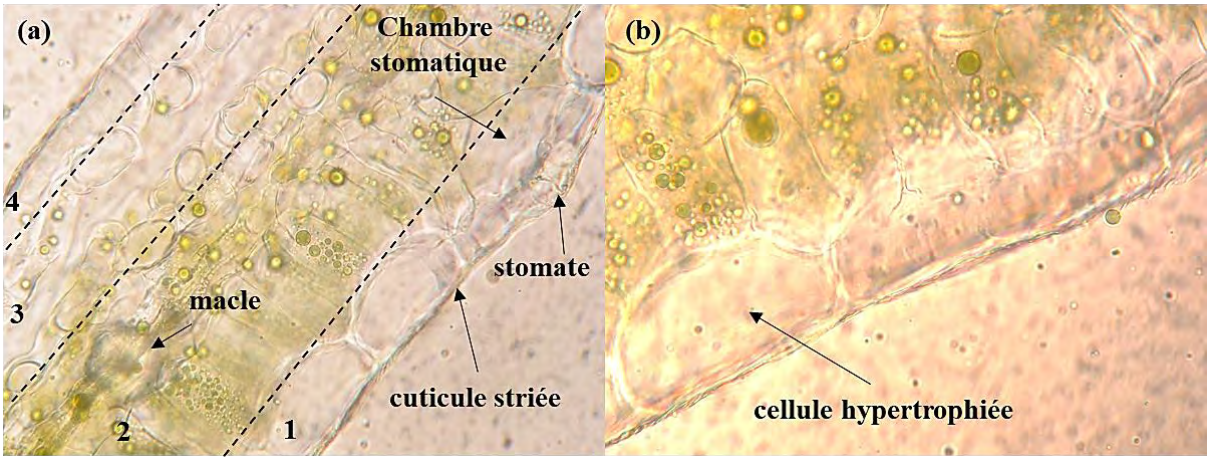


Figure 46. Coupe transversale mettant en évidence les différentes strates (a) : la cuticule striée, l'épiderme supérieur [1], le parenchyme palissadique [2], le parenchyme lacuneux [3] et l'épiderme inférieur [4]. Cellule hypertrophiée de l'épiderme non striée (b)

II.2.4.3.2.2 Viola suavis

Pour *V. suavis* les critères d'identification déterminés sont : une cuticule striée et des stomates de type anisocytique de 25 µm dont les cellules présentent une forme arrondie, quelques cellules hypertrophiées de l'épiderme de l'ordre de 90 µm de long, de fins poils tecteurs à l'extrémité pointue de l'ordre de 0,4 mm au niveau de la marge et des macles d'environ 20 µm.

II.2.4.3.2.3 Viola odorata

Pour *V. odorata* les critères d'identification déterminés sont : des stomates de type anisocytique de 25 µm dont les cellules présentent une forme arrondie, de fins poils tecteurs à l'extrémité plus arrondie de l'ordre de 0,2 mm au niveau de la marge et de 0,1 mm au niveau des nervures, ainsi que des macles d'environ 20 µm. La cuticule est striée et très peu de cellules hypertrophiées de l'épiderme ont été observées.

II.2.4.3.2.4 *Viola sp* Hearn

Pour *V. sp* Hearn les critères d'identification déterminés sont : des stomates de type anisocytique de l'ordre de 33 µm dont les cellules présentent une forme lobée, de très rares larges poils tecteurs à l'extrémité plutôt arrondie de l'ordre de 0,2 mm au niveau des nervures, ainsi que des macles d'environ 25 à 30 µm. La cuticule est striée et de nombreuses cellules hypertrophiées de l'épiderme ont été observées.

II.2.4.3.2.5 Viola labradorica

Pour *V. labradorica*, de la sous-section *Rostratae*, les critères d'identification déterminés sont : des stomates de type anisocytique de l'ordre de 27 µm dont les cellules présentent une forme arrondie à lobée, de fins poils tecteurs à cuticule transversalement striée à l'extrémité plutôt arrondie de l'ordre de 0,2 mm au niveau des nervures, ainsi que des macles d'environ 20 µm. La cuticule est striée et très peu de cellules épidermiques hypertrophiées ont été observées.

II.2.4.3.2.6 Viola grypoceras

Pour *V. grypoceras*, de la sous-section *Rostratae*, les critères d'identification déterminés sont : des stomates de type anisocytique de l'ordre de 25 µm dont les cellules présentent une forme plutôt arrondie, ainsi que des macles d'environ 20 µm. La cuticule est striée et très peu de cellules hypertrophiées de l'épiderme ont été observées. Par ailleurs, aucun poil tecteur n'a été détecté.

L'étude des feuilles est intéressante dans l'objectif de différenciation des violettes puisque c'est une ressource disponible toute l'année. Elle a été suffisante puisque quelques divergences ont été observées. La localisation des poils tecteurs observés sur la marge et/ou les nervures et la forme de leur extrémité, pointue ou plus arrondie, varient. La taille des stomates de type anisocytique diffère également avec l'observation de cellules annexes tantôt arrondies, tantôt de la forme des cellules épidermiques. Des macles plus imposantes ont été décrites chez *V. sp* Hearn tandis que seule *V. alba* ne semble pas présenter de cuticule striée. La présence de cellules épidermiques hypertrophiées est plus ou moins importante selon les espèces. L'étude de la bractéole située au niveau du pédoncule floral présente le principal avantage de pouvoir étudier l'organe dans son intégralité. Les mêmes constituants que dans l'épiderme ont été observés avec la présence additionnelle de poils sécréteurs à pied long ou court. Chez certaines espèces, des cellules à mucilages ont été détectées au niveau du parenchyme chlorophyllien et des prismes d'oxalate de calcium ont été uniquement identifiés chez *V. alba*. En complétant ces informations anatomiques avec celles morphologiques au niveau de la forme, de la couleur et de la taille des fleurs et des feuilles, des distinctions sont clairement décelables entre les différentes violettes. Le Tableau 19 résume les ressemblances et divergences observées et l'ensemble des observations microscopiques est présenté en annexe 5, page 224.

Tableau 19. Tableau des ressemblances-divergences des caractères morphologiques et anatomiques des six espèces de violettes majoritaires

	Caractère / Espèces	*V. alba*	*V. suavis*	*V. odorata*	*V. subs Rostratae*		*V. sp* Hearn
					V. labradorica	*V. grypoceras*	
Bractéole	Taille (cm)	0,6-0,8	0,5	0,5-0,6	0,4-0,5	0,5	0,6-0,7
	Distance de la fleur (cm)	6-7	2-3	4-5	1-2	1-2	7-7,5
	Cuticule striée	×	✓	✓	✓	✓	✓
	Prisme d'oxalate de Ca (taille)	✓ (25 µm)	×	×	×	×	×
	Macle (taille)	✓(25 µm)	✓(30 µm)	✓(25 µm)	✓(30 µm)	✓(25 µm)	✓(30 µm)
	Stomates de type anisocytique	✓ (Peu nombreux)	✓ (Nombreux)	✓ (Peu nombreux)	✓ (Peu nombreux)	✓ (Peu nombreux)	✓ (Peu nombreux)
	Cellule à mucilages	✓ (Nombreuses)	✓ (Peu nombreuses)	×	✓ (Nombreuses)	×	×
	Poils sécréteurs (taille)	✓ (Longs)	✓ (Courts)	✓ (Courts)	×	✓ (Longs)	✓ (Courts)
	Poils tecteurs (caractéristiques)	✓ (Courts, pointus)	✓ (Courts, arrondis)	✓ (Longs, arrondis)	×	×	×
	Hydathode	✓	✓	✓	✓	✓	✓
Epiderme	Cuticule striée	×	✓	✓	✓	✓	✓
	Stomates de type anisocytique (taille, forme des cellules)	✓ (23 µm, épidermiques)	✓ (25 µm, arrondies)	✓ (25 µm, arrondies)	✓ (27 µm, arrondies à épidermique)	✓ (25 µm, arrondies)	✓ (33 µm, épidermiques)
	Macles (taille)	✓ (20 µm)	✓ (20 µm)	✓ (20 µm)	✓ (20 µm)	✓ (20 µm)	✓ (25 - 30 µm)
	Poils tecteurs (Localisation, forme, taille)	✓ Marge et nervures Fins et pointus ~ 0,16 mm	✓ Marge Fins et arrondis ~ 0,4 mm	✓ Marge et nervures Fins et arrondis ~ 0,2 mm	✓ Nervures ~ 0,2 mm	×	✓ (très rares) Nervures Larges et arrondis ~ 0,2 mm
	Cellules hypertrophiées de l'épiderme	✓ Nombreuses	✓ Nombreuses	✓ Peu nombreuses	✓ Peu nombreuses	✓ Peu nombreuses	✓ Nombreuses
Fleur	Pétale inférieur plus échancré	✓	✓	✓	×	×	×
	Couleur	Blanc, bleu violacé ou violet avec centre blanc	Bleu violet	Violet, lilas, blanc	Violet lavande, violet	Violet à blanc	Violet à cœur blanc
	Taille (cm)	1-2	1-2	1-2	1	1-1,5	2-3
Feuille	Forme	Cordée, extrémité arrondie	Cordée, extrémité arrondie	Cordée, extrémité arrondie	Cordée, extrémité pointue	Cordée, extrémité pointue	Cordée, extrémité pointue
	Taille (cm)	3-4,5	4-4,5	3-4	2-2,5	2-2,5	3-5

II.2.5 Conclusion

Cette étude combinant les aspects génétique et phytochimique des composés volatils et non volatils, a permis de mettre en évidence la limite de l'analyse génétique dans notre tentative de classification des violettes de la collection nationale. Une première analyse génétique, non reportée ici et basée sur les analyses séparées des ITS1 et ITS2 avait permis d'obtenir une première base de classification. Cependant, la présence d'identifications non équivalentes entre les deux espaces pour un même plant a conduit à revoir la méthode de séquençage en intégrant l'espace central 5,8S pouvant apporter des informations complémentaires pour l'identification. Ainsi, environ 60% de la collection ont pu être identifiés par un nom d'espèce. Par ailleurs, l'utilisation d'amorces universelles pour l'amplification de l'ADN a conduit à la mise en place de fortes contraintes de stérilité, tant au niveau de l'environnement de travail que du matériel, en raison des hauts risques de contamination par tout ADN extérieur. Les protocoles de préparation ont donc été alourdis pour répondre à ces contraintes. C'est ainsi que la combinaison avec des analyses chimiotaxonomiques a été envisagée pour tenter d'étendre la classification sur la base des profils chimiques, de manière moins contraignante. Les résultats de l'analyse des composés volatils n'ont pas permis d'identifier une part très importante des violettes inconnues, notamment pour des raisons de sensibilité de la technique et du nombre restreint de variables discriminantes. La période de floraison étant limitée à environ deux mois sur l'année et les protocoles d'analyse standardisés étant relativement longs entre le processus d'émission des composés et l'analyse, seulement environ 30% de la collection ont pu être intégralement caractérisés. En revanche l'analyse des métabolites secondaires non volatils des parties aériennes a été plus discriminante puisqu'à terme, seulement 4% des violettes de la collection sont restés non indexés (Figure 47). En effet, les protocoles de mise en œuvre de cette technique sont moins contraignants et utilisent une matière végétale plus disponible assurant l'analyse de la quasi-totalité de la collection. Le seul facteur limitant reste la présence suffisante de feuilles sur le plant, pour assurer une récolte raisonnable. Sept espèces et l'ensemble *V. sp* Hearn ont donc été discernés et présentant, pour la plupart, des profils volatils et non volatils bien distincts avec l'identification de biomarqueurs. Une carte d'identité phytochimique qualitative de chaque espèce a donc été obtenue. Finalement, la combinaison des trois analyses a permis une totale discrimination des espèces, la génétique étant le point de départ à l'identification des plants, avec extension par l'intermédiaire de la comparaison des profils chimiques. L'étude botanique portant sur les bractéoles du pédoncule floral et un échantillon d'épiderme de la feuille a permis de confirmer la présence d'espèces différentes par l'observation de divergences morphologiques et anatomiques, et de valider les identifications obtenues par l'analyse génétique par comparaison de certains aspects avec la littérature.

La limite principale de la méthode employée émane de la non-identification d'éventuelles variétés ou sous-espèces, suspectées au sein des plants. Le principal objectif dans ce travail était de combiner différentes approches pour améliorer l'identification des plants et la phylogénie n'était qu'un premier outil. C'est pourquoi il a été décidé de mettre en œuvre le séquençage direct des ITS. La dénomination des espèces était alors suffisante, mais des analyses complémentaires seraient à envisager dans le cadre de l'approfondissement des génotypes de la collection en séquençant par exemple quelques régions de l'ADN chloroplastique (cpDNA) comme effectué chez des violettes au Japon (**Nakamura et al., 2015**). Par ailleurs, la présence d'hybrides détectés par le polymorphisme de certains pourrait être étudiée plus en détail en clonant les produits PCR comme dans le cas de la violette *Viola woosanensis* de Corée (**Gil et Kim, 2016**). En combinant les résultats des ITS et cpDNA, il serait alors possible de définir la lignée maternelle et celle paternelle à l'origine de l'hybride.

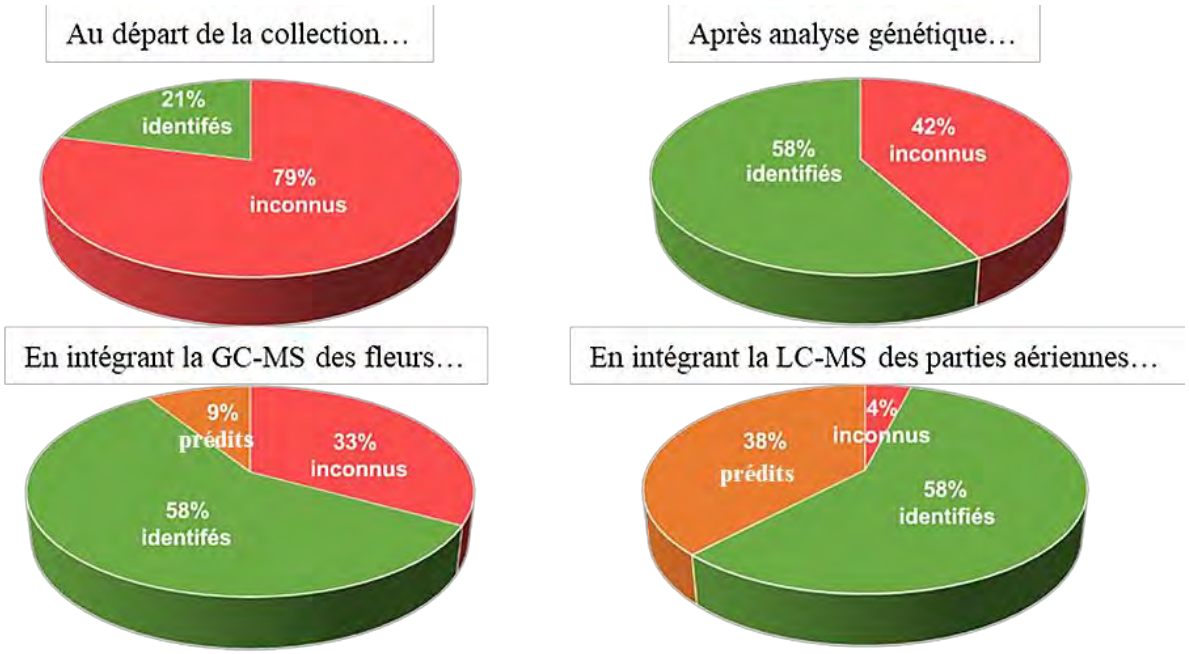


Figure 47. Identification et classification des violettes selon les méthodes employées

À présent qu'une vision plus claire de la constitution de la collection a été obtenue par une combinaison des données génétiques, chimiotaxonomiques et botaniques, une étude plus approfondie du potentiel biologique des espèces a été entreprise. Une première analyse détaillée de la violette de Toulouse, disponible en plus grande quantité aux serres, a été réalisée afin de mettre au point les protocoles de traitement, puis l'ensemble de la collection a été criblée afin de mettre en évidence les composés d'intérêt ainsi que d'éventuelles relations espèces-activités.

II.3 ETUDE DETAILLEE DE LA VIOLETTE DE TOULOUSE

II.3.1 CONTEXTE

L'ensemble des méthodes d'analyse a été dans un premier temps mis au point sur la violette de Toulouse. Rappelons que l'un des principaux objectifs de cette thèse est la caractérisation de l'activité antioxydante des parties aériennes de la violette[7]. En effet, une précédente étude au laboratoire a démontré de très bonnes activités antioxydantes des extraits éthanoliques des fleurs ainsi que des parties aériennes (feuilles et stolons) de violettes, contrairement aux racines qui n'ont présenté aucune action particulière. Le terme « antioxydant » a été formulé comme « une substance qui en faibles concentrations, en présence d'un substrat oxydable, ralentit ou empêche significativement l'oxydation des substrats matériels » (**Halliwell, 1995**). En cosmétique, les extraits traditionnels restent encore aujourd'hui des extraits hydroalcooliques, et plus particulièrement des extraits éthanoliques conformes aux restrictions cosmétiques. Cette extraction a donc été la première employée et étudiée en détail afin de mettre en évidence les composés impliqués dans l'activité antioxydante de la violette de Toulouse.

II.3.2 FRACTIONNEMENT SUR CARTOUCHE SPE C_{18}

L'extrait brut de violette de Toulouse, obtenu par une extraction éthanolique à 80% assisté par ultrasons, a été fractionné sur cartouche de silice C_{18} pour le décompléxifier et obtenir des fractions

[7] Cette étude a fait l'objet d'une publication en juillet 2017 dans le journal *Metabolomics* (Chervin et al., 2017).

enrichies en composés polaires et apolaires (cf. IV.5.2.2.6, page 182). Cinq fractions de polarités décroissantes ont été obtenues et dont les chromatogrammes UHPLC sont présentés Figure 48. Les fractions Fr1 et Fr2 sont enrichies en composés polaires aux faibles temps de rétention, tandis que les fractions Fr4 et Fr5 le sont en composés apolaires présentant des temps de rétention plus élevés en phase C_{18}. Les proportions moyennes de chaque fraction recueillie sont présentées Figure 49. La masse d'extrait brut introduite au départ, de l'ordre de 800 mg, n'est pas entièrement retrouvée en additionnant l'ensemble des poids des fractions. Environ 25% de l'extrait n'est pas retrouvé et est probablement resté au niveau de la cartouche SPE. Néanmoins, ce résidu apparait non-intéressant pour cette étude, puisque plutôt concentré en composés lipophiles.
On constate que la fraction Fr1 correspondant au solvant le plus polaire est la plus importante. Ainsi, les extraits de violettes semblent majoritairement constitués de composés très polaires représentant près de 60% de l'extrait brut, tandis que les autres fractions représentent toutes une proportion inférieure à 10%, voire même 5%.

Cette étape de fractionnement apparait primordiale dans l'objectif de détermination du potentiel antioxydant de la violette, puisqu'il devrait permettre de cibler la zone du chromatogramme, et donc le type de composés, qui semblent conférer aux extraits cette activité.

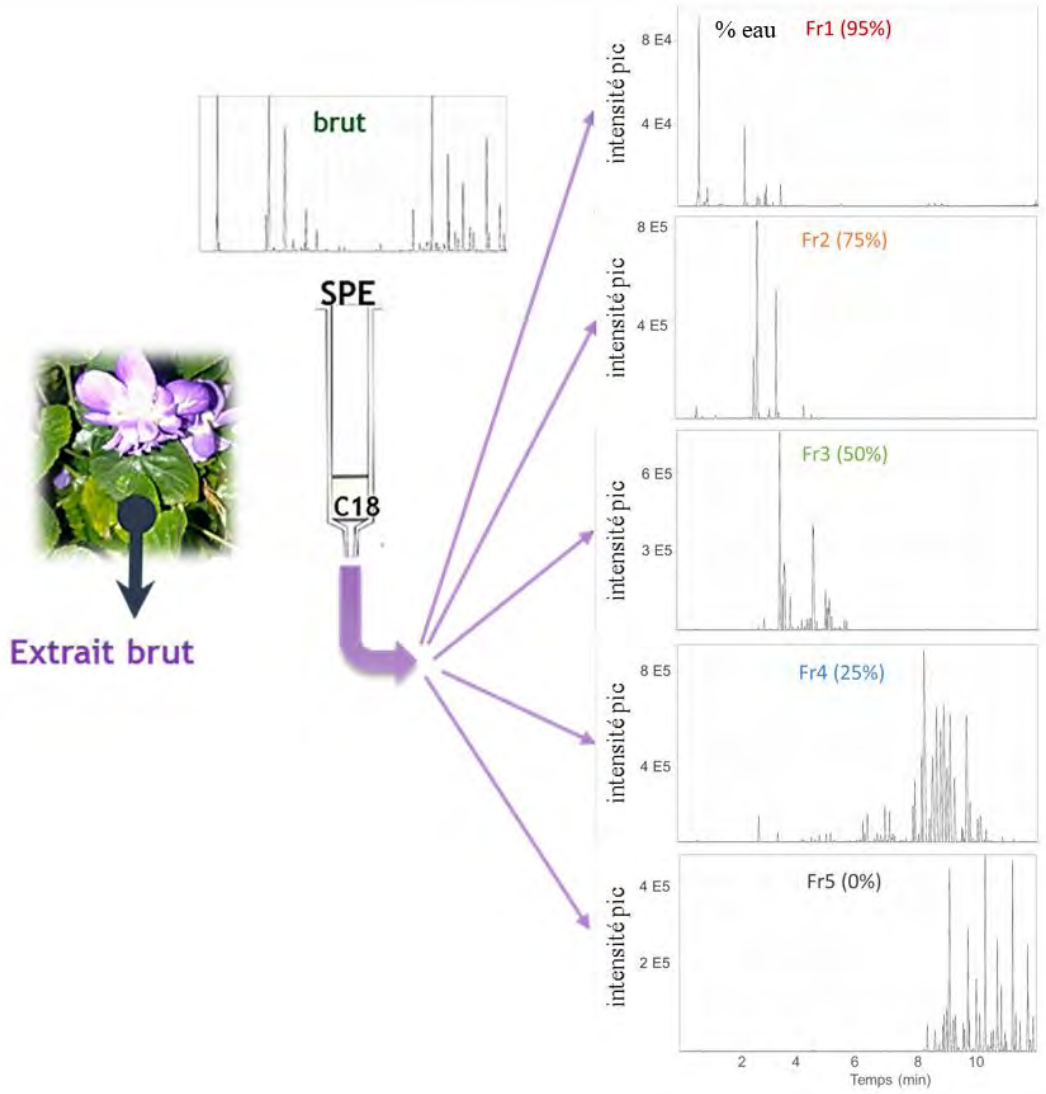


Figure 48. Fractionnement de l'extrait éthanolique de la violette de Toulouse et profils chromatographiques UHPLC des fractions SPE collectées

(a)

	Fr1	Fr2	Fr3	Fr4	Fr5
Masse(g)	$4{,}55\times10^{1}$	$3{,}60\times10^{2}$	$2{,}15\times10^{2}$	$6{,}19\times10^{3}$	$5{,}85\times10^{2}$
Ecart-type	$9{,}93\times10^{2}$	$7{,}70\times10^{3}$	$8{,}42\times10^{3}$	$1{,}90\times10^{3}$	$1{,}76\times10^{2}$
Rendement	59%	4%	3%	1%	8%

(b)

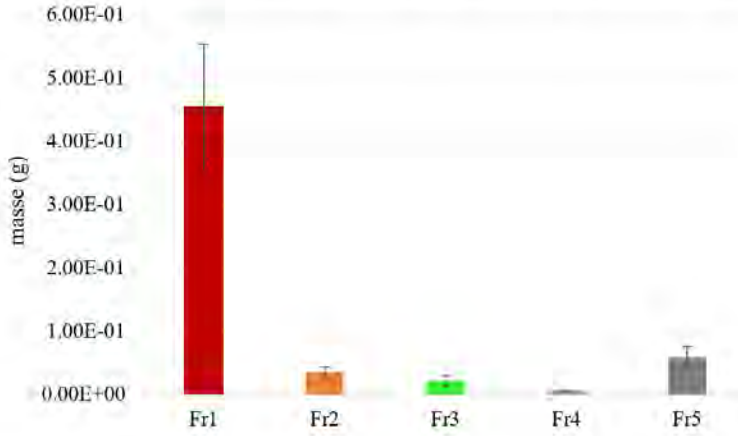


Figure 49. Bilan de masse moyen (a) et graphique associé (b) du fractionnement de l'extrait brut sur cartouche C_{18}

II.3.3 Analyse UHPLC-HRMS/MS

Chaque fraction collectée a été séchée, calibrée à 2 mg/mL dans l'éthanol 80% et injectée en UHPLC-HRMS (cf. IV.5.2.3.1.1, page 184). Les données ont ensuite été analysées à l'aide du logiciel MS-DIAL dans le but d'obtenir une liste de pics globale de l'ensemble des métabolites détectés dans les fractions. Ce sont finalement 434 pics caractérisés par la paire RT/mz qui ont été obtenus en mode négatif, contre 527 en positif (cf. IV.5.2.4.1, page 186).

Une analyse statistique de type ACP a permis de mettre en évidence la reproductibilité de la méthode et de la composition chimique des individus étudiés par le regroupement homogène de chaque fraction (Figure 50). Plusieurs méthodes de normalisation ont été testées pour finalement sélectionner celle avec les données logarithmées et mises en échelle pareto, puisqu'il traduisait le mieux l'ensemble des données. En effet, les extraits bruts et les QC, qui sont un pool de l'ensemble des fractions analysées, sont regroupés au centre de l'ellipse, et les fractions sont organisées autour, puisque les premiers comprennent l'ensemble des métabolites tandis que chaque fraction n'en comprend qu'une partie.

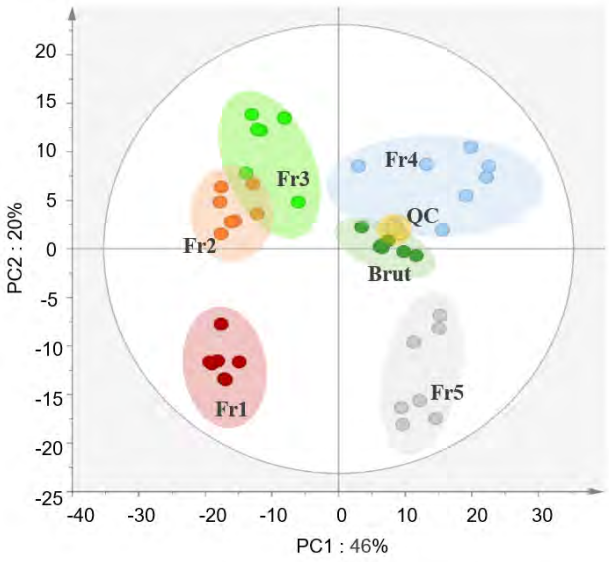


Figure 50. ACP des fractions obtenues par SPE, des extraits bruts et des contrôles qualité

Des analyses biologiques ont ensuite été entreprises sur l'ensemble de ces fractions afin de mettre en évidence celles à travailler plus en profondeur pour déterminer les molécules actives.

II.3.4 POTENTIEL ANTIOXYDANT DE L'EXTRAIT HYDROALCOOLIQUE

L'objectif a été de confirmer le fort potentiel antioxydant des parties aériennes de la violette de Toulouse par l'intermédiaire de deux méthodes, le test DPPH et la résonance paramagnétique électronique (RPE), et de mettre en évidence les composés responsables de cette activité. Pour cela, une approche métabolomique, faisant usage de la projection orthogonale des structures latentes (OPLS) avec interprétation des coefficients de cette régression, a été mise en œuvre pour révéler les composés d'intérêt en mettant en regard les profils chimiques des résultats biologiques. Deux classes chimiques ont ainsi été mises en évidence, avec confirmation par la spectrométrie de masse couplée avec un lecteur de plaques de chromatographie sur couche mince (TLC-MS) par l'intermédiaire d'une révélation DPPH, assurant l'analyse directe des composés élués sur la plaque.

II.3.4.1 Résultats des essais DPPH et RPE

Chaque fraction a été testée pour son activité antioxydante à l'aide du radical 2,2-diphényl-1-picrylhydrasil (DPPH) à une longueur d'onde de 517 nm. La gamme de concentration inhibitrice médiane (CI_{50}) était comprise entre 5 mg/L et 1000 mg/L. La fraction possédant la valeur de concentration inhibitrice 50% (CI_{50}) la plus faible est l'extrait possédant la réserve antioxydante la plus importante (cf. IV.6.1.1, page 191).

L'essai DPPH est une technique principalement basée sur le transfert d'électrons et est adapté aux molécules hydrophiles et lipophiles. Cependant, l'encombrement stérique peut conduire à une réaction plus lente voire une inertie de la molécule en raison de l'accessibilité plus restreinte au site radical, l'atome d'azote. C'est pourquoi un deuxième essai antioxydant a été mis en œuvre par RPE, afin d'étudier le comportement des extraits avec un test basé sur un mécanisme différent centré sur le radical superoxyde (cf. IV.6.1.2, 192). Pour cette étude, le potentiel antioxydant a été déterminé à partir de l'intégration double du signal observé : plus l'intégration est petite, plus l'extrait présente un fort potentiel antioxydant. Les valeurs moyennes de Cl_{50} obtenues par la méthode DPPH et les intégrations des signaux RPE de chaque fraction sont présentées Figure 51. La différence de capacité de piégeage radicalaire du trolox entre les essais DPPH et RPE peut être facilement expliquée en considérant les constantes de vitesse des réactions de piégeage: kTrolox/DPPH = $1{,}1 \times 10^4$ $M^{-1}s^{-1}$ (**Friaa et Brault 2006**) et kTrolox/$O_2^{\bullet -}$ = 0,1 $M^{-1}s^{-1}$ (**Cabelli et Bielski, 1986**). Ainsi, pour le Trolox a été la référence choisie pour le test DPPH et la rutine celle du test RPE.

(a)

	Brut	Fr1	Fr2	Fr3	Fr4	Fr5	Trolox	Rutine
CI_{50} (mg/L)	467	2605	53	153	429	1337	7	14
Écart type	14	380	5	9	10	26	0	0
Intégration RPE	12,7	12,8	2,5	4,7	11,3	14,5	15	1,8
Écart-type	0,7	0,9	0,3	0,4	0,8	2,5	1	0,1

(b)

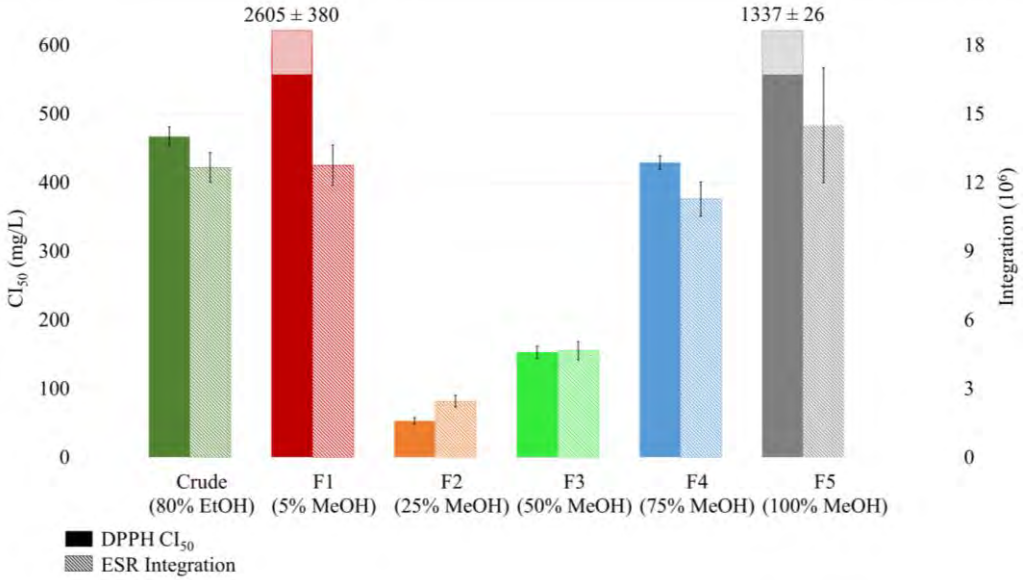


Figure 51. Valeur des CI_{50} et intégration des signaux RPE de chaque fraction (a) et graphique associé (b)

Les résultats permettent de constater que les deux essais (DPPH et RPE) conduisent à des résultats comparables. D'après les expériences, il a pu être constaté que les fractions Fr1 et Fr5, qui sont respectivement la fraction la plus polaire et celle la plus apolaire, ne présentent aucun potentiel antioxydant. On remarque d'ailleurs que leurs valeurs de CI_{50} sont en dehors du domaine d'étude, car supérieures à 1000 mg/L et de 100 à 400 fois supérieures aux valeurs des composés de référence. En revanche, les fractions Fr2 et Fr3 sont celles qui semblent présenter les réserves antioxydantes les plus intéressantes puisqu'elles possèdent les valeurs de CI_{50} et d'intégration des signaux RPE les plus faibles, de seulement 3 à 20 fois supérieures aux références. De plus, ces dernières sont inférieures à celle de l'extrait brut, mettant ainsi en évidence un effet de concentration. Ces molécules ne doivent très probablement pas être des constituants majoritaires, mais présents tout de même en concentration suffisante pour conférer une CI_{50} moyenne de l'ordre de 450 mg/mL à l'extrait brut. Il parait donc pertinent de continuer l'investigation en se focalisant sur ces deux fractions afin d'identifier les molécules jouant un rôle dans cette activité.

II.3.4.2 Analyses statistiques des résultats antioxydants

L'OPLS (Figure 52 a) a permis d'obtenir un classement de l'ensemble des métabolites (variable X) sur la base des résultats antioxydants par la méthode DPPH (variable Y). En effet, le modèle obtenu (R2Y = 0,977, Q2Y = 0,952) ayant été validé par le test de permutation (Figure 52 b), il peut être utilisé à des fins prédictives.

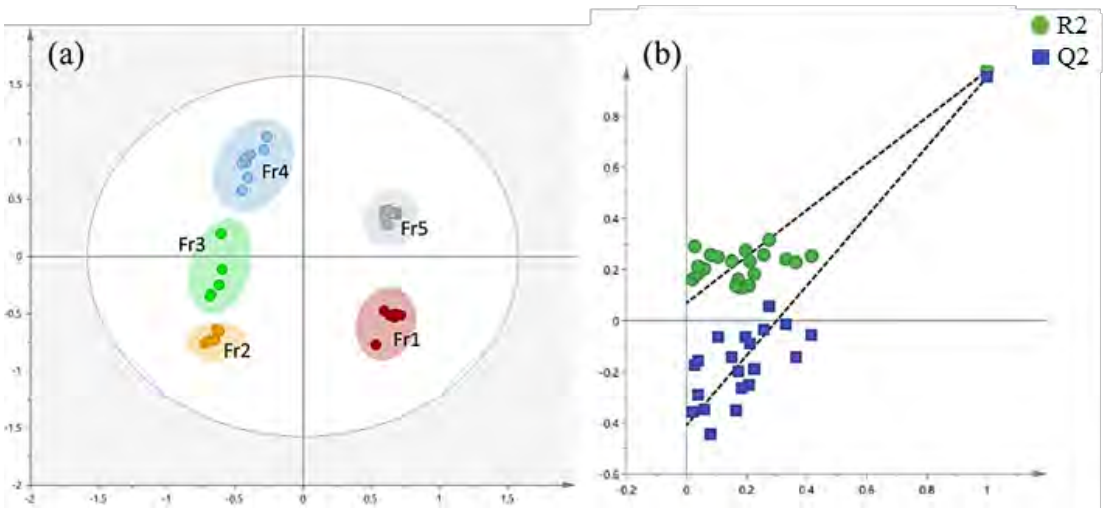
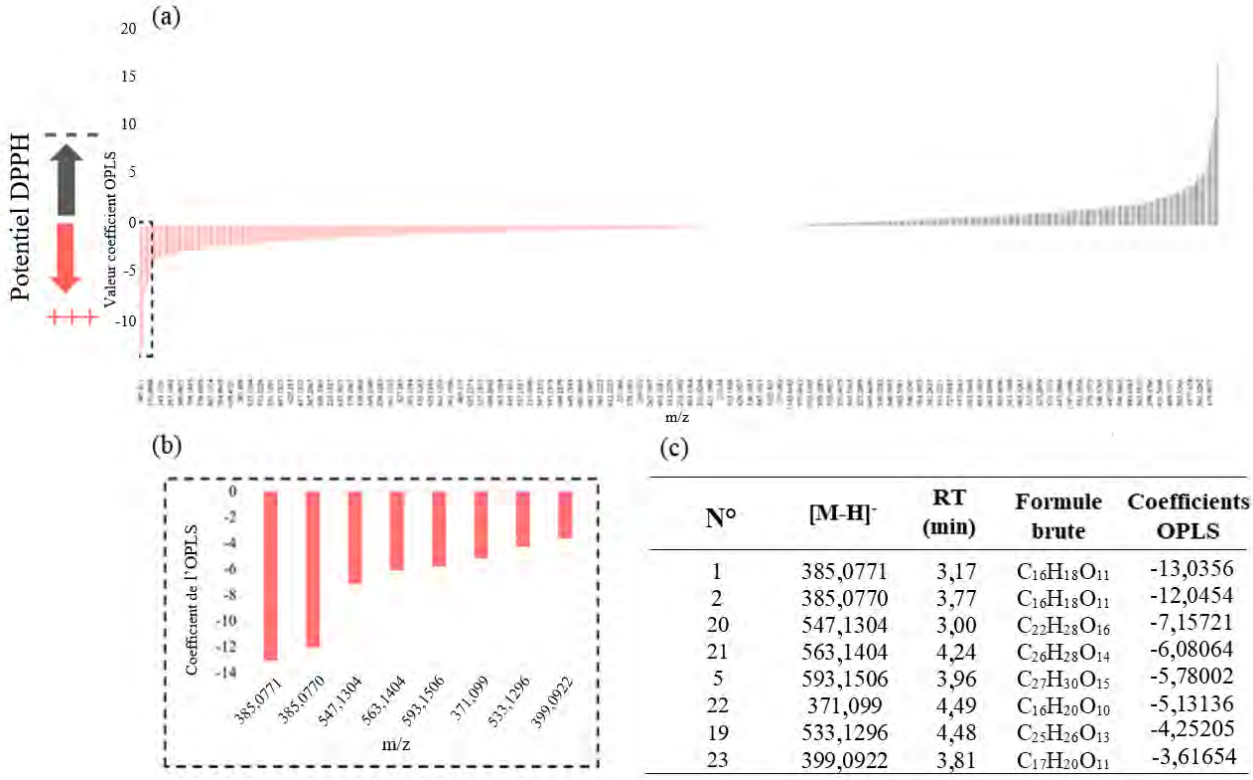


Figure 52. Modèle OPLS des données DPPH (a) et test de permutation associé (b)

Les coefficients de la régression OPLS connotent l'implication des variables dans l'activité antioxydante. Dans ce cas, les résultats de l'essai DPPH étant sous forme de CI_{50}, ce sont les coefficients négatifs qui sont associés aux molécules potentiellement antioxydantes. Les molécules qui seront essentiellement présentes au sein des fractions Fr2 et Fr3 occuperont les premiers rangs du classement, puisque ce sont les deux fractions révélées les plus riches en antioxydants. Huit composés d'intérêt ont été mis en relief, sur la base des analyses UHPLC-HRMS (Figure 53).



(c)

N°	[M-H]⁻	RT (min)	Formule brute	Coefficients OPLS
1	385,0771	3,17	$C_{16}H_{18}O_{11}$	-13,0356
2	385,0770	3,77	$C_{16}H_{18}O_{11}$	-12,0454
20	547,1304	3,00	$C_{22}H_{28}O_{16}$	-7,15721
21	563,1404	4,24	$C_{26}H_{28}O_{14}$	-6,08064
5	593,1506	3,96	$C_{27}H_{30}O_{15}$	-5,78002
22	371,099	4,49	$C_{16}H_{20}O_{10}$	-5,13136
19	533,1296	4,48	$C_{25}H_{26}O_{13}$	-4,25205
23	399,0922	3,81	$C_{17}H_{20}O_{11}$	-3,61654

Figure 53. Classement des métabolites sur la base des coefficients obtenus par l'OPLS (a) et zoom sur les huit premiers hits (b) avec détails de leurs caractéristiques (c)

II.3.4.3 Identification et déréplication des molécules d'intérêt

Huit molécules d'intérêt ont été identifiées par l'intermédiaire du logiciel MS-FINDER à partir de différentes bases de données de produits naturels (PlantCyc, UNPD, DNP, KNApSAcK et ChEBI) et de

bases personnalisées sur le genre *Viola*. Les résultats ont été présentés sous la forme d'une liste classée selon le score de similitude entre les fragments MS/MS observés et ceux *in silico*. Ce score prend en compte l'incertitude sur la différence de masse, le profil isotopique ainsi que la similarité des fragments MS/MS. Seules les structures présentant un score au-dessus de 7 ont été prises en considération. En effet, il est estimé qu'un score supérieur à 7 est lié à une hypothèse de structure très forte tandis qu'entre 5 et 7, cette hypothèse est plus discutable, mais la classe chimique est globalement pertinente. En deçà, les résultats d'identification ne sont pas pris en compte.

Un réseau moléculaire basé sur les similitudes spectrales des composés (cf. bases théoriques, annexe 1 C, page 217) a été construit à partir du logiciel en ligne MetamapR dans le but de présenter les résultats de manière plus fluide et pouvoir approfondir la déréplication des composés non identifiés par les bases de données. Étant basés sur la similitude de fragmentation des composés, seuls les composés possédant une fragmentation ont été conservés. Le réseau comprend alors environ 180 nœuds et en implémentant les classes chimiques des composés identifiés à l'aide de MS-FINDER, il a été observé la présence de composés aliphatiques et terpéniques en grande proportion, ainsi que de flavonoïdes et de coumarines auxquels les composés antioxydants semblent appartenir, d'après les résultats de l'OPLS (Figure 54).

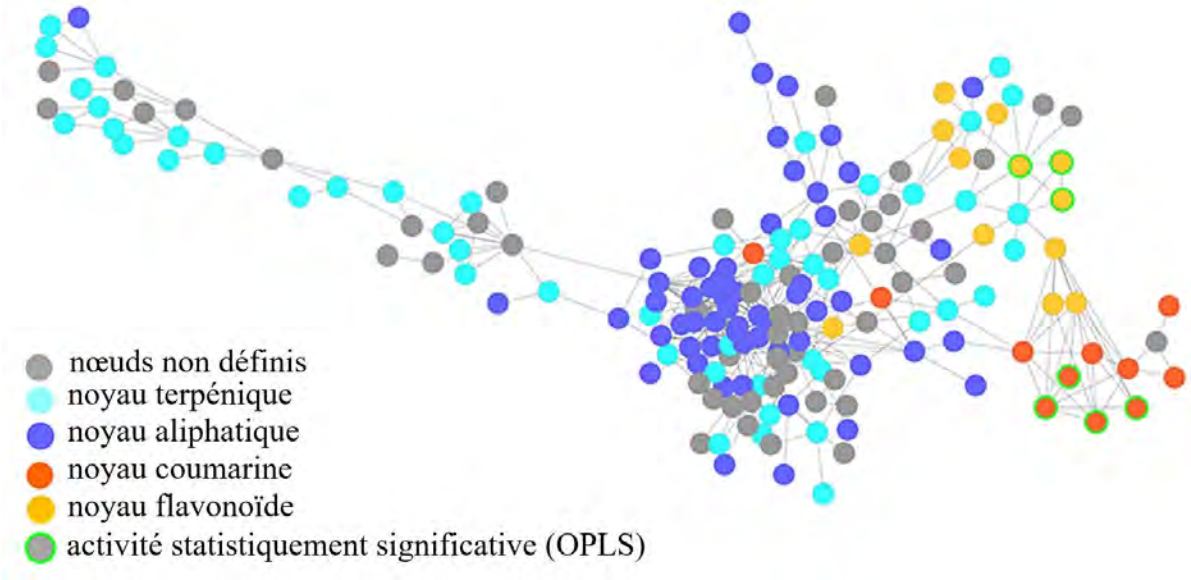


Figure 54. Réseau moléculaire des métabolites secondaires de la violette de Toulouse

À partir du réseau obtenu et des précédentes déterminations des composés antioxydants par le modèle OPLS, il a été observé deux groupes de composés antioxydants, chacun corrélé à une fraction SPE : les flavonoïdes sont principalement retrouvés dans la fraction Fr3, tandis que les coumarines sont essentiellement corrélées à la fraction Fr2.

Le groupe des flavonoïdes comprend trois composés actifs présentant des structures similaires (Figure 55). Ces trois composés possèdent des formules chimiques qui correspondent à des structures retrouvées dans les bases de données du genre *Viola*. La comparaison des spectres MS/MS expérimentaux et des fragmentations *in silico* a permis de les identifier: le schaftoside (21), le 6-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxy-8-*C*-α-L-rhamnopyranosylflavone (5) et le 6,8-Di-*C*-α-L-arabino pyranosylapigenin (19).

En ce qui concerne le groupe des coumarines, il est constitué de cinq composés statistiquement révélés antioxydants, dont une coumarine 8-*C*-glycolsylée (22) déterminée par comparaison avec les bases de données **;** et deux isomères (1 et 2) de formule brute $C_{16}H_{18}O_{11}$ clairement déterminés par analyses de

résonance magnétique nucléaire mono et bidimensionnelle (RMN 1D et 2D, cf. IV.5.2.3.3 page 185) et respectivement identifiés comme le 7-methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside et le 7-methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside. La différence entre les deux molécules réside dans la position du glucose qui influence le déplacement chimique du proton en position 8. La coumarine 23, retrouvée en mélange avec la coumarine 2, mais en quantité minoritaire, a également été partiellement authentifiée par RMN 1D et 2D.

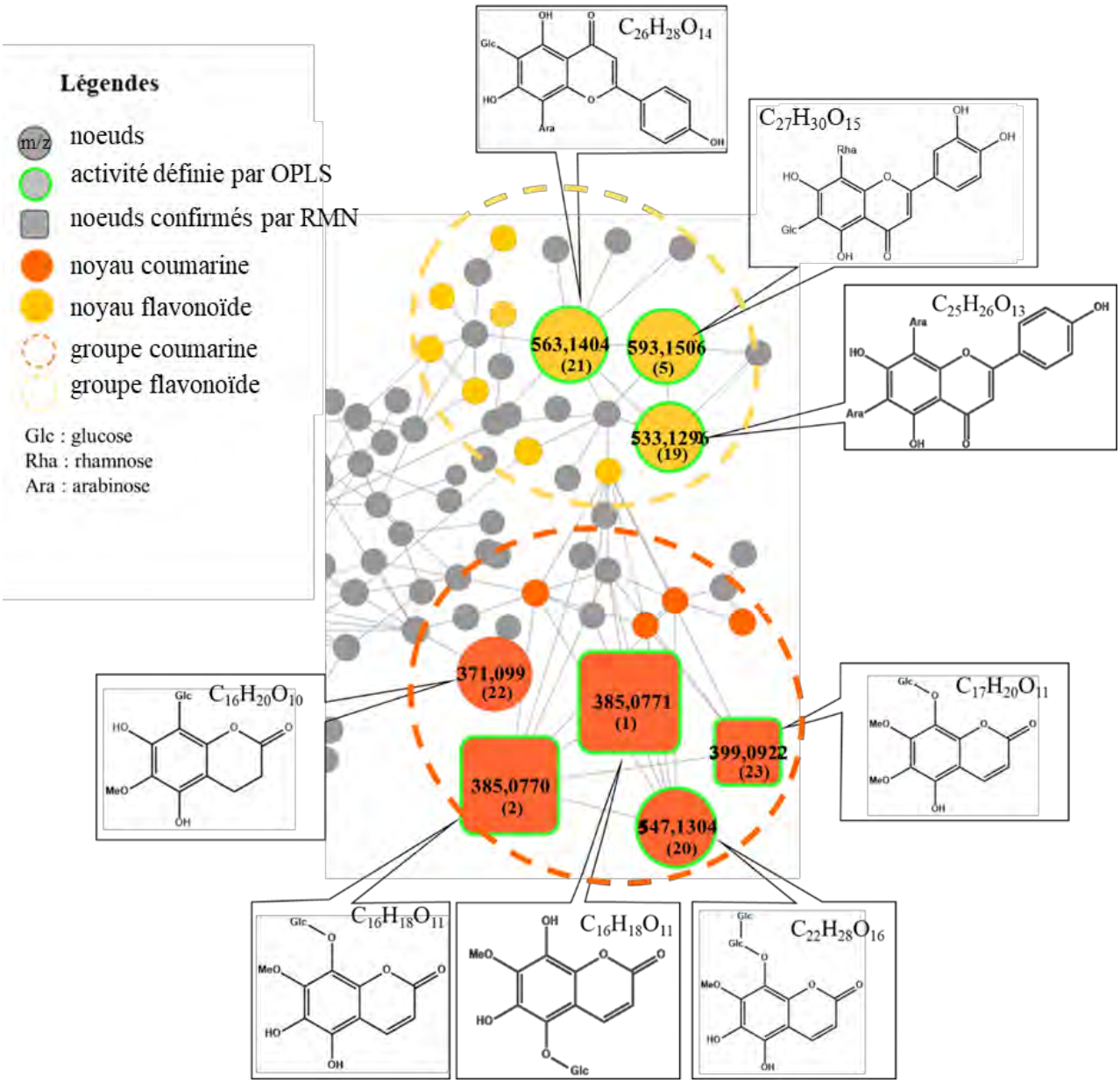


Figure 55. Emphases des groupes des flavonoïdes et des coumarines aux forts potentiels antioxydants

Les composés 1 et 2 ont été isolés par UHPLC-UV (IV.5.2.3.1.2 page 185). Les spectres RMN du proton des deux composés 1 et 2 contenaient des signaux communs caractéristiques, révélant un noyau coumarinique. Les paires de doublets de déplacements chimiques (δ) $\hat{o}_H$ vers 6 et 8 ppm avec une constante de couplage de 9,6 Hz sont caractéristiques des protons H3 et H4 d'une coumarine (**Wu et al., 1989**). Les spectres RMN ^{1}H (Figure 56) et ^{13}C (Figure 57) de 1 ont montré : un groupe méthoxy à $\hat{o}_H$ 3,88 (s, 3H) qui était en corrélation avec le signal RMN ^{13}C à δ_C 60,65 ; un glucose *O*-ß glycosylé avec un proton anomérique à $\hat{o}_H$ 4,72 (d, J = 7,56 Hz, 1H) et les autres déplacements caractéristiques à $\hat{o}_H$ 3,29 (dd, 7,59 et 8,97 Hz, 1H), $\hat{o}_H$ 3,23 (m, 1H), $\hat{o}_H$ 3,16 (dd, 7,97 et 10,09 Hz, 1H), δ_H 3,05 (ddd, 2,07,

5,42 et 9,36 Hz, 1H) ; ainsi que le système AB caractéristique du CH_2 aux protons non équivalents à δ_H 3,41 (dd, 5,46 et 11,67 Hz, 1H) et δ_H 3,59 (dd, 2,11 et 11,88 Hz, 1H). Ces déplacements chimiques étaient respectivement corrélés à δ_C 105,03 ; 73,98 ; 76,39 ; 69,77, 77,21 et 60,82 dans le spectre HSQC. Les analyses HMBC et NOESY ont permis la détermination des substitutions en montrant la non-corrélation entre le glucose et le groupe méthoxy (Figure 58). Par conséquent, il a été confirmé qu'il s'agissait d'une coumarine dihydroxy-méthoxy-5-*O*-glycosylée.

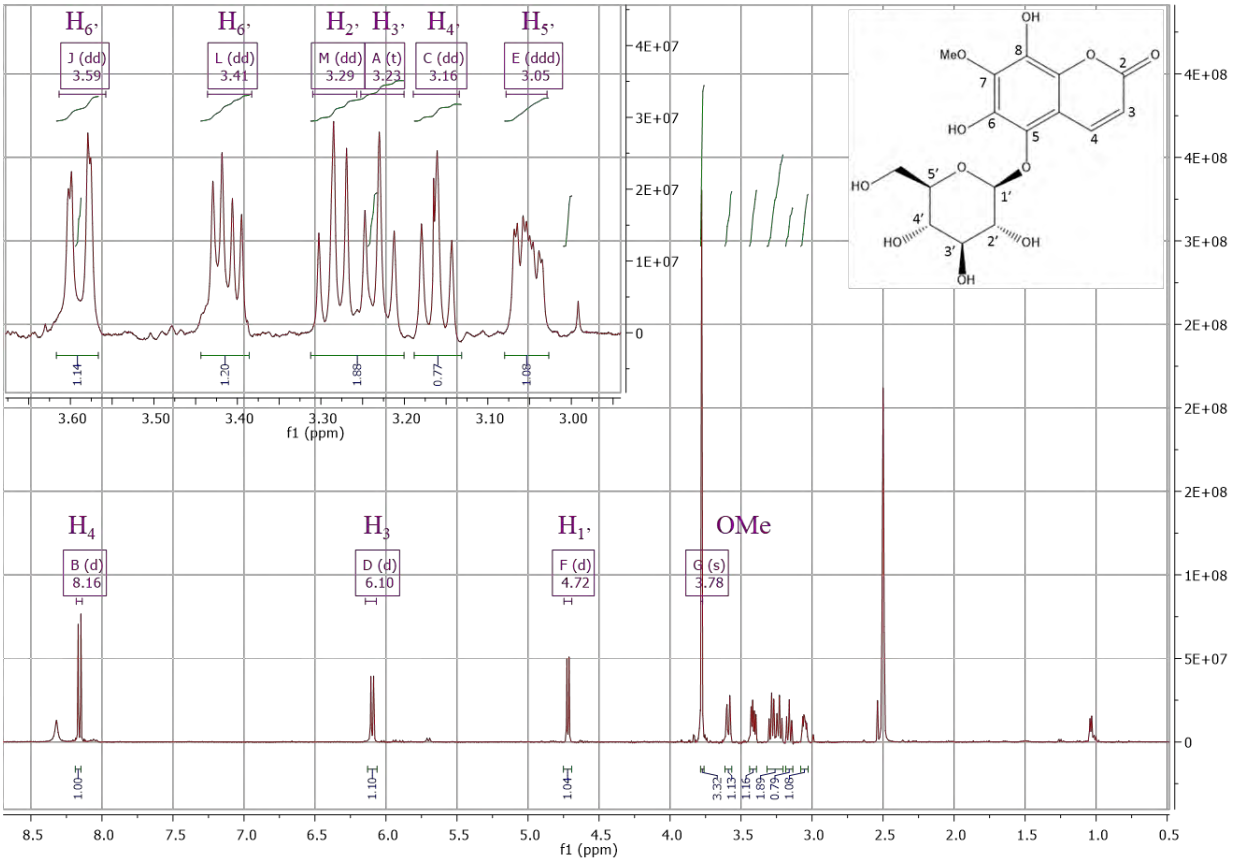


Figure 56. Spectre RMN ^{1}H du composé 1 (500 MHz, DMSO)

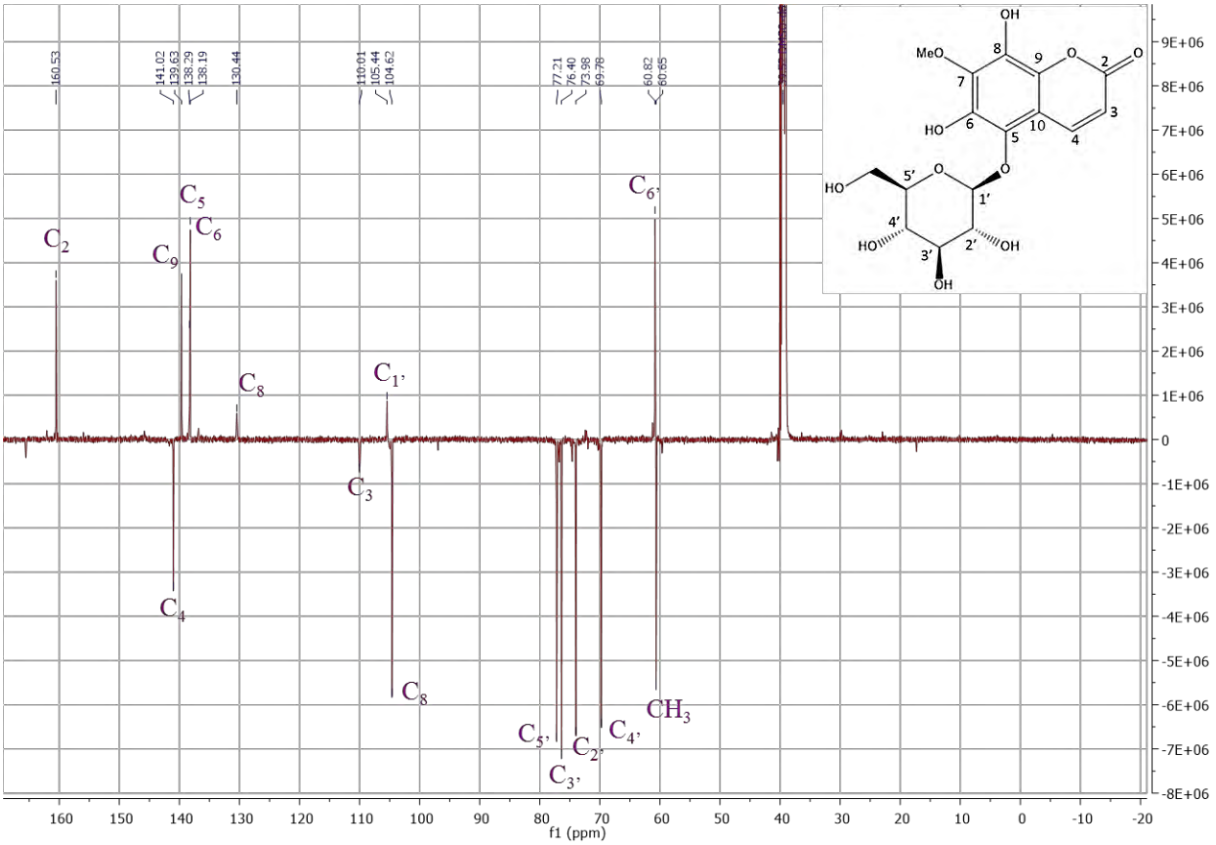


Figure 57. Spectre RMN ^{13}C du composé 1 (500 MHz, DMSO)

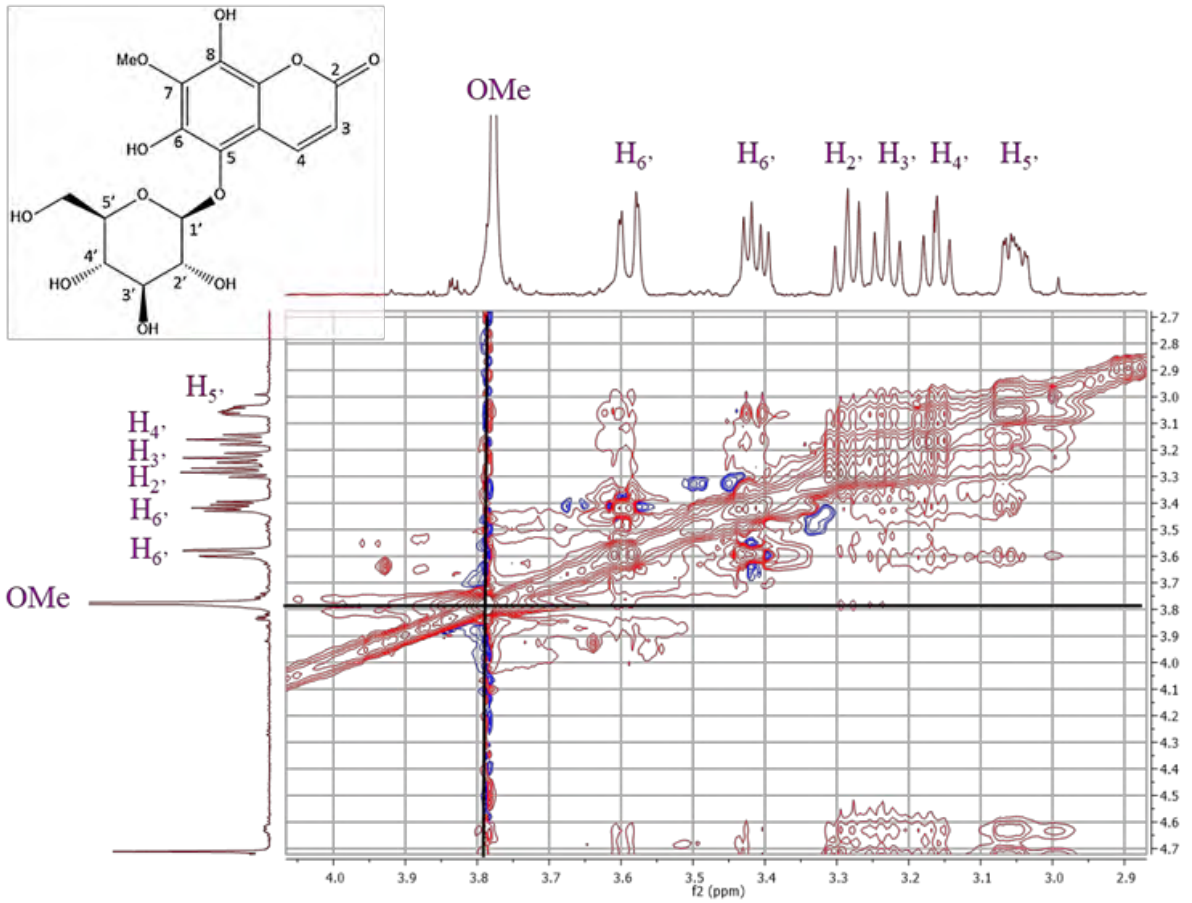


Figure 58. Spectre RMN NOESY du composé 1 (500 MHz, DMSO)

Les spectres RMN ^{1}H (Figure 59) et ^{13}C (Figure 60) de 2 (majoritaire) et 23 (minoritaire) ont révélé des signaux similaires à 1 assignés à un méthoxy à $\hat{o}_H$ 3,68 (s, 3H) et δ_C 60,25 ainsi qu'un substituant glucose avec un proton anomérique à $\hat{o}_H$ 4,63 (d, 7,75 Hz, 1H) ppm. Aucun proton aromatique n'a été observé suggérant la présence de deux groupes hydroxy sur le noyau de la coumarine. La différence de déplacement chimique pour C8 entre 1 et 2 a confirmé une coumarine dihydroxy-méthoxy-8-*O*-glycosylée pour 2. Le spectre NOESY (Figure 61) a révélé une corrélation entre le méthoxy du composé 2 et le proton 4', confirmant la substitution. Le spectre RMN ^{1}H du composé 23 en mélange avec le composé 2, mais en moindre quantité (environ 10% d'après les intégrations proton), a révélé la présence de deux groupes méthoxy avec des signaux à δ_H 3,79 (s, 3H) et δ_H 3,81 (s, 3H) ppm, ainsi qu'un glucose caractérisé par le signal d'un proton anomérique à δ_H 4,8 (d, J = 7,67 Hz, 1H) ppm. Les autres signaux de glucose n'ont pas pu être attribués en raison d'une mauvaise résolution dans la région de 3 à 4 ppm. Quelques signaux carbone ont pu être attribués à l'aide du spectre HMBC (Figure 62). Ainsi, 23 a été identifié comme étant une coumarine hydroxy-diméthoxy-*O*-glycosylée. Par ailleurs, d'après le réseau moléculaire, ce composé présente une connexion avec le composé 2 (« cosine score » > 0,9). Une différence de 14 daltons (Da) entre les deux composés peut être indicative d'un groupement méthoxy conduisant ainsi à une coumarine dimethoxy-8-*O*-glycosylée en accord avec les données RMN précédentes.

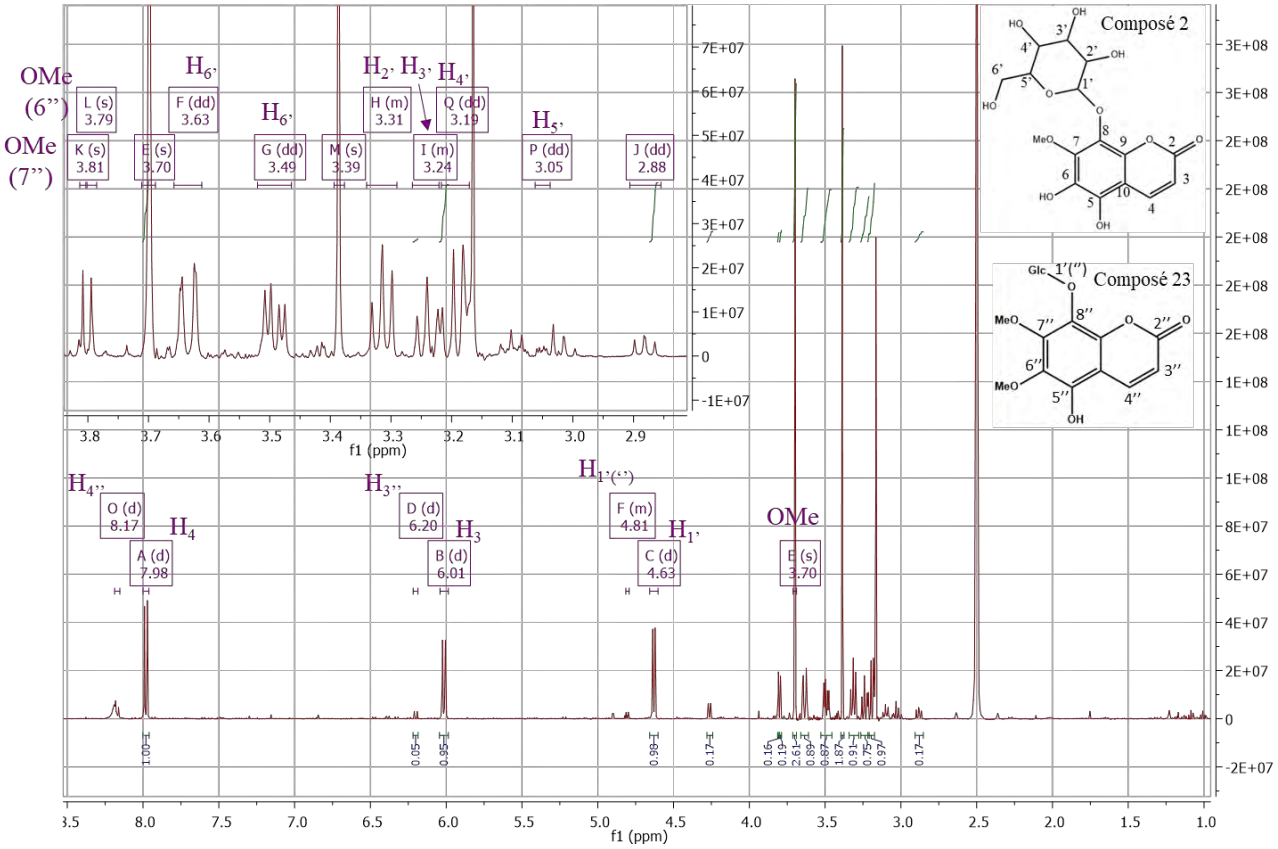


Figure 59. Spectre RMN du ^{1}H des composés 2 et 23 (500 MHz, DMSO)

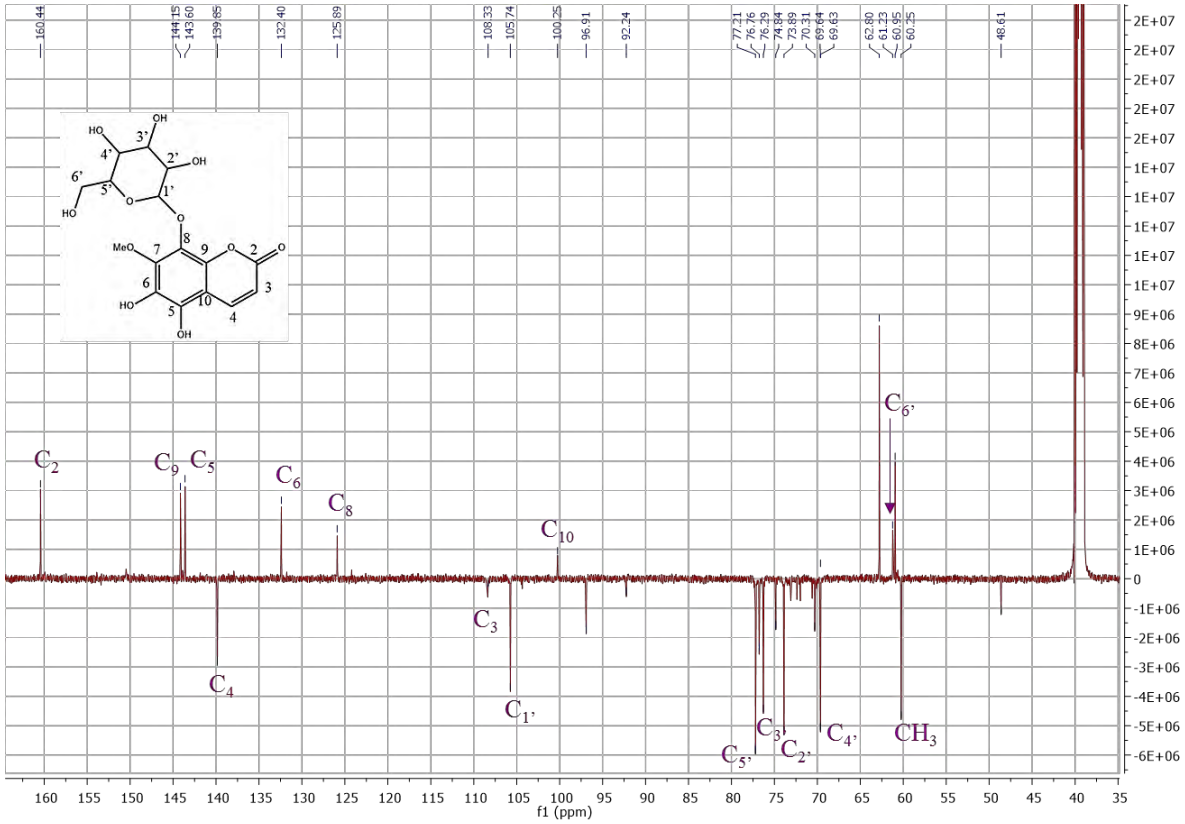


Figure 60. Spectre RMN du ^{13}C du composé 2 (500 MHz, DMSO)

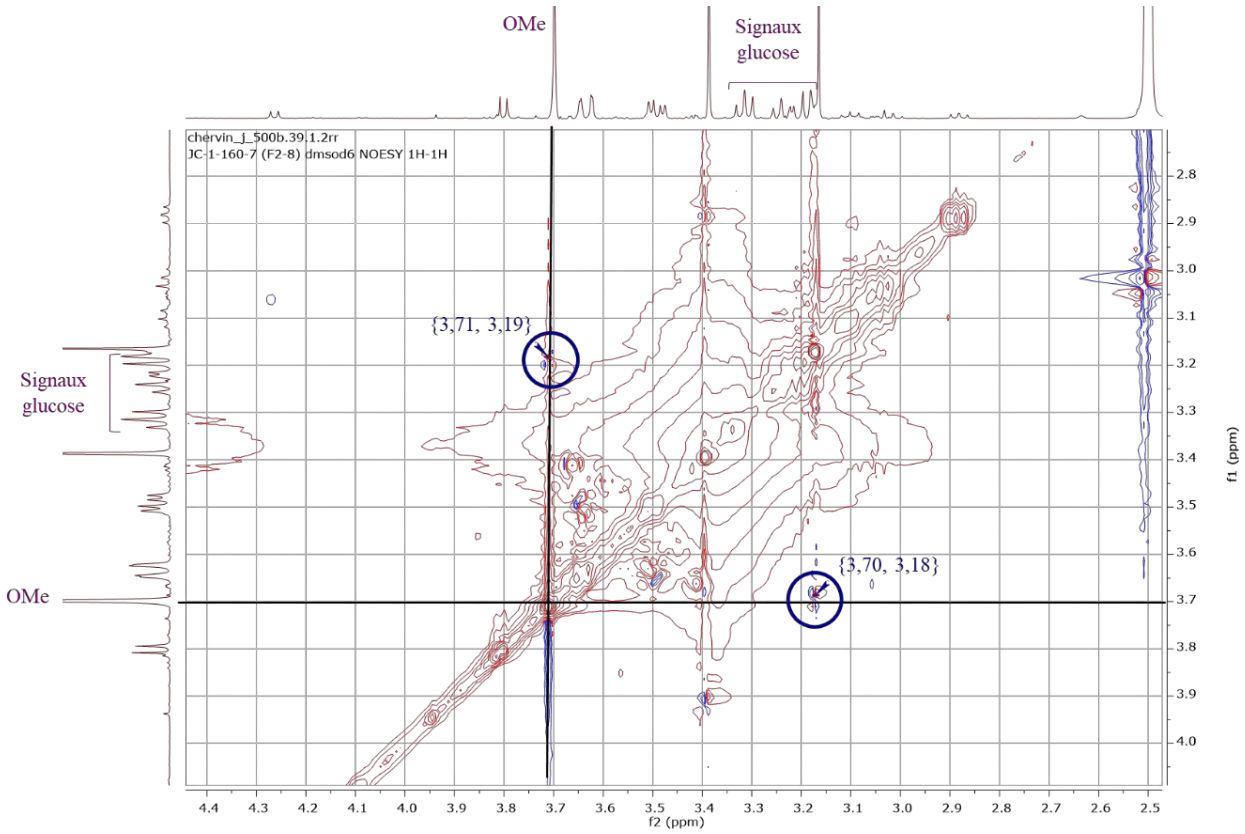


Figure 61. Spectre RMN NOESY du mélange des composés 2 et 23 (500 MHz, DMSO) avec mise en évidence d'une interaction des protons du CH_3 et du glucose justifiant la substitution proposée

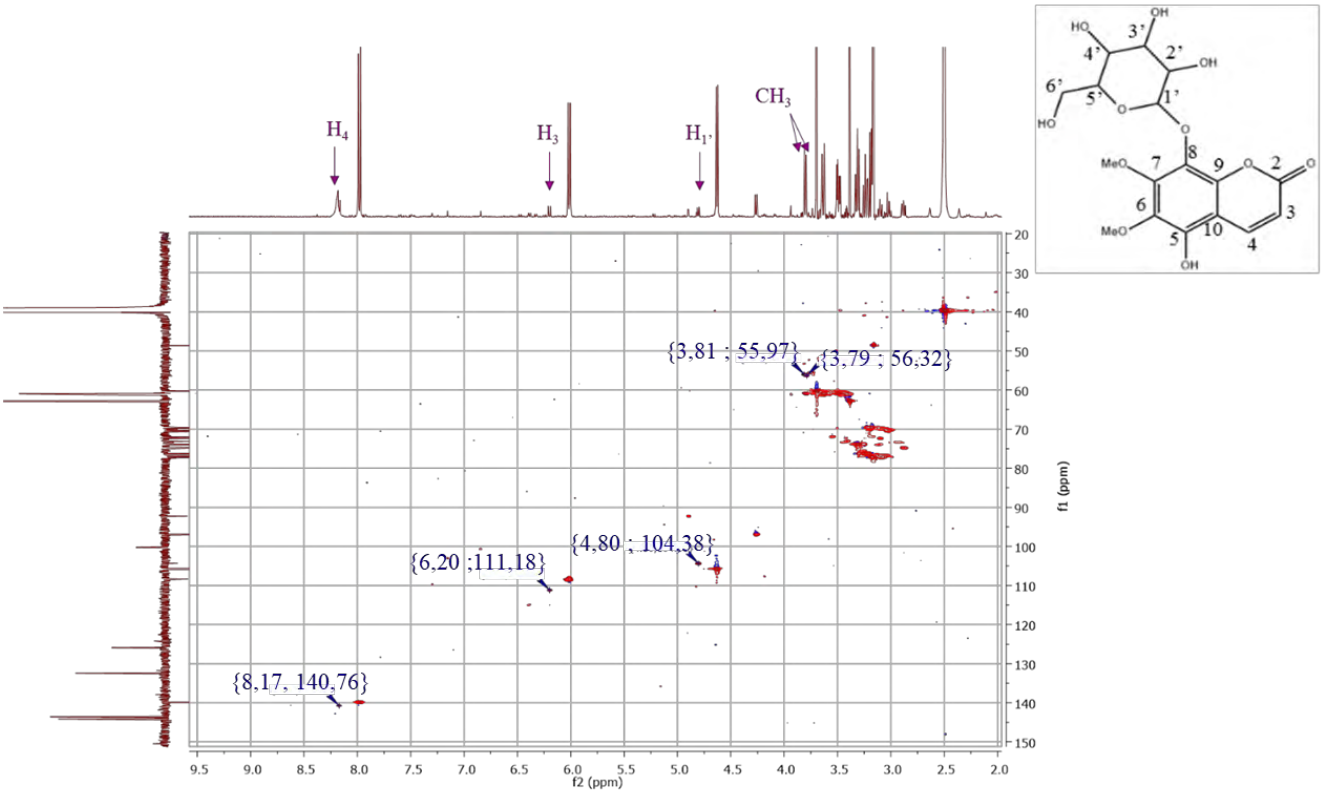


Figure 62. Extrait du spectre HMBC du composé 23 (en mélange avec 2) (500 MHz, DMSO)

L'ensemble des déplacements chimiques enregistrés pour les trois coumarines détectées est regroupé dans le Tableau 20.

Tableau 20. Résumé des déplacements chimiques des trois composés identifiés en RMN

		Composé 1		Composé 2		Composé 23	
Position	DEPT	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
Noyau							
2	CO	-	160,53	-	160,44	-	159,83
3	CH	6,10 *d* (9,63)	110,01	6,01 *d* (9,53)	108,34	6,20 d (9,72)	111,18
4	CH	8,16 *d* (9,63)	141,02	7,98 *d* (9,58)	139,85	8,17 *d* (9,69)	140,76
5	C	-	138,29	-	143,60	-	-
6	C	-	138,19	-	132,40	-	-
7	C	-	-	-	-	-	-
8	C	-	130,44	-	125,89	-	-
9	C	-	139,63	-	144,15	-	-
10	C	-	104,62	-	100,25	-	-
OMe (6)	CH3	-	-	-	-	3,81 *s*	55,97
OMe (7)	CH3	3,78 s	60,65	3,70 s	60,25	3,79 *s*	56,32
Glucose							
1'	CH	4,72 *d* (7,56)	105,03	4,63 *d* (7,75)	105,74	4,8 *d* (7,67)	104,38
2'	CH	3,29 *dd* (7,59, 8,97)	73,98	3,31 *m*	73,89	-	-
3'	CH	3,23 *m*	76,39	3,24 *m*	76,29	-	-
4'	CH	3,16 *dd* (7,97, 10,09)	69,77	3,18 *m*	69,64	-	-
5'	CH	3,05 *ddd* (2,07, 5,42, 9,36)	77,21	3,05 *m*	77,21	-	-
6' a	CH	3,41 *dd* (5,46, 11,67)	60,82	3,49 *dd* (4,94, 11,67)	61,23	-	-
6' b	CH	3,59 *dd* (2,11, 11,88)	60,82	3,63 *dd* (1,8, 11,68)	61,23	-	-

D'après le réseau moléculaire, le composé 20 présente une connexion avec le composé 2 (« cosine score » > 0,9). Sa formule brute n'ayant donné aucun résultat au niveau de l'interrogation des bases de données, une déréplication basée sur la différence de masse observée a été effectuée. Une différence de 162 Da a été interprétée comme la perte d'un glucose. Ainsi, une nouvelle 8-*O*-diglycosylée coumarine a été proposée sur la base des données MS/MS et de l'observation de la perte d'une masse de 324,12 Da s'apparentant à deux glucoses liés (Figure 63).

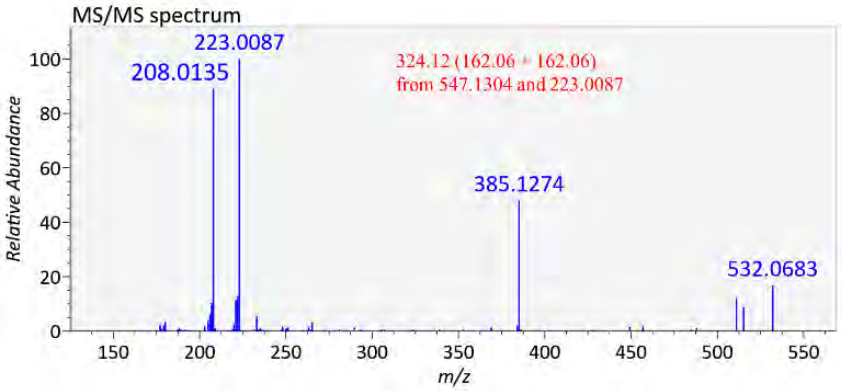


Figure 63. Spectre MS/MS du composé 20, $C_{22}H_{28}H_{16}$

L'ensemble des molécules d'intérêt a donc présenté des structures dérivées des coumarines et des flavonoïdes, connues dans la littérature pour leur potentiel antioxydant (**Procházkoá et al., 2011**; **Gacche et Jadhav, 2012**) (Tableau 21). Cette identification a été appuyée par les données UV enregistrées lors de l'analyse UHPLC-MS. En effet, les coumarines présentent deux bandes d'absorption aux alentours de 270 nm et 310 nm, les substituants au niveau du noyau pouvant avoir un effet bathochromique et conduire à une bande vers 320 nm (**Masrani et al., 1974**). Au niveau des flavonoïdes, ces bandes se situent aux alentours de 250-295 nm et 310-370 nm selon le type de flavonoïde et sa substitution (**Olsen et al., 2009**). L'ensemble des correspondances des fragments expérimentaux et *in silico* est présenté en annexe 6 page 226. La littérature a permis de confirmer les substitutions des sucres.

Tableau 21. Composés identifiés et dérépliqués ayant un potentiel lien avec l'activité antioxydante

Rang*	RT (min)	[M-H]$^-$	Formule brute	UV (nm)	Identification potentielle[a,b,c]	Classe chimique	Ref.
1	3,17	385,0771	$C_{16}H_{18}O_{11}$	260 326	7-Methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside[a,b]	Coumarine	**(Jung et al., 1995)**
2	3,77	385,0770	$C_{16}H_{18}O_{11}$	258 325	7-Methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[a,b]	Coumarine	**(Jung et al., 1995)**
20	3,00	547,1304	$C_{22}H_{28}O_{16}$	268 320	7,8-dimethoxy-5,6,8-trihydroxycoumarin-8-β glucopyranosyl-glucopyranoside[c]	Coumarine	-
21	4,24	563,1404	$C_{26}H_{28}O_{14}$	271 333	schaftozide[b]	Flavonoïde	**(Xie et al., 2003)**
5	3,96	593,1501	$C_{27}H_{30}O_{15}$	272 332	6-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxy-8-*C*-α-L-rhamnopyranosylflavone[b]	Flavonoïde	**(Vukics, Ringer, et al., 2008)**

22	4,49	371,099	$C_{16}H_{20}O_{10}$	265 320	8-C-Glucosyl-3,4-dihydro-5,7-dihydroxy-6-methoxycoumarin[b]	Coumarine	(**Marzouk at al., 2012**)
19	4,48	533,1296	$C_{25}H_{26}O_{13}$	271 331	6,8-Di-C-α-L-arabinopyranosylapigenin[b]	Flavonoïde	(**Chen Xie et al., 2003**)
23	3,81	399,0922	$C_{17}H_{20}O_{11}$	270 320	7,8-dimethoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[a,c]	Coumarine	-

*Rang basé sur le classement des coefficients de l'OPLS ; a Confirmés par RMN ; b Déterminés par fragmentation *in silico ;* c Dérépliqués à partir du réseau moléculaire

II.3.4.4 Validation de l'activité antioxydante

Un travail de purification par UHPLC-UV des fractions Fr2 et Fr3 a été amorcé (cf. IV.5.2.3.1.2, page 185) pour valider les structures des composés déterminés par MS-FINDER, et confirmer les résultats de l'OPLS quant à l'activité de ces derniers. La mise en œuvre a pu être accomplie pour la fraction Fr2, nettement moins complexe que la fraction Fr3, dont l'optimisation de la méthode d'analyse a été plus laborieuse. C'est ainsi que trois coumarines ont pu être totalement ou partiellement isolées comme évoqué précédemment. L'aboutissement de la purification de la fraction Fr3 n'a pu être obtenu en raison de nombreux problèmes techniques au niveau des instruments de purification et du temps de manipulation imparti, au cours duquel il a été observé une dégradation de certains produits. C'est pourquoi un autre protocole plus rapide, mais ne permettant de confirmer que les résultats d'activité a été proposé par spectrométrie de masse couplée avec un lecteur de plaques de chromatographie sur couche mince (TLC-MS). Une simple séparation sur plaque des fractions Fr2 avec comme solvant de migration de l'acétate d'éthyle, de l'acide formique, de l'acide acétique et de l'eau dans les proportions 50 :5,5 :5,5 :13,5 ; et Fr3 avec de l'acétate d'éthyle, de l'acide formique et de l'eau dans les proportions 60 :6,5 :6,5 a été réalisée. Le dépôt a été réalisé en bande sur une plaque 20x20 cm et une petite partie a été révélée avec le réactif DPPH à 600 M. Les bandes jaunes traduisant la présence de composés antioxydants ont ensuite été récoltées et directement injectées en masse (Figure 64).

On constate que les deux coumarines de rapport m/z égal à 385 (1 et 2) semblent être les composés majoritaires de la fraction Fr2 et les plus antioxydants, comme révélé par le modèle statistique. La coumarine de rapport m/z 547 (20) est minoritaire dans l'extrait, comme en témoigne la largeur de la tache jaune correspondante. La dernière coumarine (22) n'a pas été clairement détectée, mais pourrait se situer juste au-dessus de la tâche de 20 où l'on observe un léger éclaircissement de la plaque. Cependant, la sensibilité de la chromatographie sur couche mince et la probable faible abondance de ce constituant n'a pas permis de le confirmer.

Concernant la fraction Fr3, l'intensité des taches jaunes des flavonoïdes identifiés corrèle avec le classement statistique de l'OPLS. Il semble que cette fraction contienne d'autres composés antioxydants comme en témoigne la présence de nombreuses autres taches jaunes plus claires pouvant signaler la présence d'autres composés plus minoritaires. Encore une fois, la sensibilité de la technique n'a pas permis de les mettre en évidence ; seulement les composés majoritaires ont été détectés.

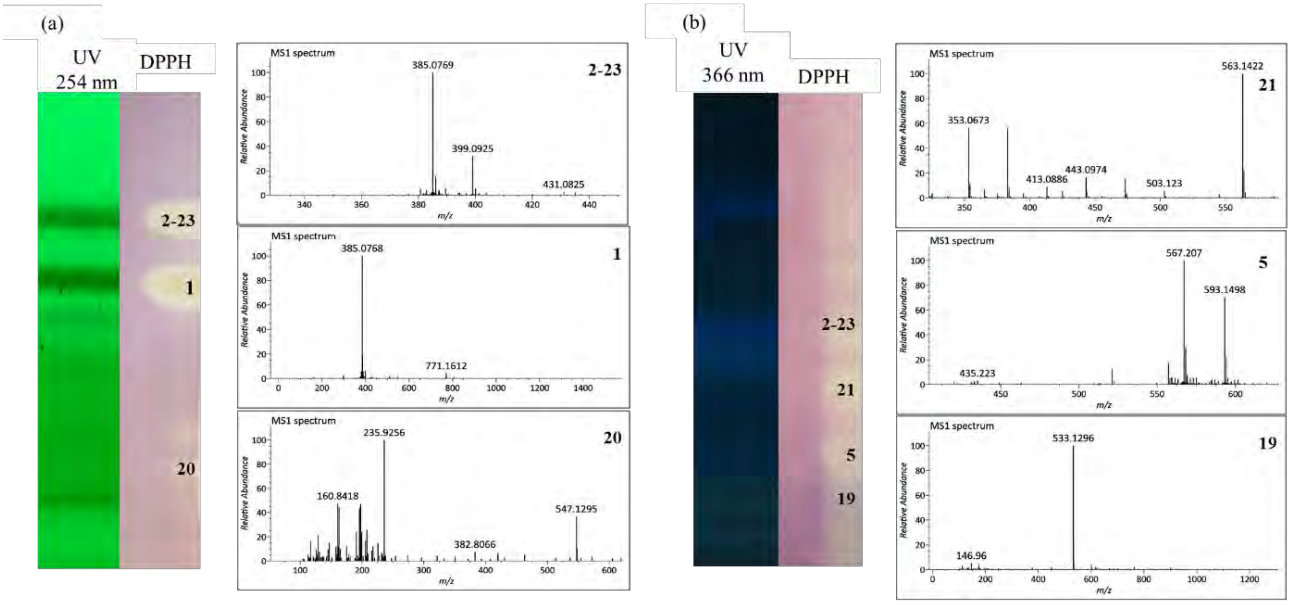


Figure 64. Séparation sur couche mince des composés antioxydants et spectres de masse associés des fractions Fr2 (a) et Fr3 (b).

Disposant tout de même des deux coumarines majoritaires purifiées, 1 et 2 (cf. Tableau 21), elles ont été testées par la méthode DPPH (diagrammes de barres simples) et par RPE avec le radical superoxyde (diagrammes de barres rayées) (Figure 65). Le trolox et la rutine ont été utilisés comme références pour les deux dosages. Les propriétés de piégeage des radicaux du composé 1 (pur à 97%) ont été confirmées par comparaison avec les témoins actifs : CI_{50} (DPPH) de 9 ± 2 mg/L contre 7 mg/L pour le trolox, et double intégration des signaux RPE de 0,88 ± 0,01 contre 1,78 ± 0,10 pour la rutine. Pour le composé 2 (pur à 90%), la CI_{50} de 49 ± 1 mg/L et la double intégration des signaux RPE de 15 ± 1 suggèrent que ce composé, bien qu'il soit moins pur que le premier, est un capteur de radicaux moins actif que le composé 1, résultat corrélant avec le classement statistique.

L'effet protecteur des coumarines contre les dommages oxydatifs dépend de la capacité donneuse d'hydrogène des groupes hydroxyles (**Kostova, 2006** ; **Borges et al., 2005**), puis de leur capacité d'oxydation. Des études antérieures sur la relation structure-activité ont permis de déterminer l'importance du nombre et de l'emplacement des groupes hydroxyles phénoliques (**Parshad et al., 2016**). Comme les composés 1 et 2 ont les mêmes substituants, des calculs de minimisation d'énergie ont été effectués en utilisant MM2 forcefield (Chemdraw 3D, Cambridge Soft, États-Unis) pour comprendre la différence entre leur capacité rédox. Ces calculs démontrent que le groupement hydroxyle en position 6 dans le composé 1 est le plus facilement oxydable en $OH^{\bullet+}$ et donc responsable de la forte capacité antioxydante mesurée pour 1 par rapport à 2.

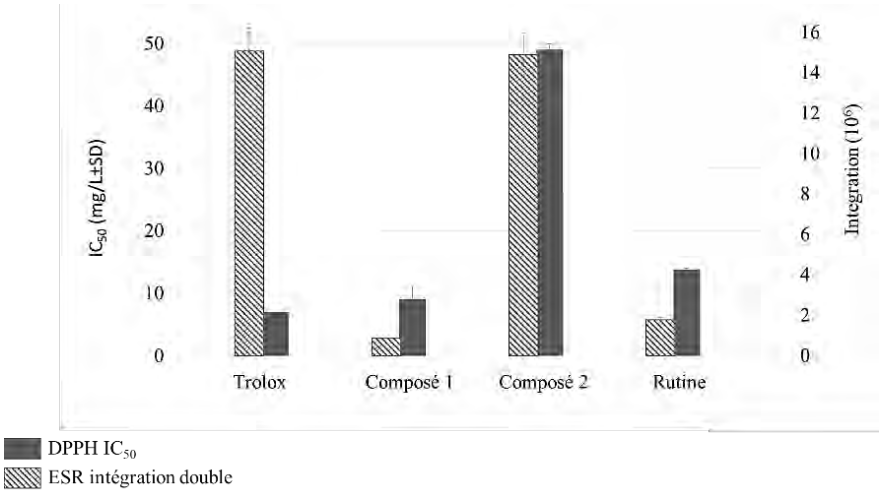


Figure 65. Capacité des composés 1 et 2 à piéger des radicaux

Ainsi, les molécules antioxydantes détectées chez la violette de Toulouse sont des composés phénoliques et principalement des flavonoïdes et des coumarines. Cette activité peut être expliquée par la présence des groupements hydroxyles OH sur la structure aromatique. En effet, l'atome d'hydrogène peut être cédé et par cette action, les molécules jouer le rôle de réducteurs.

II.3.5 Conclusion

Une étude métabolomique basée sur les profils UHPLC-HRMS combinés aux résultats biologiques et aux réseaux moléculaires a permis l'identification précoce des composés redox actifs dans un mélange complexe de violette de Toulouse. Cette approche a mis en évidence deux principaux types de métabolites, des dérivés de flavonoïdes et coumarines, présentant des groupes fonctionnels ayant des propriétés redox. Huit composés d'intérêt révélés majoritaires ont été identifiés par l'étude TLC-MS réalisée dans le cadre d'une validation de la méthode. Certaines taches bien plus minoritaires ont pu être observées, mais l'intérêt primaire étant de travailler sur le potentiel antioxydant global de la violette, l'étude des composés majoritaires a été jugée suffisante et satisfaisante avec la purification de deux coumarines avec une pureté correcte. Six d'entre eux ont été trouvés par fragmentation MS/MS et la comparaison avec des références trouvées dans les bases de données. Deux coumarines inconnues (20 et 23) ont été *de novo* dérépliquées grâce à l'interprétation du réseau moléculaire et des pertes de masse. Les résultats de déréplication ont mis principalement en évidence les coumarines dans la fraction Fr2 et les flavonoïdes dans Fr3. Ces noyaux sont connus pour avoir plusieurs activités biologiques, y compris l'activité redox comme l'ont répertorié Procházkoá et ses collaborateurs dans une revue sur les capacités antioxydantes et pro-oxydantes des flavonoïdes (**Procházkoá et al., 2011**), ou encore Gacche et Jadhav sur les dérivés de coumarines (**Gacche et Jadhav, 2012**). Un mélange de flavonoïdes *C*-glycosylés avec des coumarines *O*-glycosylées a été identifié. L'enrichissement de ces molécules dans les deux fractions explique les propriétés antioxydantes améliorées de celles-ci par rapport à l'extrait brut. Bien que certains des composés, par exemple les composés 19 et 21, soient bien connus dans le genre *Viola* (**Vukics et al., 2008** ; **Xie et al., 2003**), les autres composés identifiés n'ont pas été précédemment décrits pour ce genre.

À présent que les violettes de la collection ont été identifiées, que la méthode d'extraction et l'approche métabolomique ont été validées sur la violette de Toulouse, l'ensemble des violettes de la collection a été traité afin d'identifier les violettes, autre que la violette de Toulouse, présentant un fort potentiel biologique en général et antioxydant en particulier.

II.4 POTENTIELS BIOLOGIQUES DES VIOLETTES DE LA COLLECTION

II.4.1 CONTEXTE

La violette de Toulouse étant très sujette aux maladies, sa culture reste délicate. C'est pourquoi le criblage antioxydant de la collection de violettes a été proposé afin de mettre en évidence d'éventuelles autres espèces pouvant égaler, voire supplanter l'activité de la violette de Toulouse. Si d'autres violettes présentent un potentiel aussi intéressant et qu'elles possèdent une résistance plus importante aux phytopathogènes, il pourrait être envisagé d'augmenter leur production au sein des serres pour une éventuelle valorisation industrielle.

En complément de l'activité antioxydante, des activités phytosanitaires, plus spécifiquement antifongiques et inductrices de défense chez la plante, ont été étudiées. En effet, comme évoqué pour la violette de Toulouse, les plantes sont sujettes à de nombreuses maladies telles que le mildiou ou encore celles engendrées par *Botrytis*. Ces maladies réduisent notablement la quantité et la qualité des plantations. Elles sont aujourd'hui essentiellement contrôlées par des pesticides. Pour des raisons de santé publique, des stratégies alternatives sont recherchées, comme la stimulation des défenses des plantes à l'aide de produits naturels. Ainsi, certains métabolites secondaires ont été recherchés pour leur rôle dans les défenses des plantes contre les attaques parasitaires, en ayant soit un effet inhibiteur directement sur le parasite, soit un effet stimulant des réactions de défenses de la plante. Des études métabolomiques associées aux réponses biologiques observées ont été conduites dans le but d'identifier ces composés, et de mettre éventuellement en évidence certaines relations espèce-activité. En effet, l'observation de profils chimiques différents entre les espèces amène à penser qu'elles pourraient présenter des réponses biologiques tout aussi variées.

Finalement, les extraits étant destinés à une valorisation en dermo-cosmétique, une étude de leur cytotoxicité a été réalisée afin de valider leur innocuité.

II.4.2 CRIBLAGE DE L'ACTIVITE ANTIOXYDANTE

II.4.2.1 Résultats du test DPPH

Un premier criblage du potentiel antioxydant de la majorité des violettes de la collection a été réalisé par la méthode DPPH sur plaques 96 puits (cf. IV.6.1.1, page 191). Les résultats pour l'ensemble de la collection sont disponibles en annexe 10 page 232. Dix-huit violettes n'ont pas pu être analysées puisqu'au moment des récoltes, les pots ne présentaient pas suffisamment de feuilles pour une collecte sans dégrader le plant. Chaque échantillon de feuilles a subi le même traitement en commençant par la lyophilisation jusqu'à l'extrait éthanolique allégé en chlorophylle sur silice C_{18} (cf. IV.5.2.2.1, page 179). Finalement, onze violettes présentant une inhibition[8] équivalente ou supérieure à 50%, avec par conséquent une réserve antioxydante intéressante, ont été identifiées (Tableau 22).

[8] $\%\ inhibition = \frac{A_{éthanol} - A_{extrait}}{A_{éthanol}} * 100$, avec A = absorbance

Tableau 22. Violettes présentant un potentiel antioxydant d'intérêt

N°	Espèce	% inhibition	Origine	Couleur	Résistance
V24	*subs Rostratae*	79%	Am Nd	Violet foncé	+
V80	*subs Rostratae*	78%	ND*	ND	+
V37	*alba*	78%	ND	Blanche	-
V49	*subs Rostratae*	77%	Eurasie	Violet	-
V110	*alba*	72%	France	Violet parme	-
V109	*alba*	69%	ND	Violet	-
V7	*odorata*	63%	Angleterre	Violet	-
V9	*odorata*	60%	France	Rose	-
V69	*odorata*	56%	Angleterre	Blanche-rose	-
V16	*odorata*	49%	France	Blanche	-
V19	*alba*	48%	Italie	Blanche	-

*ND : non défini

Quatre sont particulièrement pertinentes puisqu'elles présentent de meilleurs résultats que la violette de Toulouse (V110), étudiée au paragraphe II.3, avec une inhibition d'environ 80%. Trois d'entre elles appartiennent à la sous-section *Rostratae* qui semble donc particulièrement riche en composés antioxydants. La seconde famille d'intérêt est *V. alba*, dont la violette de Toulouse fait partie, puisque trois d'entre elles ont démontré un fort potentiel antioxydant, avec une inhibition supérieure ou égale à 70%. Il est à noter que deux violettes parmi les plus antioxydantes présentent une bonne résistance aux phytopathogènes, d'après les indications de la responsable du conservatoire national des violettes aux serres municipales, et seraient donc de bons candidats pour une mise en culture.

Des études qualitatives et quantitatives ont découlé de ces premiers résultats afin de mettre en évidence les composés actifs et leur concentration dans les différentes violettes. Sur le modèle de la violette de Toulouse (II.3.4.2, page 89), une analyse métabolomique a été réalisée pour étudier plus en détail les profils chimiques des différentes violettes.

II.4.2.2 ANOVA des données DPPH

Une première analyse de variance (cf. IV.4.3.1, page 172), a été entreprise pour déterminer s'il existe une différence significative entre les différentes espèces et mettre ainsi en évidence une potentielle relation espèce/activité. Cette étape préliminaire permettra de conclure sur la pertinence d'étudier les différences de profils chimiques par l'intermédiaire des modèles statistiques ACP et OPLS, pour la détermination de métabolites d'intérêt. La configuration d'ANOVA à un facteur correspondant au résultat antioxydant est observée. La variable dépendante quantitative correspond ici au résultat biologique, qui sera expliqué par la variable qualitative, soit l'espèce des violettes. Dans ce cas, l'ANOVA à un facteur est adaptée puisque l'on dispose de plusieurs groupes indépendants (les espèces) et d'une variable quantitative (le pourcentage d'inhibition DPPH). L'hypothèse de départ (H0) est que l'ensemble des groupes sont égaux et l'hypothèse alternative (H1) est, au contraire, qu'il existe une différence permettant ainsi d'établir une relation espèce/activité. Les principaux résultats qui permettront de conclure sur la pertinence du modèle et à la différence entre les espèces sont :

- le coefficient de détermination R^2 qui donne une idée de la variabilité de la variable à modéliser. Plus il est proche de 1 et meilleur est le modèle,
- le tableau d'analyse de la variance qui permet de vérifier si la variable explicative apporte une information significative au modèle,

- le tableau d'analyse des résidus qui permet de mettre en évidence des données suspectes puisque 95% des résidus doivent se trouver dans l'intervalle [-1,96-1,96] (loi normale),
- le test de Tukey HSD (Honestly Significantly Different) qui compare deux à deux les espèces pour mettre en évidence les différences les plus significatives.

Afin d'entrer dans le cadre des ANOVA stipulant que les données doivent être normalisées, les données des résultats biologiques ont été préalablement logarithmées. L'hypothèse de départ H0 est que la variable suit une loi normale. Pour les résultats DPPH, le test de normalité Shapiro-Wilk sur les résidus centrés réduits a conduit à une *p*-value de 0,257, bien supérieure à la valeur seuil $\alpha = 0,05$. Par ailleurs, sur le graphique Q-Q (Quantile-Quantile) (Figure 66) la répartition empirique est globalement proche de la bissectrice. Ainsi, ces résultats permettent de conclure que les données suivent bien une loi normale, justifiant l'approfondissement de l'interprétation de l'ANOVA.

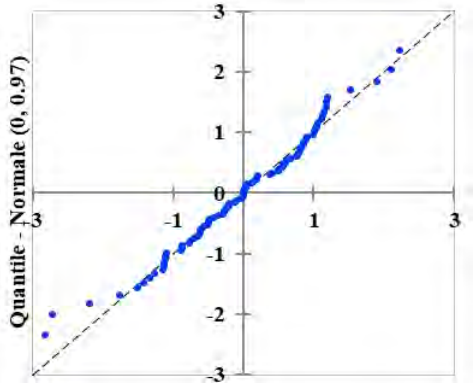


Figure 66. Graphique Q-Q de l'analyse de normalité des résidus centrés réduits des données DPPH

L'analyse ANOVA a conduit à l'obtention d'un coefficient d'ajustement R^2 égal à 0,579 indiquant que 58% de la variabilité est expliquée par l'espèce. Le test de Fisher utilisé dans l'analyse de la variance est inférieur à 0,0001 indiquant que l'on prend un risque de 0,01% en concluant que l'espèce des violettes apporte une quantité d'information significative au modèle. Finalement la présence d'une différence significative entre les espèces est confirmée par le test de Tukey HSD qui, suite aux comparaisons deux à deux des espèces, a classé les *V. alba* et les *V. subs Rostratae* dans un premier groupe (A) et toutes les autres dans un second (B) (Tableau 23).

Tableau 23. Classement des espèces sur la base du test de Tukey

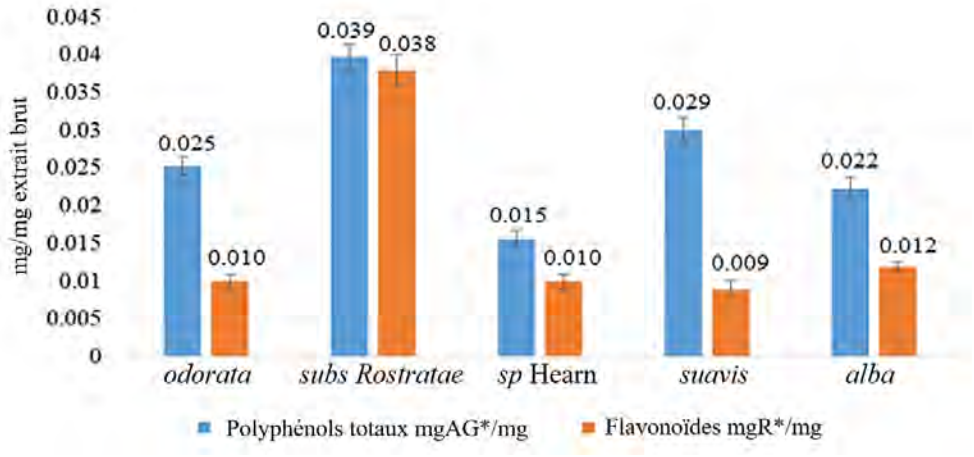
Modalité	Moyennes estimées	Erreur standard	Groupes
V. subs Rostratae	78,315	6,044	A
V. alba	61,566	4,682	A
V. odorata	30,938	1,309	B
***V. sp.* Hearn**	25,047	3,311	B
V. suavis	22,058	3,701	B
V. mandshurica	6,150	10,469	B

Ces résultats sont conformes à ce qui était attendu puisque les deux ensembles *V. alba* et *V. subs Rostratae* sont les seuls à avoir présenté un fort potentiel antioxydant, ce que le test ANOVA a permis de confirmer. Pour faire suite à la mise en évidence d'une relation espèce/activité, l'étude plus fine des profils chimiques est requise pour déterminer les métabolites d'intérêt qui feraient partie, d'après les études précédentes, de la famille des polyphénols.

II.4.2.3 Dosage des polyphénols et flavonoïdes totaux

Pour quelques échantillons de la collection, à savoir trois violettes de chaque espèce, les polyphénols totaux ont été dosés à 750 nm par la méthode de Folin-Ciocalteu (cf. IV.5.2.6.3, page 189) et les flavonoïdes par la méthode du chlorure d'aluminium $AlCl_3$ (cf. IV.5.2.6.4, page 189). Les résultats sont respectivement présentés Figure 67 en mg équivalent d'acide gallique (AG) et de rutine (R).

Le dosage Folin-Ciocalteu des polyphénols totaux a révélé que les *V. odorata* présentent une teneur moyenne de l'ordre de 2,5% dans l'extrait brut contre 1,5% pour les *V. sp* Hearn, 2,9% pour les *V. suavis*, 2,2% pour les *V. alba* et 4% pour *V. subs Rostratae*. En ce qui concerne le dosage $AlCl_3$ des flavonoïdes, les teneurs respectives déterminées ont été en moyenne de 1% pour les *V. odorata*, *V. suavis* et les *V. sp* Hearn, contre 1,2% pour *V. alba* et 3,8% pour *V. subs Rostratae*.



* en équivalent acide gallique (AG) ; ** en équivalent rutine (R)

Figure 67. Polyphénols et flavonoïdes totaux déterminés par dosage colorimétrique pour certains échantillons

L'hypothèse portant sur la teneur des polyphénols pour expliquer l'activité antioxydante n'est pas totalement vérifiée ici. En effet, bien que le dosage Folin-Ciocalteu a révélé que ce sont les *V.subs Rostratae* qui semblent les plus riches, il a montré que les *V. suavis*, *V. odorata* et *V. alba* présentent des richesses similaires. Ce constat pourrait s'expliquer par la non-sélectivité du dosage Folin-Ciocalteu qui fait intervenir des composés autres que ceux ciblés, comme les acides aminés ou les sucres libres en forte concentration (**Singleton et al., 1999**). Le dosage par la méthode $AlCl_3$ a conduit à des conclusions similaires. *V. alba* présente tout de même une teneur légèrement supérieure aux trois autres espèces, et la disparité de structures des flavonoïdes pourrait expliquer l'activité antioxydante prépondérante de *V. alba* et *V. subs Rostratae*. Une étude plus détaillée a donc ensuite été entreprise par la mise en œuvre de modèles statistiques pour identifier les métabolites actifs.

II.4.2.4 ACP et OPLS des données DPPH

L'ACP réalisée pour la phylogénie (Figure 30, page 64) a permis d'attester que les espèces de violettes présentaient des profils chimiques différents par leur séparation dans l'espace de variance, avec la formation d'un groupe homogène par espèce. Ainsi, une analyse statistique basée sur les coefficients de régression OPLS, avec pour données Y les résultats du test DPPH, devrait permettre de mettre en évidence les molécules qui seraient les plus impliquées dans cette activité, sur la même stratégie de modélisation décrite au paragraphe II.3.4.2, page 89. Les résultats étant présentés en pourcentage d'inhibition, ce sont les coefficients les plus élevés qui sont liés aux composés statistiquement antioxydants.

Les structures ont été déterminées par fragmentation *in silico* avec MS-FINDER et avec l'utilisation de diverses bases de données de produits naturels, dont certaines dérivées du genre *Viola* (Tableau 24). Elles restent hypothétiques à ce stade de l'étude, mais donnent une bonne idée de la structure et surtout de la classe chimique des composés. Ainsi, il ressort que ce sont principalement des coumarines et des flavonoïdes qui semblent conférer cette activité antioxydante, puisque ce sont ces types de structures

qui caractérisent les premiers rangs du classement des coefficients OPLS. Ces résultats sont en accord avec la précédente étude de la violette de Toulouse, qui avait mis en relief l'activité de quatre coumarines *O*-glycosylées et de trois flavonoïdes *C*-glycosylées.

Dans cette étude complète de la collection, six composés d'intérêt, majoritaires chez les violettes issues de *V. alba* et *V. subs Rostratae*, ont été révélés. Le résultat des moyennes des intensités conduit à remarquer que les coumarines semblent être majoritairement retrouvées chez les violettes appartenant à l'espèce *V. alba,* tandis que les flavonoïdes antioxydants seraient plutôt caractéristiques des *V. subs Rostratae*. La majorité d'entre eux (composés 1, 2, 5 et 21) avait été précédemment identifiée chez la violette de Toulouse et leur activité confirmée par TLC-MS. La rutine (24), connue dans la littérature pour son activité antioxydante et produit ubiquitaire de la majorité des extraits foliaires, a été retrouvée chez les *V. subs Rostratae* et classée en première position du classement des coefficients de la régression OPLS, tandis qu'un autre flavonoïde *O*-glycosylé occupe la seconde place. Les trois flavonoïdes statistiquement révélés les plus antioxydants présentent tous un noyau catéchol lié à la structure flavonoïde. Par ailleurs les deux premiers possèdent également une double liaison 2,3 conjuguée avec une fonction oxo en position 4 (cf. Figure 9, page 41). Ainsi, l'ordre d'activité des flavonoïdes semble ici confirmé par ces différences structurales connues comme responsables de l'activité antioxydante (**Bors, 1990**).

La richesse antioxydante de la violette de Toulouse est globalement retrouvée justifiant son fort potentiel antioxydant. De nouveaux composés ont cependant été identifiés suite au traitement de la collection et ces résultats sont certainement à rapprocher de la concentration de ces derniers dans les extraits. En effet, ils ont été retrouvés dans les extraits de la violette de Toulouse, mais leur intensité est nettement inférieure à celle dans les extraits des *subs Rostratae* (en moyenne $2x10^4$ contre $5x10^2$), ce qui justifie qu'ils n'aient été décelables que suite à cette étude.

Tableau 24 Identifications des potentiels composés antioxydants déterminés par analyse statistique

N°	m/z	RT (min)	Formule brute	Classe	Identification	coefficient OPLS	Espèce
24	609,1429	4,22	$C_{27}H_{30}O_{16}$	Flavonoïde	rutine[a]	5,93248	*subs Rostratae*
4	755,2006	3,92	$C_{33}H_{40}O_{20}$	Flavonoïde	3,3',4',5,5',7-Hexahydroxyflavone; 3-*O*-[α -L-Rhamnopyranosyl-(1,2)-α -L-rhamnopyranoside], 3'-*O*-α -L-rhamnopyranoside[a]	5,86872	*subs Rostratae*
5	593,1482	3,68	$C_{27}H_{30}O_{15}$	Flavonoïde	6-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxy-8-*C*-α-L-rhamnopyranosylflavone[a]	4,60812	*subs Rostratae*
1	385,0757	2,94	$C_{16}H_{18}O_{11}$	Coumarine	7-Methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside[a,b]	4,42717	*alba*
21	563,1380	3,92	$C_{26}H_{28}O_{14}$	Flavonoïde	8-*C*-β-D-Glucopyranosyl-4',5,7-trihydroxyflavone; 6"-*O*-α-L-Arabinopyranosyl[a]	3,72688	*subs Rostratae*
2	385,0757	3,47	$C_{16}H_{18}O_{11}$	Coumarine	7-Methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[a,b]	3,71295	*alba*

[a] Composé trouvé en interrogeant la base de données *Viola* extraite du DNP ; [b] Composé confirmé par RMN

Pour compléter cette étude qualitative, des dosages des coumarines et des flavonoïdes antioxydants par UHPLC-UV à 325 et 360 nm (cf. IV.5.2.6.1 page 188 et IV.5.2.6.2 page 188) ont été réalisés pour estimer leur concentration dans les différentes espèces de violettes.

II.4.2.5 Dosage UHPLC-UV des flavonoïdes et coumarines antioxydants

Quatre coumarines (Tableau 25) et cinq flavonoïdes (Tableau 26**)** antioxydants (cf. II.3.4.3 page 90 et II.4.2.4 page 105) ont été identifiés, à partir des données HRMS/MS et des spectres UV. Les solutions ayant été injectées directement après extraction et purification sur silice, les résultats sont présentés en pourcentage de polyphénols dans la poudre en prenant en compte le volume final de solution obtenue après traitement. En effet, la silice absorbe environ 35% de solvant conduisant à une réduction de volume et donc à une concentration des composés, qui a été prise en compte.

Tableau 25. Coumarines dosées dans les extraits de violette de la collection

N°	m/z	RT (min)	Formule brute	Identification*
1	385,07	2,9	$C_{16}H_{18}O_{11}$	7-methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside
2	385,07	3,4	$C_{16}H_{18}O_{11}$	7-methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside
23	399,09	3,4	$C_{17}H_{20}O_{11}$	7,8-dimethoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside
20	547,13	2,8	$C_{22}H_{28}O_{16}$	7,8-dimethoxy-5,6,8-trihydroxycoumarin-8-β glucopyranosyl-glucopyranoside

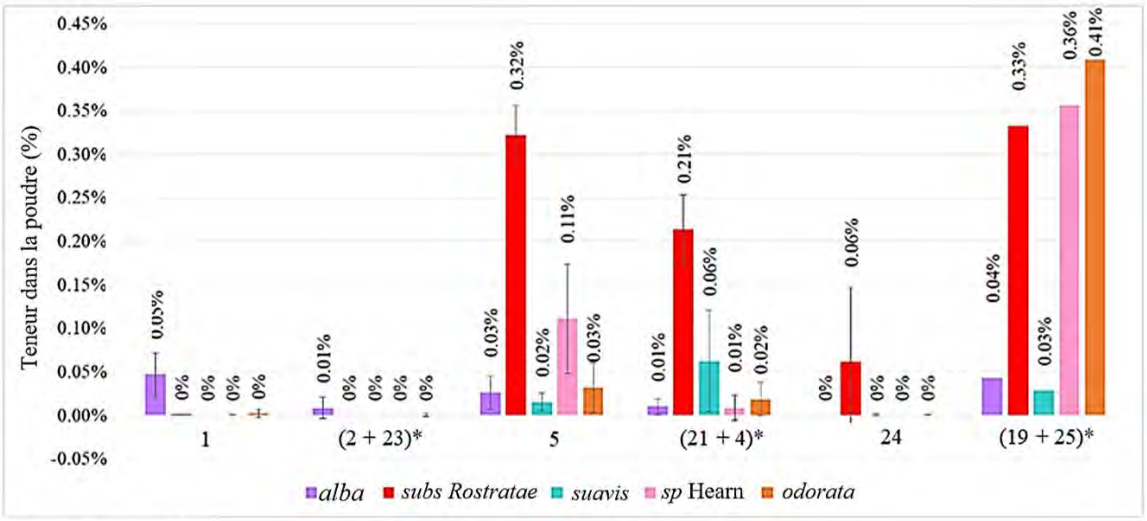
* identifications basées sur la comparaison des spectres de masse expérimentaux et *in silico*

Tableau 26 : Flavonoïdes dosés dans les extraits de violette de la collection

N°	m/z	RT (min)	Formule brute	Identification*
5	593,14	3,66	$C_{27}H_{30}O_{15}$	6-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxy-8-*C*-α-L-rhamnopyranosylflavone
21	563,14	3.84	$C_{26}H_{28}O_{14}$	8-*C*-β-D-glucopyranosyl-4',5,7-trihydroxyflavone; 6"-*O*-α-L-arabinopyranosyl
4	755,20	3,90	$C_{33}H_{40}O_{20}$	3,3',4',5,5',7-Hexahydroxyflavone; 3-*O*-[α -L-Rhamnopyranosyl-(1,2)-α -L-rhamnopyranoside], 3'-*O*-α -L-rhamnopyranoside
19	533,12	4,11	$C_{25}H_{26}O_{13}$	6,8-Di-*C*-α-L-arabinopyranosylapigenin
24	609,14	4,22	$C_{27}H_{30}O_{16}$	rutine
25	577,15	4,13	$C_{27}H_{30}O_{14}$	8-*C*-β-D-glucopyranosyl-4',5,7-trihydroxyflavone; 2"-*O*-α-L-rhamnopyranosyl

* identifications basées sur la comparaison des spectres de masse expérimentaux et *in silico*

Les teneurs moyennes des composés dans la poudre sont présentées Figure 68, pour les cinq espèces. Il est à noter que le flavonoïde 25 a été identifié en co-élution avec le composé 19. Le résultat du dosage de ces composés n'est donc pas spécifique. Il a été observé que le composé 19 était prépondérant chez *V. subs Rostratae* et *V. alba* tandis que le composé 25 l'était chez *V. odorata* et *V. sp* Hearn.



* co-élution en UV

Figure 68. Moyenne de la teneur dans la poudre des coumarines et des flavonoïdes antioxydants déterminés par dosage UHPLC-UV pour chaque espèce

Certaines violettes qui avaient été révélées très antioxydantes par la méthode DPPH ont présenté de forts taux en coumarines, dont les deux coumarines 1 et 2. La coumarine 20 n'a pas été détectée par la méthode automatique, probablement en raison de sa faible concentration ne permettant pas une bonne résolution du pic chromatographique. Une vérification manuelle a donc été effectuée, mais il s'est avéré que comme suspecté, le pic n'était pas suffisamment défini pour obtenir une valeur d'intégration correcte. Ainsi, il a été décidé de l'écarter de l'analyse puisqu'il n'impacterait que très faiblement le résultat final. Les deux coumarines 1 et 2 et la 23 sont majoritaires chez les violettes de l'espèce *V. alba* qui occupent cinq des six premiers rangs du classement, avec une teneur environ cinq fois supérieure aux autres violettes. En moyenne, les coumarines représentent près de 0,05% de la matière végétale lyophilisée de cette espèce. L'une d'entre elles présente même une teneur supérieure de l'ordre de 0,1%. Pour les autres violettes, le pourcentage de coumarines ne dépasse pas 0,01%. Par ce résultat, on confirme que les coumarines sont des composés plutôt spécifiques de *V. alba* car contenues en quantité nettement plus importante, en moyenne cinq fois plus que chez les autres espèces.

Les précédents résultats du dosage des flavonoïdes ne concordaient pas totalement avec les violettes les plus antioxydantes puisque même si les *V. subs Rostratae* sont les plus riches, les autres espèces présentent des teneurs comparables. Par conséquent, il apparait que la teneur en flavonoïdes ne semble pas directement corréler au potentiel antioxydant des violettes, et ceci pourrait tout à fait s'expliquer par la présence de flavonoïdes abondants, mais présentant un plus faible potentiel antioxydant que d'autres plus minoritaires. En étudiant plus en détail la structure des flavonoïdes en présence (Figure 69), on constate que deux d'entre eux, les 4 et 24, présentent une liaison 2,3 conjuguée avec une fonction oxo en position 4, et que deux composés, les 5 et 24, présentent une autre caractéristique d'intérêt, un noyau catéchol relié à la structure flavonoïde. Ainsi, la capacité antioxydante est issue de la richesse en flavonoïdes aux structures du même type. Ce constat est confirmé par l'identification d'autres flavonoïdes, dont le composé 25 prépondérant chez *V. odorata* et *V. sp* Hearn, jugés non antioxydants et dont la structure n'a révélé aucune des entités évoquées, ainsi qu'une fonction hydroxyle en moins.

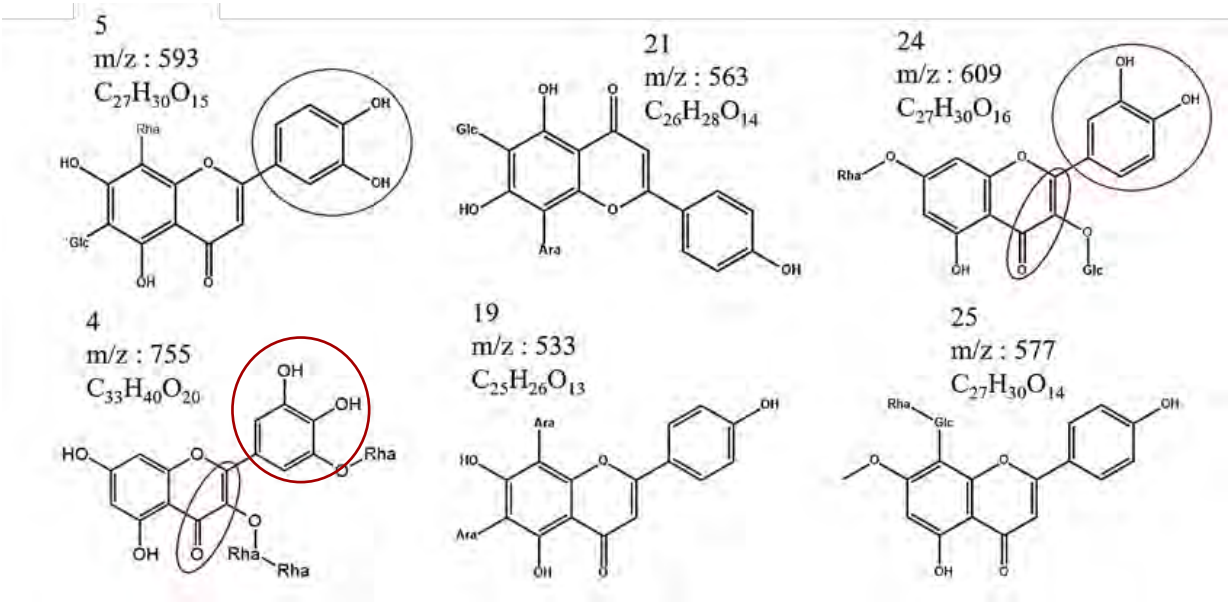


Figure 69. Structures chimiques des flavonoïdes dosés

En moyenne, les flavonoïdes antioxydants représentent 0,9% de la matière végétale lyophilisée pour *V. subs Rostratae*, c'est-à-dire un quart des flavonoïdes totaux. Les autres violettes sont riches en flavonoïdes, mais sont moins concentrées en flavonoïdes antioxydants, avec un rapport moyen de 1/10 par rapport aux *subs Rostratae*. Par ce résultat, on atteste que les flavonoïdes ne sont pas spécifiques d'une espèce en particulier, mais ceux antioxydants caractérisés par certaines entités structurales, sont nettement majoritaires chez *V. subs Rostratae*, justifiant le meilleur potentiel biologique de cet ensemble par rapport aux autres espèces. La capacité antioxydante des *V. alba* serait à rapprocher de sa concentration en coumarines puisque sa teneur en flavonoïdes antioxydants reste comparable aux autres.

II.4.2.6 Conclusion

Le criblage antioxydant de la majeure partie de la collection nationale de violettes a permis de mettre en évidence une dizaine de violettes d'intérêt, dont six avec un fort potentiel : trois appartenant à *V. subs Rostratae* et trois à *V. alba*. Les principaux composés suspectés actifs sont les coumarines et les flavonoïdes, comme révélés par l'étude de la violette de Toulouse. Le dosage des polyphénols a permis de mettre en évidence que les coumarines sont des composés majoritaires chez *V. alba.* Concernant les flavonoïdes, cette classe chimique est représentée chez toutes les espèces à des concentrations variables. L'activité antioxydante n'est pas directement liée à la teneur totale en flavonoïdes, mais certains sont majoritaires chez *V. subs Rostratae* et confèrent très certainement ce potentiel supérieur par rapport aux autres espèces. Deux d'entre eux, les flavonoïdes 5 et 24, possèdent un noyau catéchol, connu comme déterminant structural de l'activité antioxydante, puisque ce sont de bons pièges en raison de la stabilité des radicaux formés. Leur caractéristique et leur teneur chez les *V. subs Rostratae* font d'eux d'excellents discriminants antioxydants, dont l'activité de l'un (5) avait été précédemment validée, et celle de l'autre (24) est largement relatée dans la littérature (**Yang et al., 2008**).

Dans une optique de développement d'une gamme cosmétique liée à des propriétés antioxydantes, il sera donc essentiel de travailler à partir de violettes identifiées chez *V. alba* ou *V. subs Rostratae*. Cependant, les autres espèces pourraient présenter un intérêt tout autre, par exemple dans le domaine phytosanitaire exploré en collaboration avec le laboratoire de recherche en sciences végétales (LRSV). C'est pourquoi un criblage complémentaire de la collection a été effectué sur d'autres types d'activités biologiques.

II.4.3 CRIBLAGE DES ACTIVITES PHYTOSANITAIRES

L'objectif de cette étude était de mettre en évidence des violettes présentant des propriétés phytosanitaires et de pouvoir identifier les molécules responsables de ces activités par un criblage LC-HRMS couplé à une approche métabolomique. En effet, les violettes sont connues dans la littérature pour présenter de nombreuses propriétés biologiques et il a été décidé, en lien avec nos partenaires du LRSV, d'étudier deux types d'activités à savoir, une activité antifongique basée sur le criblage de cinq souches de champignons, et une activité de défense des plantes basée sur l'expression du gène Gus. En effet, une méthode alternative aux pesticides consisterait à booster le système immunitaire de la plante à l'aide d'élliciteurs (**Burketova et al., 2015**) par exemple le methylsalicylate constituant de l'huile essentielle de *Gaultheria procumbens* (**Vergnes et al., 2014**) ou encore les oxylipines, produits naturels dérivées des acides gras insaturés qui sembleraient contribuer aussi bien à une activité antimicrobienne directement sur les pathogènes, qu'à une activité inductrice de défense de la plante (**Prost, 2005**).

II.4.3.1 Criblage de l'activité antifongique

II.4.3.1.1 Résultats d'inhibition de la croissance de cinq souches

Un premier criblage de la collection a été fait sur trois souches phytopathogènes de la famille des Ascomycètes, *Fusarium graminearum*, *Colletotrichum higginsianum* et *Botrytis cinerea,* par Mme Audonnet, stagiaire de licence professionnelle productions végétales (parcours génome et biotechnologie pour l'amélioration des plantes, GeBAP) (cf. IV.6.3.1, page 195). Pour étendre la gamme de souches étudiées, certains échantillons, ceux révélés d'intérêt lors des premiers tests et un échantillonnage ayant présenté une réponse négative, ont ensuite été testés sur deux autres souches phytopathogènes, l'Ascomycète *Alternaria brassicicola* et l'Oomycète *Phytophthora capsici*. Une description des cinq souches est disponible ci-dessous :

- *Botrytis cinerea e*st un champignon nécrotrophe de la famille des Sclerotiniaceae, de la division des Ascomycètes (Figure 70 a). Il est responsable de la pourriture grise, une pathologie qui touche de nombreuses cultures d'intérêt tels que la vigne, le tournesol ou encore la tomate (Figure 70 b**)**.

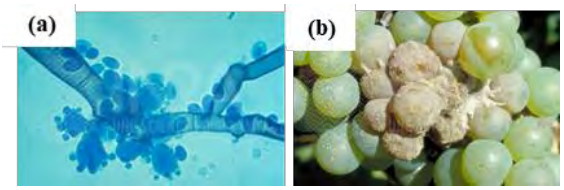


Figure 70. *Botrytis cinerea* vu au microscope (a) et foyer de *Botrytis* sur une grappe (b) (images issues de la médiathèque de l'INRA)

- Le genre *Colletotrichum* (division des Ascomycètes), contient de nombreuses espèces pathogènes sur de nombreuses plantes, dont les céréales et légumineuses. Après avoir envahi la plante à l'aide de cellules spécialisées, le champignon se développe dans une première phase à l'intérieur des cellules végétales sans dommages puis dans une seconde phase, il détruit les tissus végétaux colonisés dont il se nourrit. *Colletotrichum higginsianum* (Figure 71 a) s'attaque préférentiellement aux plantes de la famille des Brassicacées et est responsable de la maladie appelée anthracnose, caractérisée par des tâches nécrotiques (Figure 71 b).

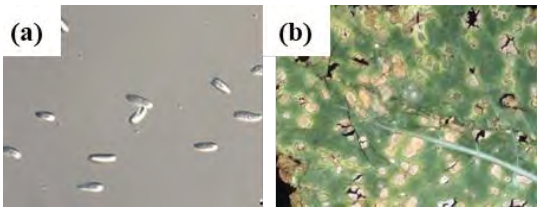


Figure 71. *Colletotrichum higginianum* vu au microscope (a) et anthracnose sur une feuille de chou (image issue de University of Georgia Plant Pathology, University of Georgia, Bugwood.org) (b)

- *Fusarium graminearum* est un champignon phytopathogène de la famille des Deutéromycètes (champignons imparfaits filamenteux). Son nom est tiré de la forme de ses spores en forme de fuseau. *Fusarium graminearum* (Figure 72 a) se développe sur de nombreuses plantes, principalement les graminées. Il est responsable de la fusariose (Figure 72 b) se manifestant par des tâches de décoloration virant au noir et accélérant le flétrissement de la plante.

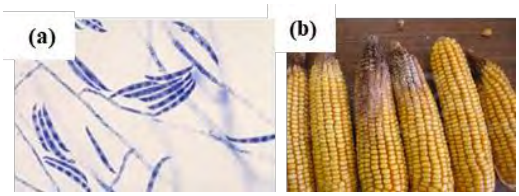


Figure 72. *Fusarium graminearum* vu au microscope (a) et divers niveaux d'attaque sur épis de maïs (b) (images issues de la médiathèque de l'INRA)

- *Phytophthora capsici* est un Oomycète (Figure 73 a) principalement connu en France pour affecter les poivrons ; mais il a également une incidence sur les Cucurbitacées. Il attaque aussi bien les racines que les tiges, les feuilles ou encore les fruits de la plante hôte. La maladie se manifeste par des lésions au niveau des tiges avec une décoloration des tissus, des tâches vertes au niveau des feuilles qui vont ensuite brunir (Figure 73 b), ou encore des spores blanches qui font penser à du sucre à glacer à la surface du fruit.

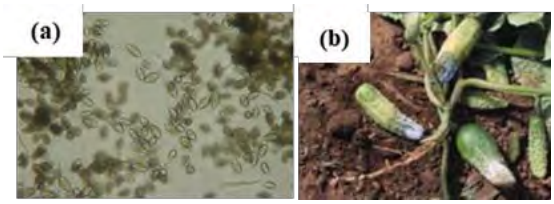


Figure 73. *Phytophtora capsici* vu au microscope (a) et manifestation de l'infection sur des concombres (b) (images issues de Hausbeck et Lamour, 2004)

- *Alternaria brassicicola* appartient au genre *Alternaria* qui comprend plus de 100 espèces ubiquitaires répandues dans les aliments et dans l'environnement. La majorité des espèces est fortement liée à une famille de plante, voire à une plante en particulier. *Alternaria brassicicola* (Figure 74 a) est responsable de la maladie de la tache noire (ou « black spot ») des Brassicacées. Cette tâche nécrotique est entourée par un halo jaune résultat de l'action des toxines qui tuent les cellules végétales en périphérie de la lésion (Figure 74 b).

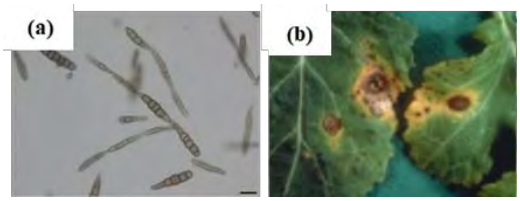


Figure 74. *Alternaria brassicicola* vu au microscope (a) et illustration de la maladie de la tâche noire sur une feuille de colza (b) (images issues de la médiathèque de l'INRA)

Le pourcentage d'inhibition a été déterminé avec les solutions à 10 mg/mL d'extrait brut dans le méthanol 50%, diluées au 10ème en solution finale. La mesure de la densité optique (DO) obtenue au 5ème jour a servi à évaluer l'activité. Une correction des DO mesurées a été faite en soustrayant la valeur obtenue au 1er jour. Une solution de MeOH 50%, solvant de dilution des extraits, a dans un premier temps été testé afin de vérifier son innocuité par rapport au test effectué, et les résultats ont été concluants puisque l'absorption était comparable à celle de l'eau, avec un écart-type de l'ordre de 2 à 4% (Figure 75).

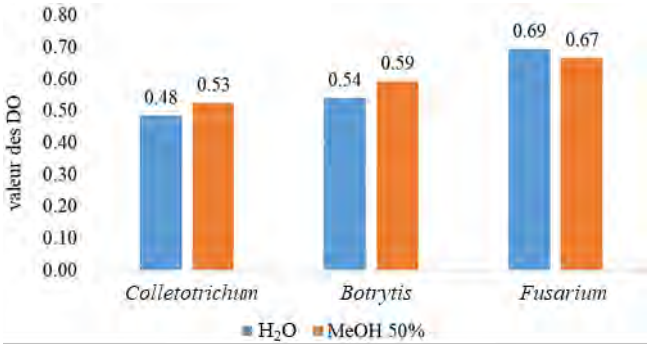
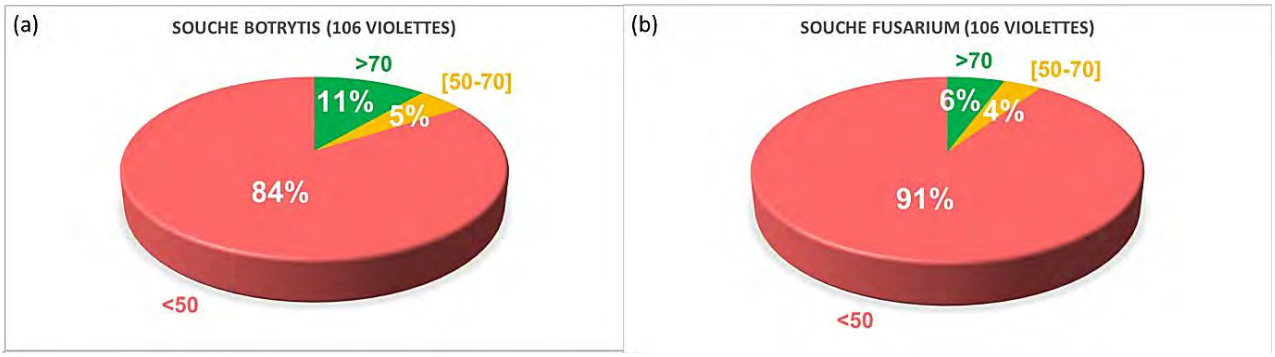


Figure 75. Valeurs des densités optiques de l'eau et du méthanol 50% sur les souches *Colletotrichum*, *Fusarium* et *Botrytis*.

En effet, dans le cas de la souche *Fusarium* par exemple (Figure 76), il est constaté la même couleur beige foncé au niveau des puits de l'eau et du méthanol 50%, contre un milieu incolore pour l'antifongique (AGN 23), révélant la destruction totale du champignon.

Figure 76. Photos du milieu au contact de l'antifongique, du méthanol 50% et de l'eau sur *Fusarium*

Les cinq souches ont présenté des sensibilités variées aux extraits de violette. Le premier résultat obtenu a été la non-activité de l'ensemble des extraits de la collection sur la souche *Colletotrichum*. En effet, tous les extraits ont présenté des inhibitions inférieures à 50%. De plus, une grande majorité a même donné des inhibitions négatives, traduisant une stimulation de la croissance du champignon. En ce qui concerne les autres souches, quelques échantillons se sont démarqués puisque respectivement 10% et 16% des 106 violettes analysées ont présenté une inhibition supérieure à 50% pour *Fusarium* et *Botrytis*, dont respectivement 6% et 11% supérieures à 70% (Figure 77 a et b). Concernant les deux autres souches testées, seulement 28 extraits de violettes ont été analysés, en raison du temps de manipulation imparti dans le cadre du stage de Mme Audonnet, et les résultats ont montré que ces deux souches semblent particulièrement sensibles avec respectivement 43% et 61% d'extraits ayant présenté une inhibition supérieure à 70% (Figure 77 c et d), dont 6 ayant donné une réponse négative pour les deux premières souches étudiées, soit un quart des extraits testés.



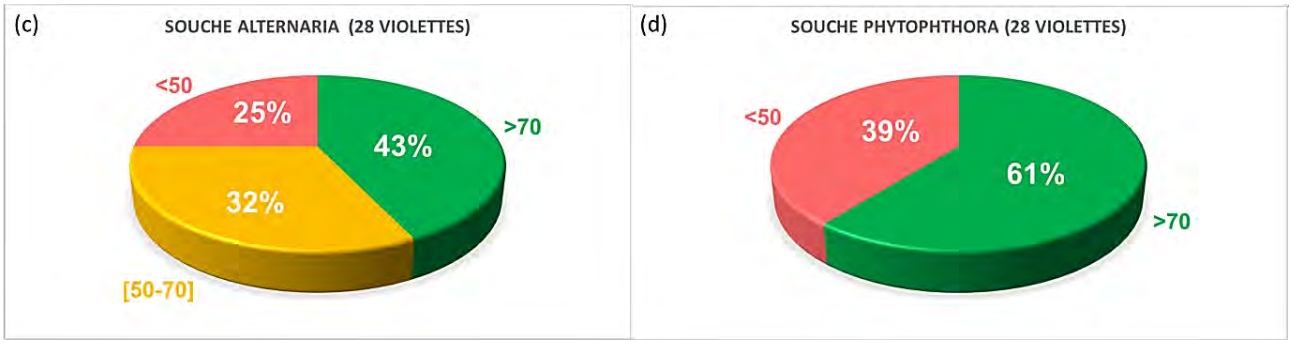


Figure 77 : Répartition des inhibitions mesurées pour *Botrytis* (a), *Fusarium* (b), *Alternaria* (c) et *Phytophthora* (d)

Ces résultats sont plutôt intéressants puisque ce sont des souches pathogènes pour la violette. Au final, une vingtaine de violettes a montré une activité avec une inhibition supérieure à 70% chez une des souches étudiées, mis à part *Colletotrichum* (Tableau 27). Ces extraits complexes obtenus par une extraction avec du méthanol 50% présentent donc un fort potentiel antifongique par rapport aux résultats du témoin positif.

Tableau 27. Résultats des activités phytosanitaires sur les quatre souches*

N°	Nom	Espèce	% inhibition *Botrytis*	% inhibition *Fusarium*	% inhibition *Alternaria*	% inhibition *Phytophthora*
V19	Viola Comte de Brazza	*V. alba*	98 ± 0	-	-	-
V36	Ash Vale Blue	*V. alba*	-	-	98 ± 26	-
V110	Violette de Toulouse	*V. alba*	97 ± 0	-	-	-
V1	Doreen	*V. odorata*	98 ± 0	-	-	-
V2	Alice	*V. odorata*	99 ± 0	-	125 ± 7	98 ± 2
V3	Viola Baronne de Rothschild	*V. odorata*	97 ± 0	-	122 ± 7	92 ± 8
V6	Charles Winston Groves	*V. odorata*	99 ± 0	-	117 ± 11	99 ± 3
V12	Annie	*V. odorata*	99 ± 0	-	-	-
V27	Crepuscule	*V. odorata*	80 ± 8	-	118 ± 9	96 ± 3
V33	Viola Nora Church	*V. odorata*	98 ± 0	-	-	-
V102	Coven garden	*V. odorata*	-	-	-	101 ± 0
V30	Viola sororia dark freckles	*V. sp* Hearn	-	100 ± 0	-	-
V31	Sororia blanche	*V. sp* Hearn	115 ± 10	97 ± 9	124 ± 3	96 ± 5
V83	Viola palmata	*V. sp* Hearn	-	101 ± 1	127 ± 14	98 ± 4
V106	Odorata blanche	*V. sp* Hearn	-	100 ± 0	135 ± 0	95 ± 2
V68	Viola San Remo	*V. suavis*	-	-	-	97 ± 2
V79	Elisabeth Bailes	*V. suavis*	-	-	121 ± 5	-
V100	Viola Amiral Avellan	*V. suavis*	-	-	-	89 ± 3
V34	Viola Green Jade	ND**	-	101 ± 1	118 ± 2	95 ± 5
V85	Viola striatta	ND	-	-	123 ± 1	96 ± 3
V101	Viola Hederacea epertion	ND	98 ± 11	85 ± 4	131 ± 1	73 ± 7
V111	Lady Hume Cambell	ND	79 ± 14	-	-	-

* *Colletotrichum* insensible aux extraits de violette
**ND : non déterminé

L'ensemble des espèces identifiées par analyse ITS est représenté, excepté les violettes appartenant à *V. subs Rostratae* qui ne semblent présenter aucune activité phytosanitaire sur les souches étudiées.

Suite au criblage de la collection sur *Botrytis* et *Fusarium*, il apparait que des violettes appartenant à *V. odorata* et *V. sp* Hearn ont la meilleure activité antifongique.
Seulement deux *V. alba* sur cinq et une *V. sp* Hearn sur vingt ont révélé une activité sur *Botrytis*, ne permettant donc pas de conclure d'une réelle relation espèce-activité. En revanche, concernant la souche *Fusarium*, un tiers des *V. sp* Hearn de la collection a présenté une activité supérieure à 50% et seulement des violettes de cette espèce se sont révélées actives. Ainsi, une certaine spécificité des *V. sp* Hearn à la souche *Fusarium* semble être mise en évidence, avec cependant un tiers des violettes actives restant non identifiées en génétique ou par prédiction avec les données UHPLC-HRMS. Une totale spécificité ne peut donc être avancée, mais cette souche est visiblement très sensible à cette espèce. Les *V. suavis* n'ont présenté aucune activité sur ces deux souches (Figure 78 a et b).

Les deux autres souches testées sur un échantillonnage réduit de la collection, comprenant les violettes actives sur *Fusarium* et *Botrytis* ainsi que d'autres n'ayant pas présenté de résultats concluants, se sont révélées plus sensibles aux extraits de violettes, mais toutefois sans réelle spécificité (Figure 78 c et d). En effet, respectivement 12 et 17 violettes sur les 28 analysées ont présenté une activité sur ces deux souches contre seulement 12 et 6 pour *Botrytis* et *Fusarium* sur l'ensemble de la collection. On peut donc supposer que ces nombres seraient augmentés avec le criblage de l'ensemble de la collection. À ce stade, l'échantillonnage des espèces n'étant pas représentatif de l'ensemble de la collection, une relation espèce/activité ne peut pas être plus profondément étudiée. Le principal résultat qui ressort de cette deuxième étude reste la sensibilité plus importante de ces deux souches à l'ensemble des violettes.

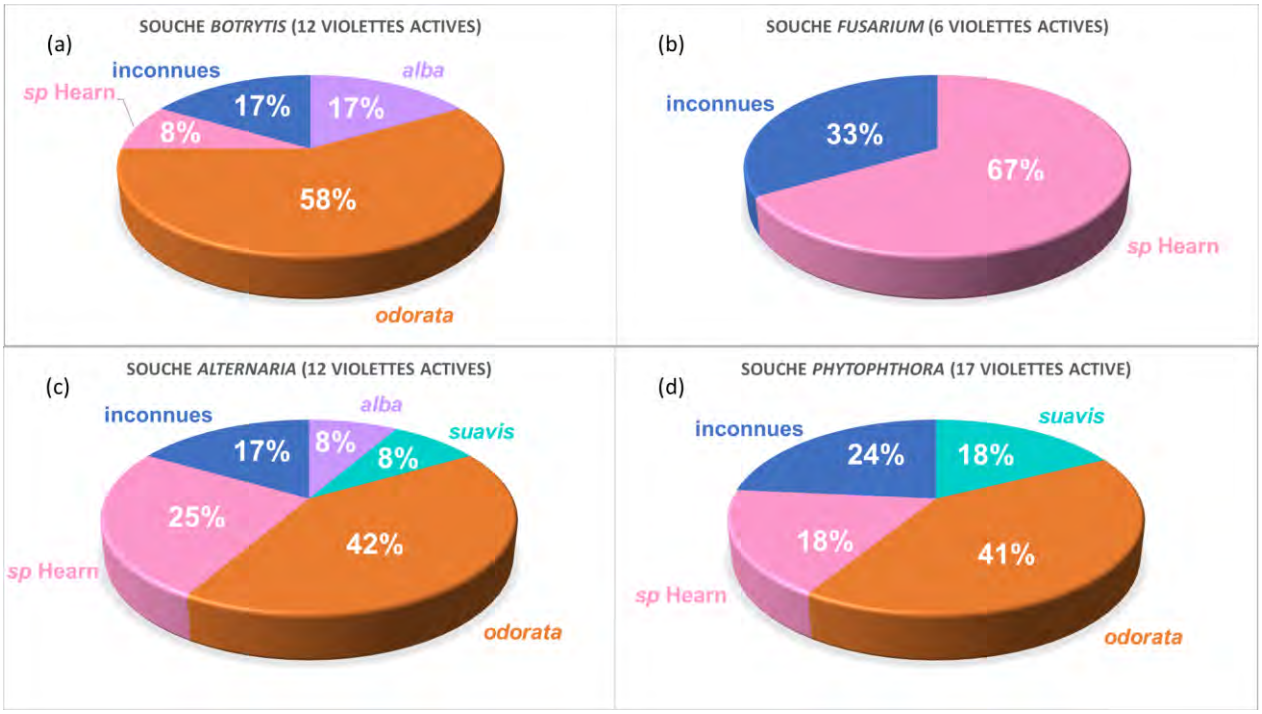


Figure 78. Relation espèce/activité pour les souches *Botrytis* (a), *Fusarium* (b), *Alternaria* (c) et *Phytophthora* (d)

Pour conclure sur l'activité antifongique des extraits de violettes, les *V. odorata* sont majoritairement représentées et constituent globalement un tiers des violettes actives. La deuxième espèce d'intérêt est *V.sp* Hearn représentant près de 20% des violettes actives (Figure 79). Quelques violettes *V. alba* et *V. suavis* ont été mises en relief, mais des études complémentaires sont nécessaires pour tirer des conclusions plus pertinentes sur la présence d'une relation entre les espèces et leur activité sur les souches étudiées. Les premiers résultats ont mis l'accent sur la sensibilité de la souche *Fusarium* aux extraits issus de *V. sp* Hearn, et de *Botrytis* à ceux issus de *V. odorata*.

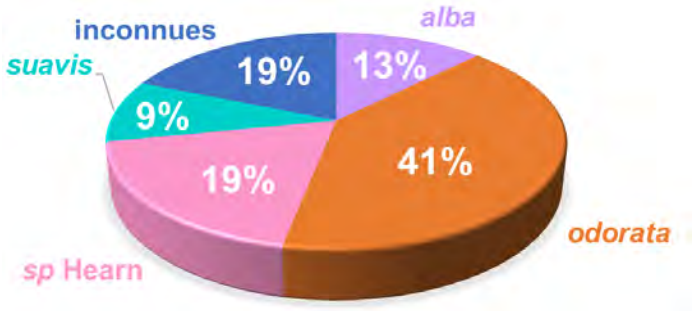


Figure 79. Répartition des espèces actives sur au moins l'une des souches étudiées

II.4.3.1.2 Analyse statistique des données antifongiques

II.4.3.1.2.1 ANOVA

De même que pour le test DPPH, une analyse ANOVA des résultats antifongiques a été réalisée afin de déterminer s'il existe une différence significative entre les différentes espèces et de mettre en évidence une potentielle relation espèce/activité pour l'ensemble des souches étudiées. Les résultats de normalisation et de l'analyse ANOVA sont résumés Figure 80 et Tableau 28.

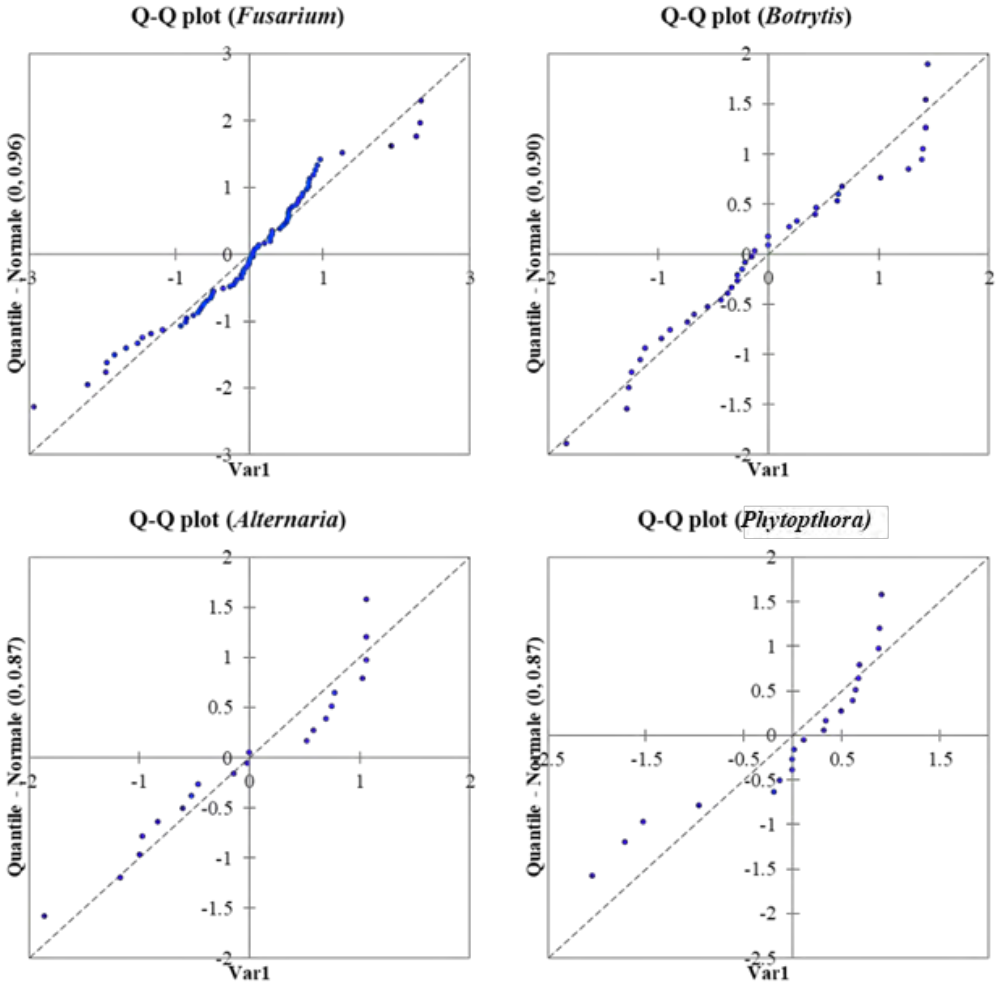


Figure 80. Graphique Q-Q de l'analyse de normalité des résidus centrés réduits des données antifongiques

Tableau 28. Résultats de l'analyse ANOVA pour chaque souche

Souche	Obs.	Obs -log	*p*-value Shapiro	p-value Jarque-Bera	R^2	Test de Fischer Pr > F	Test de Tukey
Fusarium	88	80	0,015	0,119	0,116	9,9%	Aucune différence
Phytophthora	22	20	0,004	0,119	0,185	51,4%	Aucune différence
Botrytis	90	38	0,093	0,507	0,115	67,1%	Aucune différence
Alternaria	22	20	0,110	0,529	0,060	91,1%	Aucune différence

Deux tests de normalisation ont été appliqués et leurs conclusions sont divergentes puisque pour le test de Jarque-Bera, toutes les données semblent normalisées avec un risque de 50% de rejeter l'hypothèse alors qu'elle est vraie pour *Botrytis* et *Alternaria* et de 12% pour *Fusarium* et *Phytophthora* ; tandis que pour Shapiro, seuls les résultats pour *Botrytis* et *Alternaria* concluent à des données normalisées avec un risque de rejeter H0 alors qu'elle est vraie d'environ 10%. Le test de Shapiro étant très puissant, il fournit généralement des valeurs de *p* inférieures aux autres tests. Lorsque l'échantillonnage est grand, par exemple dans ce cas presqu'une centaine d'extraits pour *Fusarium*, il est de rigueur de prendre en compte le graphique Q-Q pour affiner les conclusions. Ainsi, en combinant les résultats de ces graphes et les *p*-value de Shapiro, on peut tout de même considérer la normalité de l'ensemble des données. En effet, pour *Fusarium*, l'éloignement de la bissectrice reste limité et la tendance est globalement respectée tout comme pour *Botrytis* et *Alternaria* dont la normalité est confirmée par les tests numériques. Il en est de même pour *Phytophthora* dont le graphique Q-Q est sensiblement similaire à celui d'*Alternaria*. Sur ces considérations, l'ANOVA a donc pu être interprétée en détail et la valeur du R^2 est inférieure à 20% pour l'ensemble des souches, allant de 6% pour *Alternaria* à 18% pour *Phytophthora*, signifiant que moins du quart de la variabilité semble ici expliquée par l'espèce. Ainsi la seule appartenance à une espèce n'explique que faiblement les différences biologiques mesurées. Pour le test de Fisher, le seuil de signification (0,05) est largement dépassé pour l'ensemble des souches, excepté pour *Fusarium* avec une valeur de 10%, impliquant que l'hypothèse H0 (aucune différence entre les espèces) ne peut être rejetée et donc aucune relation espèce/activité ne peut être déterminée. L'ensemble des résultats peu significatifs issus de l'ANOVA pourraient résulter de la disparité des activités au sein des groupes. En effet, pour exemple, certaines *V. odorata* ont été révélées actives tandis que d'autres non. Ainsi, les moyennes des variables calculées pour chaque espèce (Figure 81) comprennent des intervalles parfois très étendus résultant d'écarts-types importants. Par conséquent, la différence entre les moyennes des espèces est clairement amenuisée n'assurant pas une bonne discrimination entre les groupes. De plus, le fait que plusieurs espèces ont réagi à une même souche n'arrange aucunement la mise en évidence d'une pertinente relation espèce/activité. En ce qui concerne *Botrytis*, le groupe *V. alba* semble se démarquer par rapport à *V. odorata*, ce qui ne corrèle pas avec la Figure 78a et ce résultat s'explique par le nombre des violettes de chaque ensemble. En effet, la collection ne comprenant que cinq *V. alba*, la moyenne des réponses reste élevée puisque deux violettes ont présenté une inhibition supérieure à 90%. Cependant, l'important écart-type indique que les autres violettes de cet ensemble n'ont présenté aucune activité.

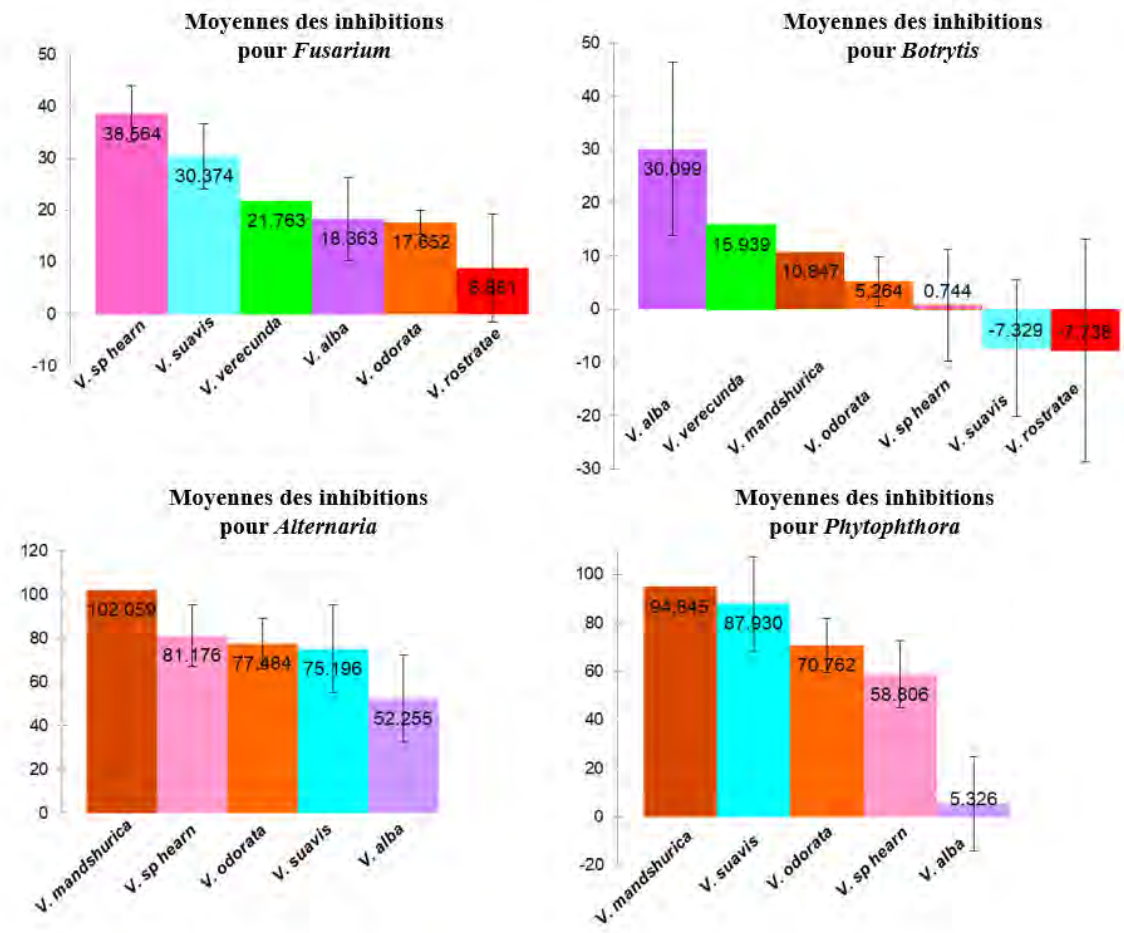


Figure 81. Moyenne des pourcentages d'inhibition avec standard de chaque espèce pour les quatre souches *Fusarium*, *Botrytis*, *Alternaria* et *Phytophtora*

La valeur nettement inférieure du résultat du test de Fisher pour *Fusarium* en comparaison aux autres souches peut être expliquée par le fait que ce sont principalement des violettes *V. sp.* Hearn qui ont été actives sur cette souche. Néanmoins, l'intégralité des violettes de cette espèce n'ayant pas présenté la même réponse, les résultats de l'ANOVA ne permettent pas de conclure sur une différence par rapport aux autres. En creusant ces résultats, il s'est avéré que les trois violettes qui avaient montré une activité sur *Fusarium* semblent présenter quelques différences au niveau de leur profil chimique, comme en témoigne leur séparation des autres violettes sur l'ACP, Figure 82. Ainsi, il pourrait être supposé la présence de variétés différentes qui n'ont pas pu être mises en évidence lors des analyses génétiques sur les ITS.

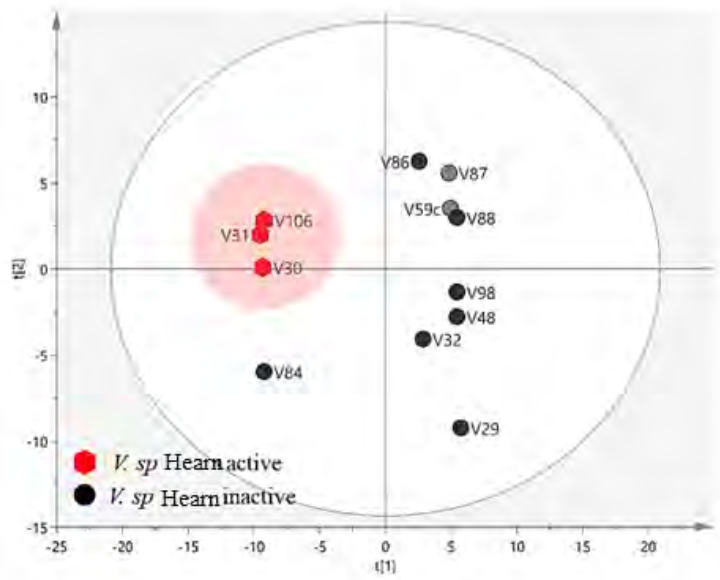


Figure 82. ACP des *V. sp* Hearn à partir des données UHPLC-HRMS et coloration des plants actifs sur *Fusarium*

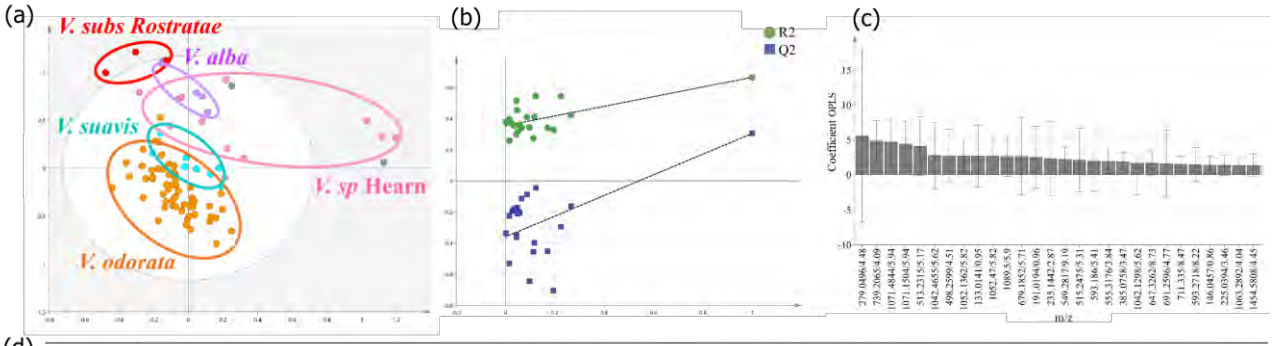
II.4.3.1.2.2 OPLS et identification des métabolites d'intérêt

Malgré les résultats peu significatifs du test de Tukey obtenus pour les différentes souches, ne permettant pas de mettre en évidence une relation espèce/activité antifongique, des modèles OPLS avec les données non loguées et l'échelle pareto ont été réalisés afin d'obtenir une classification globale de l'ensemble des métabolites, en se focalisant sur les composés majoritaires. Les résultats de l'OPLS sont présentés Tableau 29.

Tableau 29. Résultats des paramètres caractérisant le modèle OPLS des données antifongiques

Souche	R2X	Q2Y	Permutation plot	CV-ANOVA : *p*-value
Fusarium g.	0,363	0,308	Conforme	$2{,}78\text{x}10^{-5}$
Phytophthora c.	0,322	0,305	Juste conforme	$5{,}39\text{x}10^{-2}$
Botrytis c.	0,260	-0,330	Non conforme	1
Alternaria b.	0,300	-0,608	Non conforme	1

Les deux critères de validation de l'OPLS sont le test de permutation et la valeur de *p* de l'analyse CV-ANOVA. Pour toutes les souches, excepté *Fusarium*, le premier n'est pas conforme aux exigences et la valeur de *p* supérieure au seuil de validation égal à 0,05. Par conséquent, l'analyse des coefficients pour ces trois souches ne pourra être effectuée puisque les modèles ne sont pas valides et donc non prédictifs. En revanche, cela est possible pour la souche *Fusarium* puisque le modèle (Figure 83 a), caractérisé par un R2X de 0,363 et un Q2Y de 0,308, présente un test de permutation conforme (Figure 83 b) et une valeur de *p* égale à $2{,}78.10^{-5} < 0{,}05$. Le modèle est donc de qualité suffisante pour pouvoir procéder à l'interprétation des coefficients (Figure 83 c) et mettre en évidence des composés d'intérêt qui seraient responsables de l'activité antifongique sur la souche *Fusarium* (Figure 83 d).



(d)

N°	Score coefficient	[M-H]⁻	[M+2H]²⁺	δm/z (ppm)	[M+3H]³⁺	δm/z (ppm)	RT (min)	Formule brute	Classe chimique	Identification
17	4,89	739,2065			-	3,6525	4,09	$C_{33}H_{40}O_{19}$	Flavonoïde	Kaempferol 3,7-diglycosides; 3-O-[β-L-Rhamnopyranosyl-(1→6)-β-D-glucopyranoside], 7-O-α-L-rhamnopyranoside
26	4,69	-	1605,7244	0,9342	1070,8164	1,4008	5,94	$C_{138}H_{216}N_{36}O_{40}S_6$	Cyclotide	Cycloviolacin O11
14	4,09	-	1570,1869	3,0570	1047,1292	1,2415	5,71	$C_{133}H_{207}N_{37}O_{39}S_6$	Cyclotide	Cycloviolacin O9
27	2,77	-	1562,1857	5,5051	1041,7974	1,3438	5,62	$C_{133}H_{207}N_{37}O_{38}S_6$	Cyclotide	Cycloviolacin O13
28	2,69	-	1577,1978	1,1423	1051,7999	2,3769	5,82	$C_{134}H_{209}N_{37}O_{39}S_6$	Cyclotide	Cycloviolacin O4

Figure 83. Modèle OPLS (a) validé par le test de permutation (b) et analyse des coefficients de la régression linéaire (c) avec emphase de 5 composés d'intérêt majoritaires chez *V. sp* Hearn (d)

Les composés présentant un écart-type trop important ont été écartés de l'analyse des coefficients OPLS et ce sont cinq composés majoritaires chez *V. sp* Hearn qui ont été principalement mis en évidence ; parmi lesquels quatre cyclotides, en particulier des cycloviolacines et un flavonoïde. Un cinquième cyclotide de masse m/z = 1089,168 correspondant au cyclotide Cter D (12) de formule brute $C_{142}H_{221}N_{35}O_{41}S_6$, est retrouvé chez *V. sp* Hearn, mais également chez *V. suavis* (inhibition autour de 50%) et *V. odorata* (inhibition autour de 30%). Les cyclotides sont essentiellement détectables en mode d'ionisation positif et ont été identifiés par la forme caractéristique du profil isotopique, ainsi que la masse exacte des fragments majoritaires $[M+2H]^{2+}$ et $[M+3H]^{3+}$. Les incertitudes de masses étant toutes inférieures ou proches de 5 ppm, les identifications obtenues restent des hypothèses solides. Les spectres de masse correspondants sont présentés en annexe 7, page 227 avec un tableau des caractéristiques des cinq cyclotides cités.

Les résultats du modèle OPLS ont conduit à la mise en évidence du fort potentiel actif des cyclotides sur la souche *Fusarium*. Ces résultats sont en adéquation avec la littérature puisque des études ont mis en évidence les potentiels antifongique et antimicrobien des cyclotides sur diverses souches de champignons et de bactéries (**Ribeiro et al., 2013**). Bien que seule cette souche ait pu être étudiée plus en détails, il semble cohérent de penser que des composés similaires soient à l'origine des résultats biologiques sur les autres souches ; d'autant plus que les *V. odorata*, dont les composés spécifiques avaient précédemment été identifiés parmi les cyclotides, sont clairement représentées au niveau des espèces de violettes ayant révélé une activité.

II.4.3.1.3 Vérification de l'activité des cyclotides

II.4.3.1.3.1 Extraction de *V. odorata* et *V. sp* Hearn

Une des violettes identifiées chez *V. odorata* (V6) et deux chez *V. sp* Hearn (V30 et V31) ont été étudiées plus en détail par Mr Khatib, chercheur-ingénieur au sein du LRSV, afin de valider l'hypothèse portant sur l'activité des cyclotides. Pour cela elles ont subi la même extraction que pour le criblage de la collection, avec une solution de MeOH 50%. La solution résultante étant très visqueuse, il a fallu ajouter deux fois plus de solvant atteignant ainsi un rapport solvant/matière de 20, au lieu de 10

d'ordinaire et les solutions ont dû subir une centrifugation pour récupération du surnageant, la filtration papier étant impossible.
Ces violettes appartenant à la collection nationale, la disponibilité de la matière est réduite et seulement quelques feuilles ont pu être récoltées conduisant au traitement de l'ordre de 1 à 2 g de matière lyophilisée. Les rendements d'extraction ont été de l'ordre de 30% en masse de matière sèche.

II.4.3.1.3.2 Fractionnement SPE des extraits bruts

Les trois extraits bruts ont été fractionnés par SPE sur cartouche C_{18} de 1 g (cf. IV.5.2.2.6 page 182) pour confirmer l'hypothèse d'activité portant sur les cyclotides. Pour tous les échantillons excepté V31, 100 mg d'extrait brut obtenu par l'extraction avec du méthanol 50% ont été dissous dans 600 µL de tampon NH_4HCO_3 et introduits dans la cartouche ; contre 155 mg en trois essais pour V31 en raison d'une saturation importante de la colonne. Les cartouches avaient été préalablement activées avec du méthanol et équilibrées avec du tampon. Le bilan de masse révèle un rendement moyen de l'ordre de 40% avec une disparité au niveau des poids de chaque fraction. La fraction Fr1 reste la plus abondante pour chacun des échantillons reflétant la présence d'abondants composés polaires tandis que la fraction Fr4 est la plus minoritaire. La fraction Fr3, supposée la plus riche en cyclotides, représente environ 13% de l'extrait brut chez *V. odorata*, contre moins de 5% pour *V. sp* Hearn. La première espèce semble donc plus particulièrement riche en ces composés. Ces résultats restent en accord avec de précédentes études ayant mis en évidence les cyclotides en tant que biomarqueurs prépondérants des *V. odorata.*

II.4.3.1.3.3 Profils UHPLC/HRMS

Une ACP a été réalisée sur les 18 extraits analysés (3 extraits de chaque fraction SPE, 3 extraits bruts, 3 QC) afin de rendre compte de l'organisation générale des données et de leur disparité. On a pu observer que les fractions SPE formaient chacune un groupe homogène justifiant ainsi la séparation de composés chimiques similaires que ce soit avec les violettes identifiées chez *V. sp* Hearn, que chez celle identifiée chez *V. odorata* (Figure 84 a). L'analyse des profils UHPLC-MS a permis de constater que la fraction Fr2 était enrichie en flavonoïdes, dont les composés 5, 21 et 25 identifiés dans les études précédentes. La fraction Fr3 semble être enrichie en cyclotides également détectés dans la fraction Fr4 en mélange avec des composés plus apolaires (Figure 84 b). Les identifications des cyclotides sont présentées dans le Tableau 30 et les spectres de masse correspondants peuvent être retrouvés en annexe 8, page 230. Une majorité de cycloviolacines a été constatée dont trois restent inconnues.

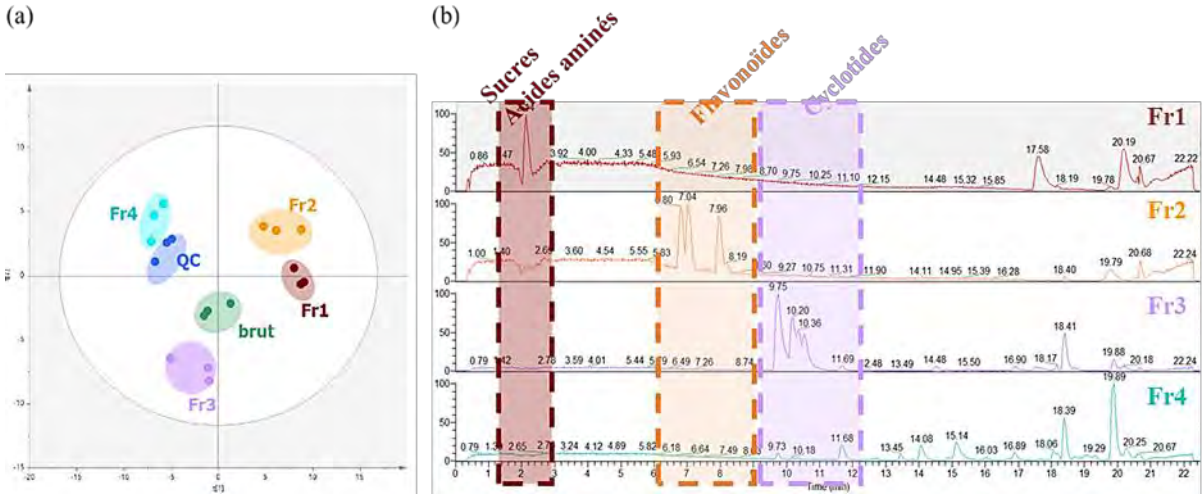


Figure 84. ACP des données UHPLC-HRMS en mode PI des fractions SPE (a) et profils MS associés (b)

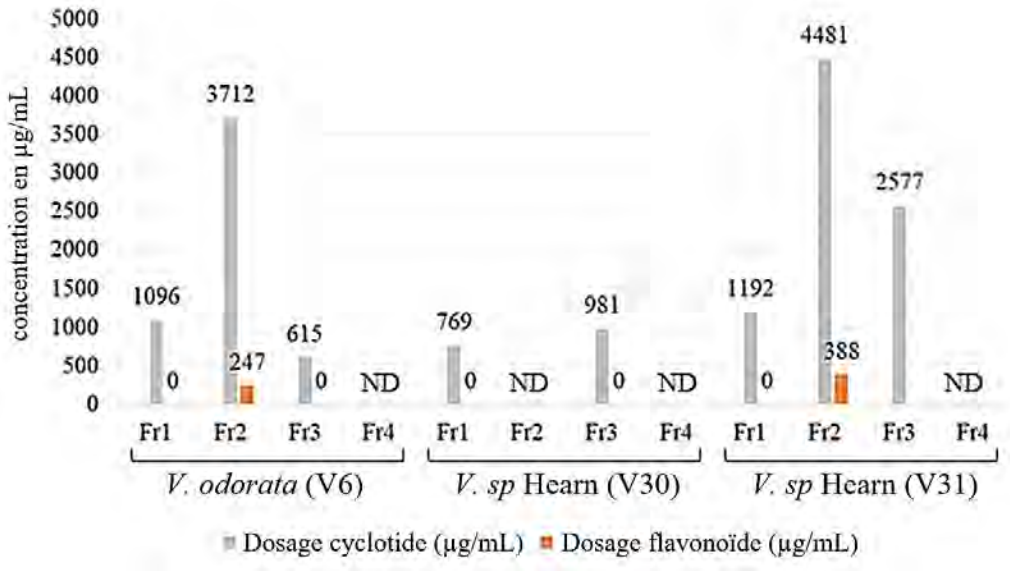
Tableau 30 Identités des flavonoïdes et des cyclotides déterminés dans les fractions Fr2, Fr3 et Fr4

N°	m/z	RT (min)	Formule brute	Fraction	Identification*
5	$[M+H]^+$ 595,1646	7,34	$C_{27}H_{30}O_{15}$	Fr2	6-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxy-8-*C*-α-L-rhamnopyranosylflavone
21	$[M+H]^+$ 565,1536	7,70	$C_{26}H_{28}O_{14}$	Fr2	8-*C*-β-D-Glucopyranosyl-4',5,7-trihydroxyflavone; 6"-*O*-α-L-Arabinopyranosyl
25	$[M+H]^+$ 579,1696	7,96	$C_{27}H_{30}O_{14}$	Fr2	8-*C*-β-D-Glucopyranosyl-4',5,7-trihydroxy-6-*C*- α-L-rhamnopyranosylflavone
29	$[M+2H]^{2+}$ 1448,0735 $[M+3H]^{3+}$ 965,7178	9,4	ND	Fr3	ND
30	$[M+2H]^{2+}$ 1455,0807 $[M+3H]^{3+}$ 970,3887	9,4/9,85	ND	Fr3	ND
31	$[M+2H]^{2+}$ 1544,1776 $[M+3H]^{3+}$ 1029,7891	9,40	$C_{132}H_{208}N_{34}O_{39}S_6$	Fr4	Cycloviolacine O18
32	$[M+2H]^{2+}$ 1625,2166 $[M+3H]^{3+}$ 1083,8062	9,40	ND	Fr4	ND
33	$[M+2H]^{2+}$ 1462,0730 $[M+3H]^{3+}$ 975,0496	9,75	$C_{116}H_{175}N_{35}O_{42}S_6$	Fr3	Vigno 2
34	$[M+2H]^{2+}$ 1613,7262 $[M+3H]^{3+}$ 1076,1498	10,35	$C_{138}H_{216}N_{36}O_{41}S_6$	Fr4	Cycloviolacine O8
35	$[M+2H]^{2+}$ 1446,0732 $[M+3H]^{3+}$ 964,3868	10,45	$C_{117}H_{179}N_{35}O_{39}S_6$	Fr3	Cycloviolacine O12
36	$M+2H]^{2+}$ 1439,0686 $[M+3H]^{3+}$ 959,7144	10,50	$C_{116}H_{177}N_{35}O_{39}S_6$	Fr3	Vaby peptide B, C, D
37	$[M+2H]^{2+}$ 1485,5924 $[M+3H]^{3+}$ 990,7306	10,84	$C_{123}H_{184}N_{34}O_{40}S_6$	Fr3	Cycloviolacine O21
38	$[M+2H]^{2+}$ 1565,6754 $[M+3H]^{3+}$ 1044,1187	11,06/11,69	$C_{132}H_{208}N_{36}O_{40}S_6$	Fr3	Cycloviolacine H1
39	$[M+2H]^{2+}$ 1573,1740 $[M+3H]^{3+}$ 1049,1198	11,47	$C_{135}H_{205}N_{37}O_{38}S_6$	Fr3	Cyclovioline D
40	$[M+2H]^{2+}$ 1575,6840 $[M+3H]^{3+}$ 1050,7855	11,67	$C_{134}H_{208}N_{38}O_{38}S_6$	Fr3	Cycloviolacine O17

* ND = non déterminé

II.4.3.1.3.4 Dosage des flavonoïdes et des cyclotides

Les différentes fractions concentrées à 10 mg/mL dans le méthanol 50% ont été dosées, dans la mesure du possible, en teneurs en flavonoïdes par la méthode $AlCl_3$ avec comme référence la rutine, et en cyclotides par la méthode de l'acide bicinchoninique (BCA) (cf. IV.5.2.6.5, page 190) avec comme référence la protéine BSA (albumine de sérum bovin) (Figure 85).



* ND : non déterminé

Figure 85. Résultats des dosages BCA et flavonoïdes des fractions des extraits cyclotides

On constate que le dosage BCA donne des résultats positifs sur les trois premières fractions. D'après les profils MS (Figure 84 b), il semble que pour la première fraction ce soient essentiellement les sucres et acides aminés qui sont dosés, comme en témoigne l'important pic d'injection à 2 min. Pour la fraction Fr3, le dosage semble bien porter sur les cyclotides, détectés en HRMS. La fraction Fr2 semble dépourvue de cyclotides mais présente tout de même une forte réponse au dosage BCA. Malgré l'absence de pic d'injection en masse, ce résultat pourrait provenir d'une teneur en acides aminés élevée. La fraction Fr4 n'a pas pu être dosée en raison d'un volume de solution limitée qui a été dédié aux tests antifongiques.

Pour le dosage des flavonoïdes, on constate que globalement les fractions Fr1, Fr3 et Fr4 sont dépourvues de flavonoïdes, essentiellement contenus dans la fraction Fr2. Pour la violette V31, une teneur dans la fraction Fr3 a été déterminée, mais la comparaison avec le signal quantitatif du détecteur Corona ne permet pas de valider la présence de flavonoïdes dans cette fraction, en raison de l'absence de pics quantitatifs dans la région de temps de rétention de ces composés. Une erreur de dosage est probablement survenue et il a été décidé de ne pas tenir compte de la valeur obtenue.

II.4.3.1.3.5 Activité antifongique

Les différents extraits de violette ont été concentrés à 10 mg/mL dans le méthanol 50% et testés sur deux souches de champignons précédemment criblés, à savoir *Botrytis* et *Fusarium*, puisque ce sont les deux souches qui ont été particulièrement sensibles aux extraits de *V. odorata* pour la première, et *V. sp* Hearn pour la seconde. Les mêmes conditions d'analyse que lors du criblage de la collection ont été respectées.

- Résultat antifongique sur la souche *Botrytis*

Il n'a été observé aucune influence du méthanol sur la croissance du champignon, mais une bonne réponse de l'antifongique à J5, c'est-à-dire cinq jours après contact avec les échantillons En effet, il est constaté un état similaire du milieu de couleur gris clair avec une pigmentation noire pour les puits du blanc (H_2O) et du méthanol contre un milieu incolore pour l'antifongique, révélant la destruction totale du champignon (Figure 86 a).

Pour les extraits de violette, il est constaté que les fractions Fr1 et Fr2 contribuent à la croissance du champignon pour les trois violettes. Ce fort développement est visuellement constaté par la pigmentation marron foncé à clair des puits relatifs à ces fractions. Ce résultat pourrait être expliqué tout d'abord par la présence de sucres et probablement de mucilages en quantité non négligeable dans les extraits, aux vus de leur viscosité. Ces substances sont des sources importantes de carbone contribuant à la croissance des champignons.

En revanche, une inhibition est remarquée pour les fractions Fr3 et Fr4 de *V. odorata* avec des moyennes respectives de l'ordre de 15 et 50%. Il est visuellement difficile d'en juger pour Fr3, mais un milieu plus clair est clairement noté dû à une pigmentation noire moins prononcée pour Fr4 (Figure 86 a). Aucune inhibition significative n'a été mise en évidence pour les fractions Fr3 des *V. sp* Hearn, mais une légère activité a été décelée pour leurs fractions Fr4 avec des inhibitions de l'ordre de 10%.

Ces réponses indiquant un potentiel antifongique des fractions Fr3 et Fr4 pourraient ainsi être mises en relation avec la présence de cyclotides, identifiés sur les profils UHPLC-HRMS. Par ailleurs, le potentiel antifongique plus prononcé chez *V. odorata* pourrait être à mettre en relation avec une certaine spécificité de ses cyclotides, par rapport à ceux contenus chez les *V. sp* Hearn. Cependant, ces hypothèses devront être vérifiées par l'isolement de ces composés d'autant plus que la calibration des extraits ici testés était portée sur la masse d'extrait brut et non la quantité en cyclotides qui a également une forte influence.

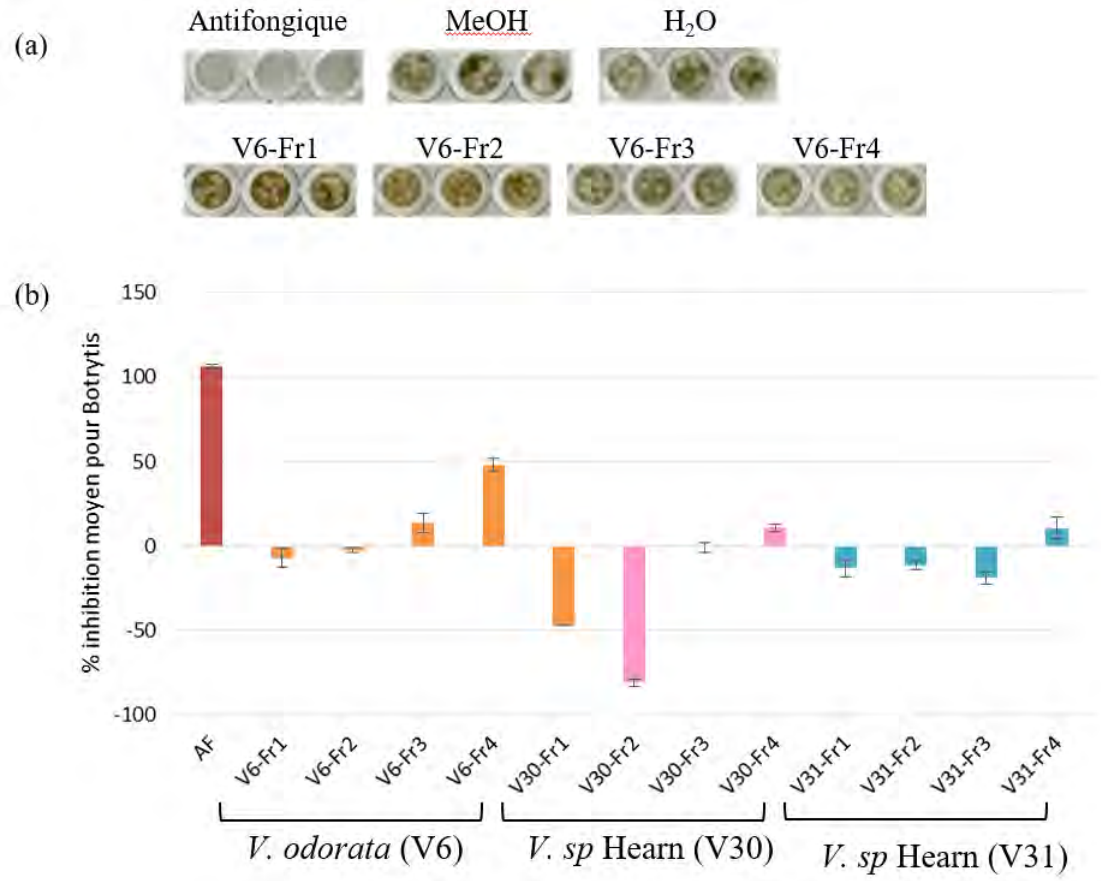


Figure 86. Photos des tests antifongiques des extraits sur *Botrytis* (a) et moyennes des pourcentages d'inhibition avec déviation standard de l'antifongique (AF) et des différents extraits de violette (b)

Les résultats de cette expérience ont permis de mettre en évidence les activités des fractions Fr3 et Fr4 de *V. odorata*. Les deux autres violettes appartenant à *V. sp* Hearn ont présenté une légère activité qui

reste tout de même moins prépondérante. Cependant, la présence manifeste de mucilages dans les extraits méthanoliques et les fractions, suspectée par leur fort aspect visqueux, a pu entacher les résultats du test. En effet, les mucilages sont des sources de carbone participant au développement du champignon. Si la balance source de carbone/composés antifongiques est positive, l'inhibition de la croissance ne sera pas significative. C'est ce qui est suspecté pour les fractions Fr3 supposées riches en cyclotides d'après les profils UHPLC-HRMS et les dosages. Ainsi, il ressort que *V. odorata* présente une activité antifongique sur *Botrytis*, en particulier les deux fractions enrichies en cyclotides. Pour *V. sp* Hearn, les conclusions sont moins évidentes et des tests complémentaires sont à envisager.

- Résultat antifongique sur la souche *Fusarium*

Il n'a été observé aucune influence du méthanol sur la croissance du champignon, mais une bonne réponse de l'antifongique. En effet, il est constaté la même couleur beige foncé au niveau des puits du blanc (H_2O) et du méthanol contre un milieu incolore pour l'antifongique, révélant la destruction totale du champignon (Figure 87 a). La fraction Fr1 est celle qui favorise le plus le développement du champignon, très certainement en raison de sa concentration en sucres, au même titre que les fractions Fr2, mais dans une moindre mesure. Cette forte croissance est visuellement constatée par la couleur marron clair des puits relatifs aux fractions Fr1 (Figure 87 a). Une inhibition de la croissance de *Fusarium* de l'ordre de 18%, est observée pour la fraction Fr4 de *V. odorata* et de la majorité des fractions Fr3 et Fr4 des *V. sp* Hearn de l'ordre de 10 à 20% (Figure 87 b). Cette inhibition est visuellement constatée par la couleur plus claire du milieu en comparaison du blanc, et la présence d'une légère pigmentation marron foncé aux extrémités (Figure 87 a). Ainsi, il semble que les fractions les plus riches en cyclotides soient les fractions les plus actives. Il est toutefois intéressant de constater que les fractions Fr3 de *V. odorata* et de la *V. sp* Hearn n°30 n'ont pas inhibé la croissance du champignon. Ce résultat pourrait ainsi évoquer la présence d'une spécificité de certains cyclotides.

Par ailleurs, la spécificité de la souche à *V. sp* Hearn évoquée lors du criblage de la collection n'est pas retrouvée ici. Tout comme évoqué précédemment, la présence de mucilages peut avoir altérée la réponse en raison de la balance sources de carbone/composés antifongiques. Par ailleurs, l'influence de la concentration en composés antifongiques peut expliquer la plus faible réponse des *V. sp* Hearn, la calibration des extraits étant portée sur la masse d'extrait brut et non la quantité en cyclotides. Des essais complémentaires sont à envisager pour approfondir les résultats. Néanmoins, le principal objectif de confirmer l'activité des cyclotides a été validé, puisque la majorité des fractions Fr3 et Fr4 enrichies en cyclotides ont présenté une inhibition de la croissance du champignon.

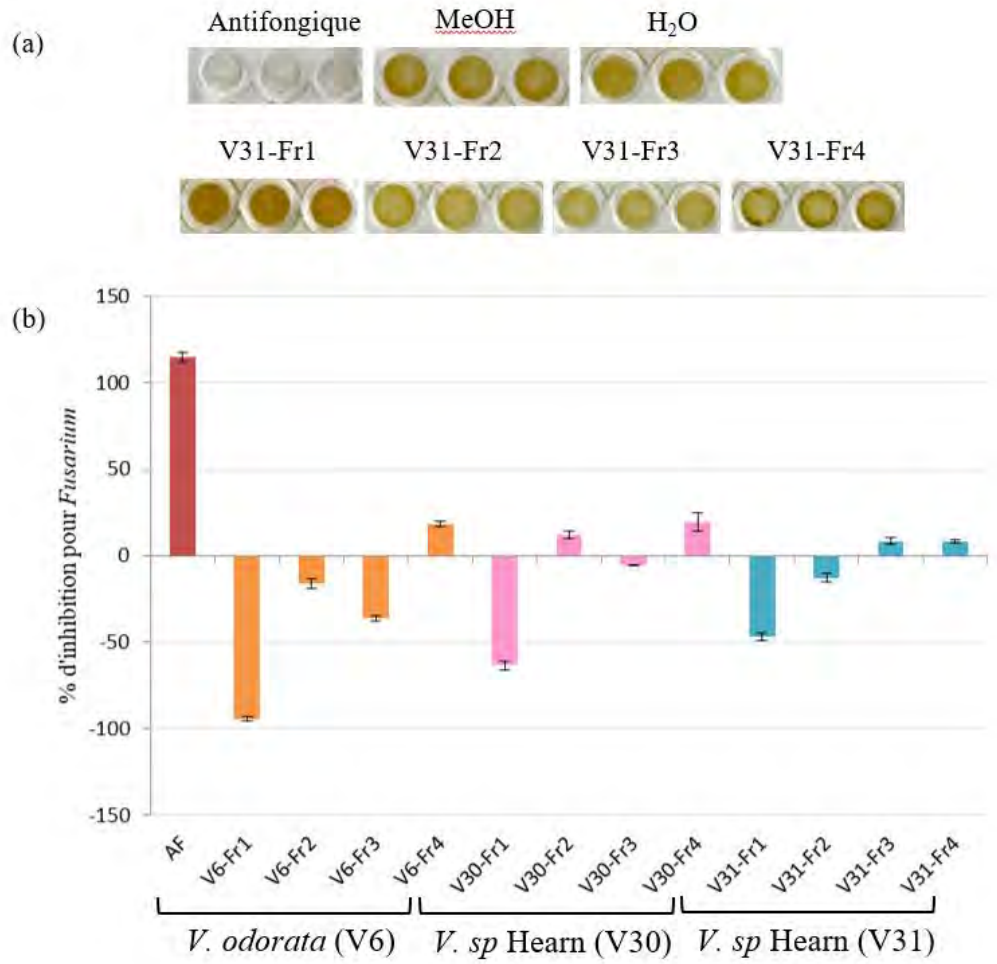


Figure 87. Photos des tests antifongiques des extraits sur *Fusarium* (a) et moyennes des pourcentages d'inhibition avec déviation standard de l'antifongique (AF) et des différents extraits de violette (b)

II.4.3.1.3.6 Conclusion

L'étude de l'activité antifongique approfondie, sur *Botrytis* et *Fusarium*, de deux violettes *V. sp* Hearn et d'une violette *V. odorata* a permis de mettre en évidence que les fractions Fr1 et Fr2, riches en sucres et acides aminés (dosage BCA positif), favorisaient globalement la croissance des champignons par l'apport de source de carbone tandis que les fractions Fr3 et Fr4 assuraient généralement une inhibition de la croissance. Or, il a été déterminé, principalement par UHPLC-HRMS et dans une moindre mesure par dosage, que ces fractions étaient enrichies en cyclotides. Ainsi, l'activité de ces extraits semble être en étroite relation avec leur concentration en cyclotides, comme supposé par les tests statistiques. Cependant, la non-calibration des extraits sur la concentration en cyclotides ne permet pas ici de conclure sur l'activité significative d'une espèce par rapport à l'autre ni même d'une fraction par rapport à l'autre, en particulier entre Fr3 et Fr4. Par ailleurs, la présence de mucilages dans les extraits, fortement suspectée par leur viscosité, pourrait perturber quelque peu la détermination des activités antifongiques des différentes fractions. C'est la présence d'eau dans le solvant d'extraction initial qui conduit à l'extraction de ces polysaccharides qui gonflent au contact de l'eau. Il serait donc intéressant de tester un autre protocole d'extraction ne faisant pas intervenir de solvant aqueux, afin de s'affranchir de ces composés et pouvoir estimer l'activité réelle des fractions SPE. Par ailleurs, les profils UHPLC des fractions Fr3 et Fr4 ont révélé la présence de cyclotides non identifiés par les bases de données. La purification et le séquençage de ces composés, par l'intermédiaire d'un clivage enzymatique afin

d'obtenir une conformation linéaire, permettraient de confirmer les identifications obtenues et de déterminer les séquences des cyclotides inconnus.

II.4.3.2 Criblage de l'activité élicitrice de défense des plantes

II.4.3.2.1 Résultat de l'activité GUS

La quasi-totalité de la collection a été analysée puisque 90 extraits de violette dans le méthanol 50% ont été testés pour cette activité de défense des plantes (cf. IV.6.3.2, page 197). Les plantules d'*Arabidopsis thaliana* génétiquement modifiées ont été utilisées car elles expriment le gène rapporteur iudA codant la β-glucuronidase sous le contrôle du promoteur de gène de défense PR1, bon marqueur des réponses de défense chez les plantes (**Shapiro et Zhang, 2001**). Le test a reposé sur l'observation de la coloration des plantules après ajout d'un substrat chromogénique de l'enzyme, à l'œil nu, pour un criblage haut débit (Figure 88).

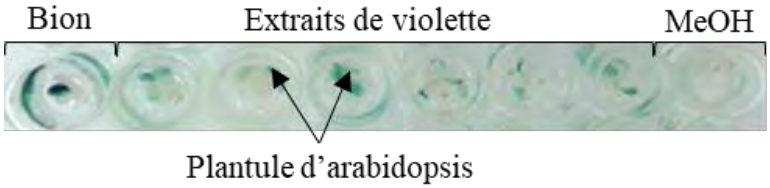


Figure 88. Révélation des plantules par la solution de coloration GUS de la solution de Bion, du méthanol et de quelques extraits de violette

Celle-ci a été évaluée sur une échelle de 0 à 4, 4 représentant la coloration la plus intense. Le témoin positif, une solution de Bion à 40 µg/ml, a présenté la coloration expérimentale la plus intense notée à 4. Les extraits de violettes ont ensuite été classés en trois catégories comme suit : [0-1,5], [1,5-2,5] et [2 ,5-4]. Environ la moitié des 90 extraits n'a présenté aucune activité élicitrice de défense, un tiers a présenté une activité modérée et seulement 14% soit 13 violettes, une coloration traduisant une activité pertinente (Figure 89 a et b). La lecture étant visuelle, les écarts-types sur certaines valeurs sont parfois importants et il a été décidé de sélectionner les extraits dont la moyenne était placée au-dessus de la valeur seuil fixée à 2,5 reflétant une activité pertinente.

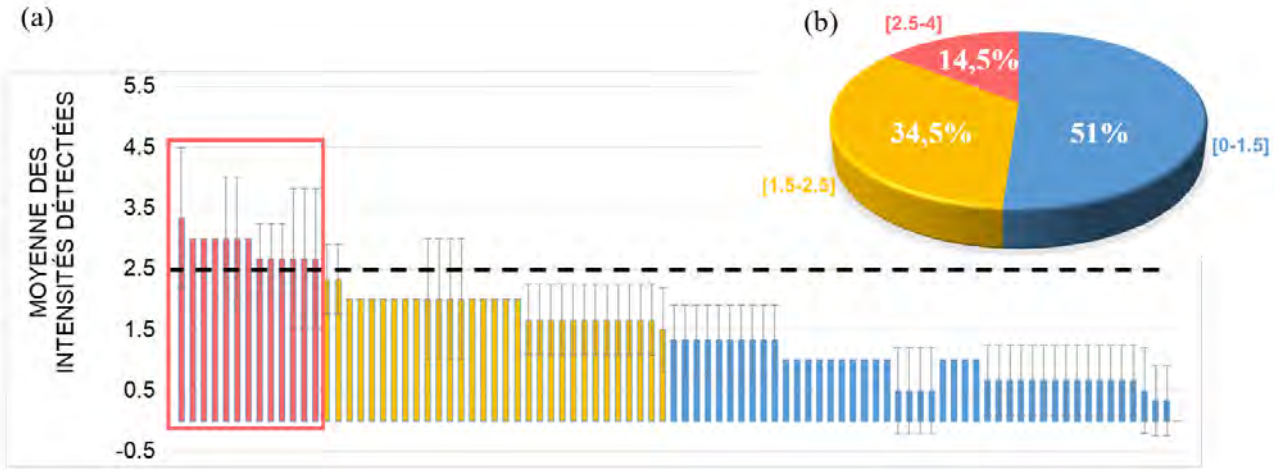


Figure 89. Courbe de tendance des moyennes des intensités mesurées (a) et répartition selon la coloration des plantules (b)

Les extraits de violettes ayant favorisé une réaction de défense chez la plante, traduite par la coloration bleu issue de l'expression du gène GUS, sont issus de deux principales espèces (Tableau 31) :

majoritairement *V. odorata* puisqu'elles représentent 38% des violettes actives et *V. sp* Hearn pour 23% d'entre elles. Une seule *V. alba* a révélé une activité, ce qui n'est pas représentatif pour assurer une activité pertinente de cette espèce (Figure 90). Par ailleurs, un tiers n'est pas identifié par manque de résultats en génétique et de prédiction par UHPLC-MS. Ainsi, l'établissement d'une relation espèce-activité semble compromis aux vus des résultats, mais il a été mis en évidence deux espèces avec un fort potentiel d'induction de mécanismes de défense chez la plante. Il est intéressant de noter que ces deux espèces sont celles largement mises en évidence pour l'activité antifongique.

Tableau 31. Identification des violettes d'intérêt dans la recherche d'activité de défense

N°	Nom vernaculaire	Espèce	Intensité de coloration
82	*Viola sororia* blue diamond	*V. sp* Hearn	3,3 ± 1,2
67	Toscane	*V. odorata*	3,0 ± 0,0
83	*Viola palmata*	*V. sp* Hearn	3,0 ± 0,0
101	*Viola Hederacea epertion*	ND*	3,0 ± 0,0
37	Hopley's white	*V. alba*	3,0 ± 1,0
72	Viola De Bruneau	ND	3,0 ± 1,0
89	Dawnie	*V. odorata*	3,0 ± 0,0
94	Zampieri	*V. odorata*	2,7 ± 0,6
60	Viola Valentré	*V. odorata*	2,7 ± 0,6
43	*Viola Eizanensis*	*V. sp* Hearn	2,7 ± 0,6
69	Viola Mrs Barton	*V. odorata*	2,7 ± 1,2
104	Kerry Girl	ND	2,7 ± 1,2
85	*Viola striatta*	ND	2,7 ± 1,2

* ND = non déterminé

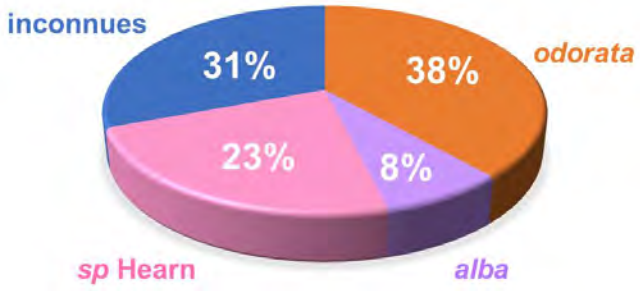


Figure 90. Relation espèce/activité pour l'étude de défense des plantes

II.4.3.2.2 Analyse statistique des données SDP

II.4.3.2.2.1 ANOVA

L'analyse ANOVA sur les données de stimulateur de défense des plantes (SDP) a conduit aux valeurs *p* respectives de 0,017 et 0,231 pour Shapiro et Jarque-Bera. Les deux résultats sont donc contradictoires ici puisque dans le premier cas, la valeur étant inférieure à 0,05, il faut rejeter l'hypothèse de départ stipulant que les données sont normalisées, à l'inverse du second résultat supérieur à la valeur seuil. En intégrant les résultats du graphique Q-Q (Figure 91), on peut finalement considérer la normalité des données puisque l'éloignement de la bissectrice reste limité et la tendance est globalement respectée.

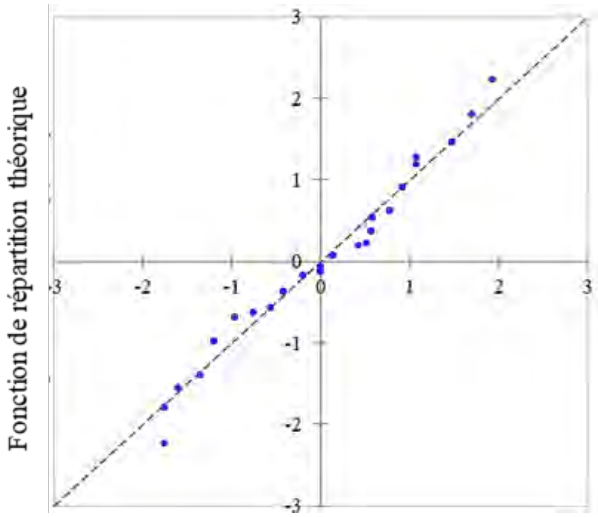


Figure 91. Graphique Q-Q des résidus centrés réduits des données logarithmées de coloration

L'interprétation des résultats de l'ANOVA a donc conduit à rendre compte que seulement 6,2% de la variabilité des résultats biologiques semble être expliquée par l'espèce (R^2 = 0 ,062). Ainsi, de même que pour les tests antifongiques, la seule appartenance à une espèce n'explique que très faiblement les différences biologiques mesurées. Le seuil de signification du test de Fisher (0,05) est clairement dépassé avec un risque de 61% de rejeter l'hypothèse H0 (aucune différence entre les espèces) alors que celle-ci est vraie. De façon évidente, le test de Tukey n'a pas permis de classifier les espèces, toutes considérées comme équivalentes en les intégrant dans un seul et même groupe. En effet, en examinant les moyennes de coloration enregistrées par espèces (Figure 92), aucune espèce ne se démarque réellement, mise à part *V. verecunda* dont le seul pot disponible ne permet pas de contrebalancer la valeur mesurée, tout comme pour *V. mandshurica*. À l'inverse, la collection comprenant de nombreux pots identifiés chez *V. odorata*, une moyenne autour de 1,4 est obtenue malgré l'émergence de certains plants avec des colorations de l'ordre de 3. Ces résultats pourraient être liés aux conclusions faites précédemment pour *V. sp* Hearn, à savoir que la présence de variétés différentes avec quelques différences chimiques pouvant expliquer ces disparités biologiques au sein de la même espèce.

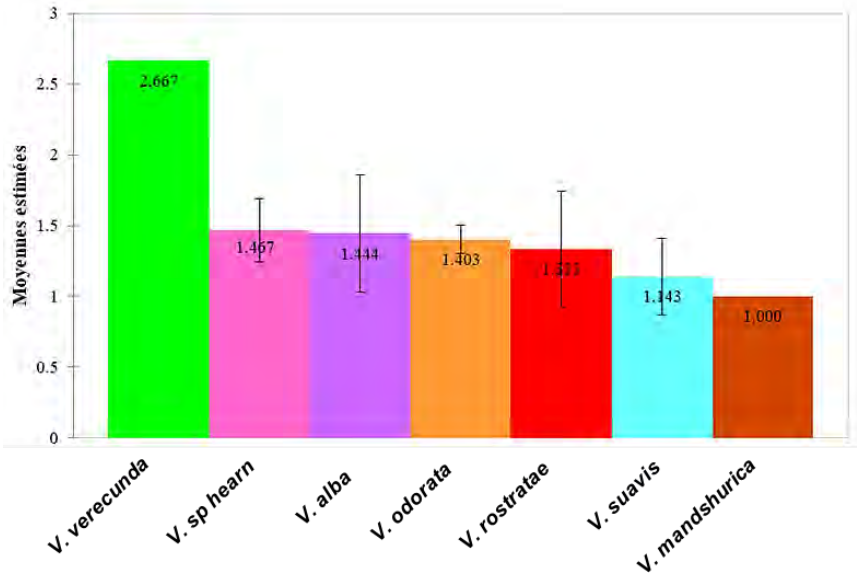


Figure 92. Moyennes des intensités de coloration avec intégration de l'erreur standard de chaque espèce

II.4.3.2.2.2 OPLS

Un modèle OPLS avec les données non logarithmées et l'échelle pareto a été construit en indiquant pour donnée Y les résultats d'intensité de coloration visuellement déterminés, toujours dans le même but d'analyser les coefficients de cette régression pour mettre en évidence des composés d'intérêt.

Le test de permutation (Figure 93) non conforme et la valeur *p* de l'analyse CV-ANOVA supérieure au seuil de validation de 0,05 ont conduit à conclure que le modèle OPLS ne peut être interprété puisqu'il ne respecte pas les prérequis pour assurer de bonnes conditions de prédiction. Par conséquent, l'identification de composés d'intérêt pouvant expliquer l'activité élicitrice de défense chez certains plants de violettes ne peut être proposée.

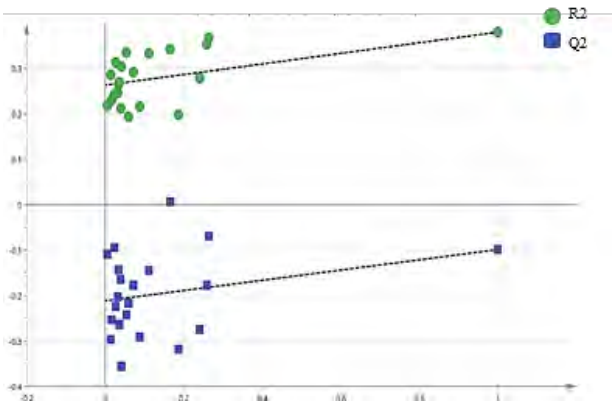


Figure 93. Test de permutation de l'OPLS des données non loguées en pareto des données SDP

II.4.3.3 Conclusion des activités phytosanitaires

Les études d'activités antifongique et de défense des plantes ont permis de mettre en évidence une trentaine de violettes avec un fort potentiel. Deux ensembles se sont particulièrement démarqués à savoir *V. odorata* et *V. sp* Hearn représentant respectivement 41 et 19% des violettes révélées actives, tandis que les *V. subs Rostratae* n'ont présenté aucune activité, tant au niveau antifongique que de défense (Figure 94). Les profils UHPLC ayant révélé que cet ensemble est particulièrement riche en flavonoïdes, mais semble pauvre en cyclotides au contraire des violettes ayant présenté une activité, en particulier *V. odorata* et *V. sp* Hearn, le postulat de départ basé sur le potentiel des cyclotides dans les activités phytosanitaires étudiées est justifié. Par ailleurs, la présence de cyclotides chez *V. suavis* mais leur faible représentation au niveau des espèces impliquées pourrait induire une certaine spécificité de quelques cyclotides pour les souches étudiées.

(a)

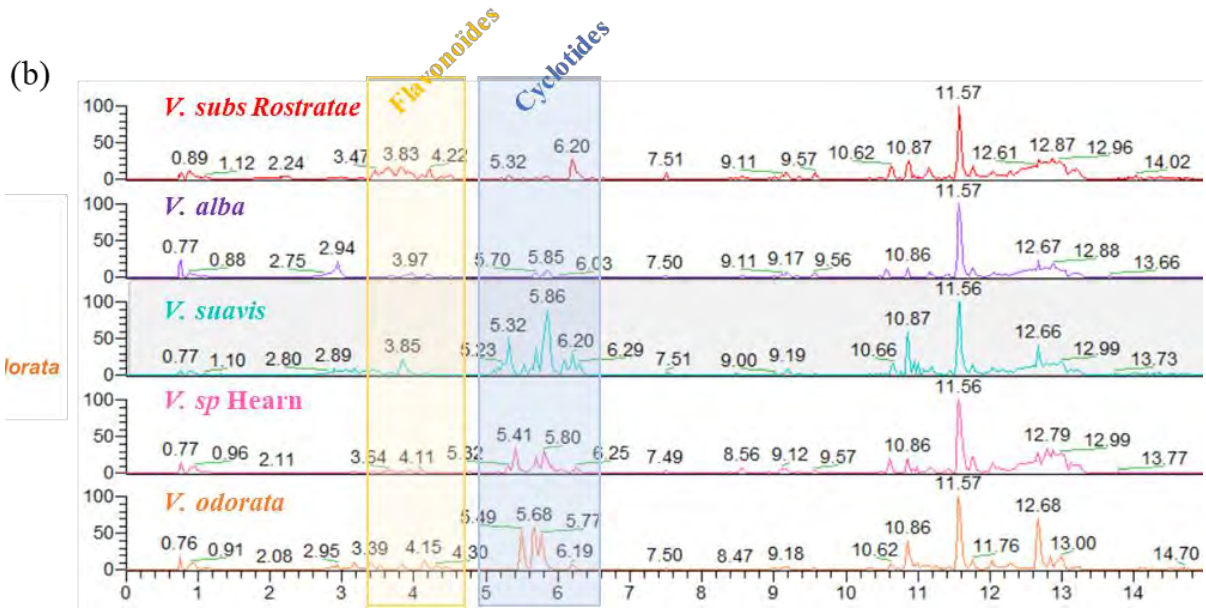


Figure 94. Proportion des espèces impliquées dans les activités antifongique et de défense des plantes (a) et profils MS associés (b)

Différentes études métabolomiques ont été menées sur les résultats biologiques obtenus et aucune relation espèce-activité n'a pu être clairement mise en évidence. En effet, plusieurs espèces ont présenté des activités sur les mêmes souches et des disparités de réponses biologiques ont été observées au sein d'une même espèce. Or, il a été démontré des différences chimiques claires entre les différentes espèces, ce qui ne permet donc pas de conclure clairement sur ces composés responsables de ces activités. Il faut néanmoins prendre en compte l'influence de la souche elle-même qui pourrait également expliquer la disparité des réponses. Cependant, la présence de variétés au sein de la collection pourrait être un élément de réponse pour expliquer ces disparités. L'exemple de la souche *Fusarium* particulièrement sensible aux extraits de *V. sp* Hearn illustre bien ce propos puisque les trois plants ayant présenté une activité sont clairement séparés des autres plants sur une ACP, mettant ainsi en évidence une certaine variabilité chimique entre les échantillons. Ces résultats pourraient donc être à mettre en relation avec la présence éventuelle de variétés au sein des espèces identifiées. En effet, l'analyse génétique réalisée dans ces travaux n'a pas permis d'identifier les variétés et sous-variétés en présence, mais uniquement les espèces. Des études génétiques complémentaires pourraient donc être envisagées pour pouvoir approfondir cette hypothèse et expliquer la disparité des réponses biologiques au sein d'une même espèce.

Il a été possible de déterminer de manière statistique les composés potentiellement actifs, mais cela uniquement pour la souche Fusarium, en raison d'un modèle OPLS valide et donc prédictif. En accord avec la bibliographie, ces derniers ont été majoritairement identifiés comme des cyclotides, et plus particulièrement des cycloviolacines connues pour leurs nombreuses propriétés biologiques, dont insecticides (**Barbeta et al., 2008** ; **Ribeiro et al., 2013** ; **Parsley et al., 2018**). Des études complémentaires et plus approfondies ont été entreprises sur deux *V. sp* Hearn et une *V. odorata* afin de valider l'hypothèse portant sur les cyclotides. Le fractionnement des extraits bruts complexes a alors permis de mettre en relief l'activité des fractions enrichies en cyclotides, d'après les données UHPLC-HRMS et les quelques résultats de dosage que l'on a pu effectuer.

II.4.4 CRIBLAGE DE L'ACTIVITÉ CYTOTOXIQUE

Des études de cytotoxicité sur cellules VERO (acronyme provenant de la combinaison de «verda» = singe africain vert et «reno» = rein en espéranto) ont été conduites pour vérifier la toxicité des extraits en vue d'une application dermo-cosmétique (cf. IV.6.2, page 193). Le doute est porté sur les coumarines. La doxorubicine (M = 543,525 g/mol) a été choisie comme molécule de référence en remplacement du Triton, préalablement utilisé, mais influençant la croissance des cellules dans les puits voisins, conduisant ainsi à des résultats erronés pour la moitié de la plaque.

II.4.4.1 Résultats du criblage cytotoxique

Une visualisation préalable des cellules a été faite afin de confirmer l'action de la doxorubicine sur le développement cellulaire et valider l'usage de ce témoin positif (Figure 95).

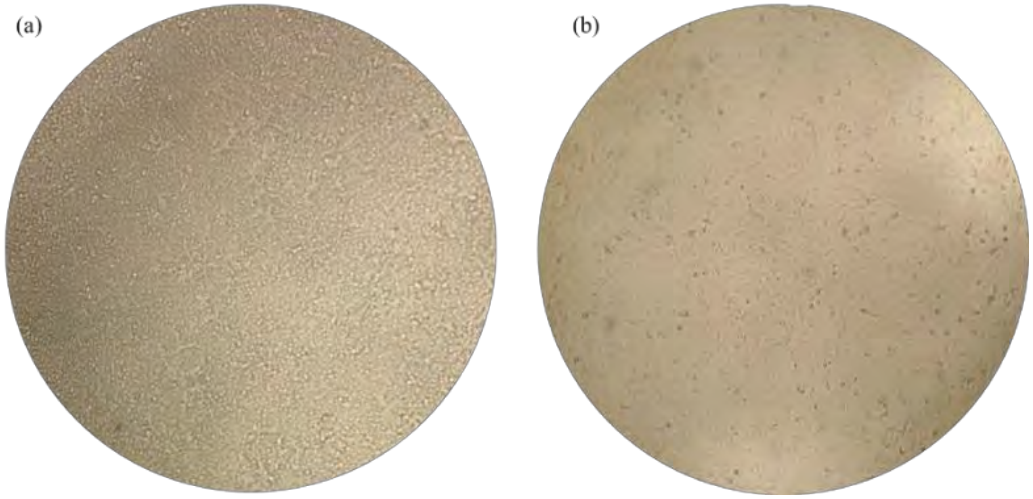


Figure 95. Visualisation des cellules d'un puit comprenant uniquement du milieu MEM (a) et d'un puit avec de la doxorubicine à 0,5 µM (b)

Il a été constaté que dans le puit de référence, les cellules sont à confluence et de forme circulaire. Au contraire, dans le puit comprenant de la doxorubicine, on observe un allongement cellulaire et donc une diminution de la densité de cellules. La doxorubicine est donc une molécule adaptée comme témoin positif pour valider les expériences puisqu'elle agit sur la croissance cellulaire. Une gamme de doxorubicine a été réalisée afin d'établir la CI_{50} de cette molécule. On observe l'obtention d'une sigmoïde (Figure 96 a) indiquant la possibilité de déterminer la CI_{50} en travaillant à partir de l'équation de la partie linéaire. Pour cela les deux points extrêmes ont été éliminés (Figure 96 b) et l'équation $y = -0,221x + 0,8663$ a été obtenue. On obtient alors une CI_{50} (doxorubicine) de 1,6 µM soit 0,9 µg/mL.

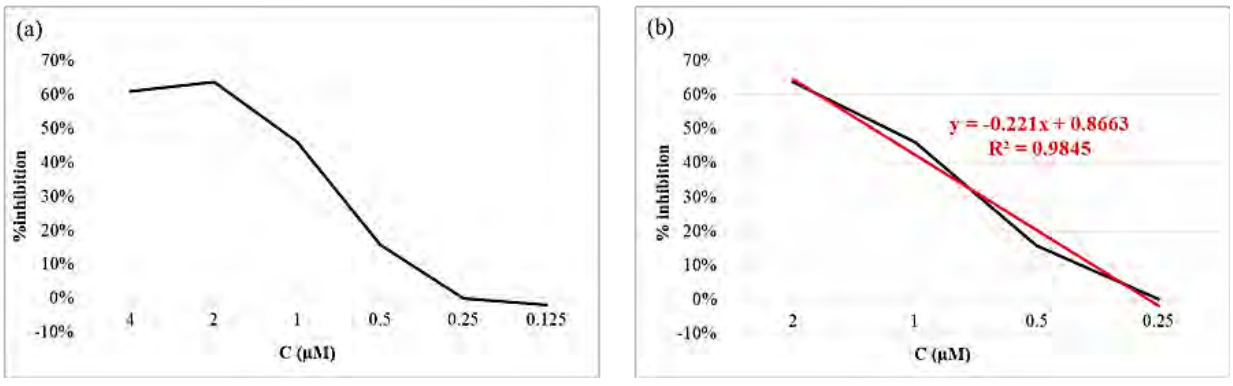


Figure 96. Courbe d'inhibition de la gamme de doxorubicine et équation de la courbe de tendance de la partie linéaire permettant le calcul de la CI_{50}

Au total, 32 échantillons ont été testés sur l'ensemble des extraits de violettes réalisés. Les 15 violettes révélées les plus antioxydantes par le criblage DPPH ainsi que les 15 les moins antioxydantes ont été sélectionnées ainsi que les deux fractions SPE de la violette de Toulouse enrichies en composés antioxydants, Fr2 et Fr3. Étant donné le nombre d'extraits, il a été décidé d'effectuer un criblage en testant 3 concentrations par extrait, le tout en triplicata. Pour des raisons de solubilité évidentes dues à la variabilité des extraits, le DMSO a été choisi comme solvant et les extraits ont été préparés à 5 mg/mL dans des conditions stériles. La contrainte première étant de ne pas dépasser 0,5% de DMSO dans le puit final pour ne pas influencer la croissance des cellules, la concentration de 25 µg/mL n'a pas pu être dépassée dans le puit final. Ainsi, les trois concentrations suivantes, 25, 6,250 et 1,5625 µg/mL ont été testées. Les échantillons ont été regroupés par espèce et les résultats moyennés (Figure 97).

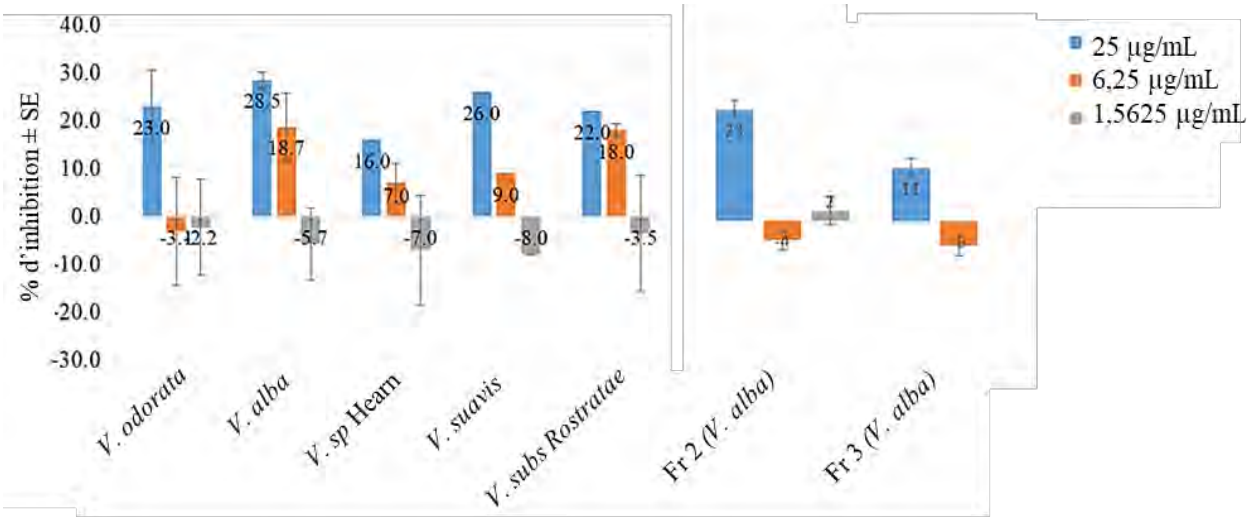


Figure 97. Résultats moyens des tests de cytotoxicité pour chaque espèce et méthodes d'extraction

Il est constaté que la CI_{50} n'est jamais atteinte puisque pour la concentration maximale de 25 µg/mL, l'inhibition maximale enregistrée est d'environ 30%. Pour pouvoir la calculer, il aurait fallu étendre la gamme à des concentrations supérieures, ce qui n'a pas pu être fait en raison de la concentration initiale des échantillons et la contrainte sur le pourcentage de DMSO autorisé dans les puits. En comparaison à la référence doxorubicine de CI_{50} égale à 0,9 µg/mL, les extraits de violettes ne semblent pas cytotoxiques puisque leur CI_{50} est plus de 25 fois supérieure. Pour la concentration la plus forte en extrait, seules les *V. sp* Hearn se démarquent avec une inhibition inférieure de l'ordre de 15% contre 25% pour les autres espèces. À la concentration de 6,25 µg/mL, les *V. alba* et *V. subs Rostratae* sont celles qui présentent les inhibitions les plus hautes, autour de 18% contre moins de 10% pour les autres, avec même une réponse négative pour *V. odorata*. Cette réponse négative est ensuite observée pour toutes les espèces à la concentration la plus faible en extrait. Les coumarines spécifiquement présentes chez *V. alba* pourraient être responsables de cette différence d'activité comme en témoigne le résultat de la fraction Fr2 avec une inhibition de 23%. Les flavonoïdes sont écartés de cette activité d'après les résultats de la fraction Fr3. Une autre catégorie de molécules est suspectée : les cyclotides qui sont particulièrement présentes chez *V. odorata et V. sp* Hearn.

En conclusion, les extraits de violette ne présentent pas de toxicité significative. Même si quelques molécules types coumarines, présentes chez *V. alba*, semblent conduire à une diminution de la croissance cellulaire, la concentration pour atteindre un seuil de cytotoxicité avéré est supérieure à 25 µg/mL soit une concentration au moins 30 fois plus élevée que pour la doxorubicine. En général, une substance est jugée cytotoxique à partir d'une CI_{50} de l'ordre de 10 µg/mL, ce qui n'est pas le cas pour l'ensemble de nos extraits, même celui enrichi en coumarines. D'autres molécules semblent affecter la

croissance cellulaire et sont présentes chez *V. odorata*. L'hypothèse de base porte sur les cyclotides et pour affiner ces suppositions, une analyse métabolomique prenant en compte les valeurs d'inhibition pour 25 µg/mL va être réalisée.

II.4.4.2 Analyse statistique des données de cytotoxicité

II.4.4.2.1 ANOVA

De même que pour le test DPPH, une analyse ANOVA a été réalisée afin de déterminer s'il existe une différence significative entre les différentes espèces et mettre en évidence une potentielle relation extrait-espèce/cytotoxicité. L'analyse de normalisation réalisée sur les résidus centrés réduits issus de l'ANOVA a conduit à des valeurs de *p* respectives de 0,32 pour le test de Shapiro et 0,60 pour le test de Jarque-Bera. Ces valeurs étant supérieures au seuil de signification α = 0,05, on ne peut pas rejeter l'hypothèse de départ qui est toujours de dire que les données sont normalisées. Ce résultat est confirmé par le graphique Q-Q sur lequel on observe une légère déviation à la diagonale (Figure 98).

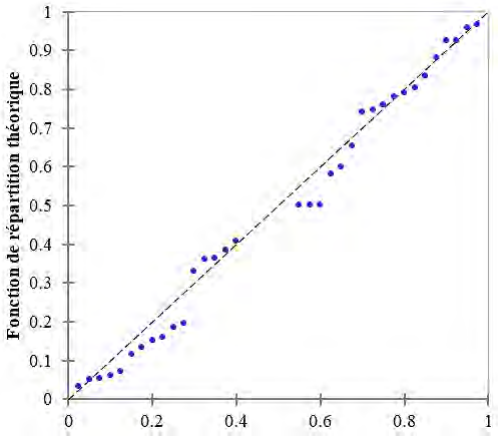


Figure 98. Graphique Q-Q de l'analyse de normalité des résidus centrés réduits de l'ANOVA des données de cytotoxicité

Sur ces considérations, l'ANOVA a donc été interprétée en détail et les résultats des différents paramètres à prendre en compte sont présentés dans le Tableau 32.

Tableau 32. Résultats de normalisation et de l'ANOVA sur les données de cytotoxicité

p-value Shapiro	p-value Jarque-Bera	R^2	Test de Fischer Pr > F	Test de Tukey
0,32	0,60	0,52	0,3%	Aucune différence significative

L'obtention d'un coefficient d'ajustement R^2 = 0,52 indique que 52% de la variabilité est expliquée par l'espèce. Le test de Fisher F utilisé dans l'analyse de la variance est inférieur à 5%, seuil de signification généralement proposé, indiquant ainsi un risque de 5% de mal conclure quant à une réelle différence. Dans ces cas, l'on prendrait un risque de 0,3% en concluant que l'espèce des violettes apporte une quantité d'information significative au modèle. Cependant, le test de Tukey HSD comparant deux à deux les échantillons ne permet pas de conclure quant à une réelle différence significative entre les espèces ni entre les différents extraits testés.

II.4.4.2.2 ACP et OPLS

Les données UHPLC-HRMS en modes négatif et positif ont été combinées par l'intermédiaire de MS-combine. Des modèles statistiques ACP et OPLS (Figure 99), avec comme données Y les pourcentages d'inhibition de croissance de cellules VERO à 25 µg/mL, ont ensuite été construits à partir des données normalisées non logarithmées et l'échelle pareto. L'objectif était d'utiliser les coefficients de la régression linéaire afin de mettre en évidence les composés qui pourraient contribuer à un effet cytotoxique, puisque les résultats ont tout de même mis en évidence quelques disparités au sein des échantillons analysés. Les coefficients étant basés sur l'inhibition de la croissance des cellules, ce sont les coefficients les plus élevés qui reflèteront le potentiel cytotoxique des composés.

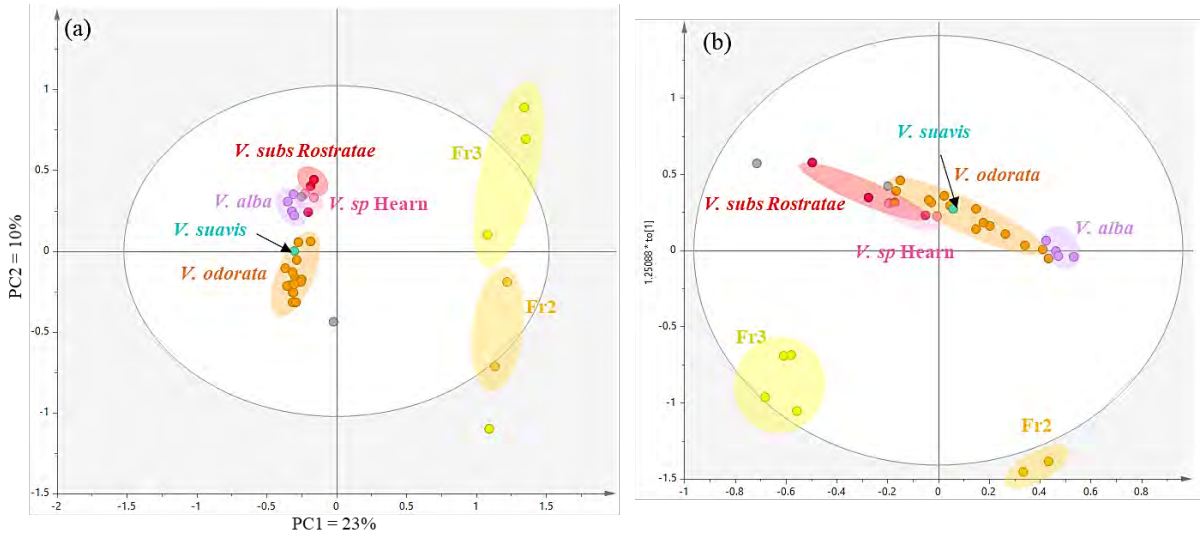


Figure 99. ACP (a) et OPLS (b) des données UHPLC-HRMS pour l'étude de la cytotoxicité

Pour l'ACP, on constate que les catégories d'échantillons sont bien définies, en particulier par l'axe PC2 : les violettes d'une même espèce forment un groupe, les fractions enrichies en coumarines (Fr2) et en flavonoïdes (Fr3) sont bien séparées des autres extraits. Cependant, le modèle OPLS n'était pas suffisamment convenable, aux regards des coefficients R2X = 0,291 et Q2Y = 0,158 ainsi que de la valeur de *p* du CV-anova égale à 0,29 (supérieur à 0,05). L'interprétation des coefficients ne peut donc être approfondie puisque notre modèle ne présente pas les critères adéquats à la mise en œuvre d'une prédiction quelconque. Ces résultats pourraient s'expliquer par la disparité d'activité au sein d'une même espèce, en particulier chez *V. odorata*, qui ne permettrait pas de bien différencier les échantillons aux vus de leur profil chimique proche. Néanmoins, les violettes appartenant aux espèces *V. alba* et *V. odorata* ont présenté les pourcentages d'inhibition parmi les plus élevés. On peut donc supposer que certaines des molécules responsables pourraient être les coumarines, spécifiques de la première espèce et dont la fraction enrichie présentait une inhibition de près de 25% ; et de certains cyclotides, abondants chez la deuxième espèce et connus dans la littérature pour cette activité. Leur concentration dans les extraits de violette reste tout de même telle que la cytotoxicité de l'extrait brut n'est pas avérée par les tests *in vitro*. Les extraits semblent donc conformes pour une utilisation en cosmétique. Des tests sur cellules de peau seraient toutefois à envisager pour confirmer ces premiers résultats.

II.5 ETUDE DE MISE A L'ECHELLE INDUSTRIELLE

Sur la base des résultats obtenus pour l'extrait éthanolique de violette de Toulouse étudiés en détail précédemment, d'autres méthodes d'extraction ont été testées sur la même violette afin de déterminer la plus adaptée à l'obtention des composés antioxydants d'intérêt.

II.5.1 DESCRIPTION DES METHODES D'EXTRACTION TESTEES

Des techniques respectant les concepts de la chimie verte, et adaptées à l'extraction des polyphénols, ont été étudiées à savoir, l'électroporation (cf. IV.5.2.2.3, page 180) mettant en œuvre l'application d'un champ électrique comme dans l'extraction aqueuse des polyphénols de la bourrache (**Segovia et al., 2015**), les micro-ondes (cf. IV.5.2.2.5, page 182) utilisées par exemple pour l'extraction des polyphénols du thé vert (**Nkhili et al., 2009**) et le CO_2 supercritique (cf. IV.5.2.2.4, page 181) tel que dans l'extraction des composés phénoliques des feuilles de l'olivier (**Le Floch et al., 1998**). Les principales caractéristiques de ces méthodes sont présentées Figure 100.

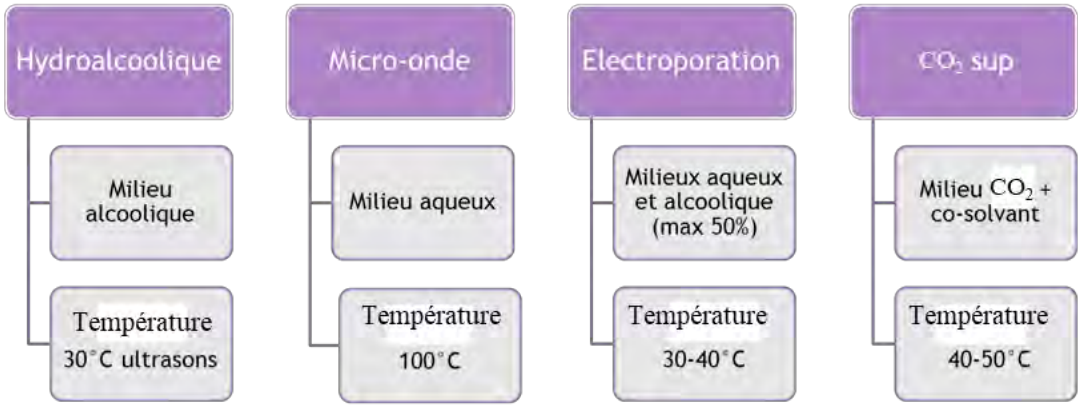


Figure 100. Principales caractéristiques des méthodes d'extraction comparées

Pour cette étude de comparaison, l'extrait brut de référence obtenu par extraction éthanolique à 80% a été comparé à :

- 20 extraits issus de l'électroporation dont la moitié obtenue à partir de feuilles fraîches et l'autre moitié à partir de feuilles congelées. Ils ont été obtenus soit par une extraction aqueuse soit éthanolique,
- 30 extraits issus du CO_2 supercritique. Ils ont été obtenus par extraction des feuilles lyophilisées broyées avec augmentation progressive du rapport de co-solvant éthanol entre 5% et 50%,
- 2 extraits issus de l'extraction assistée par micro-ondes. Ils ont été obtenus par extraction des feuilles fraiches sans ajout de solvant, mais en exploitant uniquement l'eau résiduelle des feuilles.

Un contrôle qualité pour chaque technique et un pool total (QC_{tot}) ont été préparés. Ainsi, 57 échantillons ont été analysés.

II.5.2 Profils UHPLC-HRMS

L'ensemble des échantillons a été analysé en UHPLC-HRMS/MS et traité par une analyse métabolomique afin d'étudier la disparité des profils chimiques (cf. IV.5.2.4.3, page 187). Une première ACP a été faite afin de rendre compte de la répartition générale des données (Figure 101 a). Il a ainsi été remarqué la formation de deux principaux groupes séparés par l'axe PC1 et confirmés par une analyse hiérarchique (CAH) (Figure 101 b) : les extraits issus des micro-ondes et de l'électroporation forment un premier groupe, tandis que les extraits issus de l'extraction au CO_2 supercritique forment un second ensemble bien séparé. L'extrait hydroalcoolique semble se situer plutôt au centre de l'espace. On peut donc conclure de ces premiers résultats que les méthodes mises en œuvre conduisent à des profils chimiques différents, notamment attendus plutôt apolaires pour le CO_2 supercritique et plutôt polaires pour les micro-ondes et l'électroporation, dû à l'utilisation d'un milieu plus aqueux. L'extrait hydroalcoolique présente une composition intermédiaire qui s'apparenterait au QC_{tot} en raison de sa position centrale sur le graphique. Ces premières constatations rendent pertinente l'analyse plus poussée des extraits.

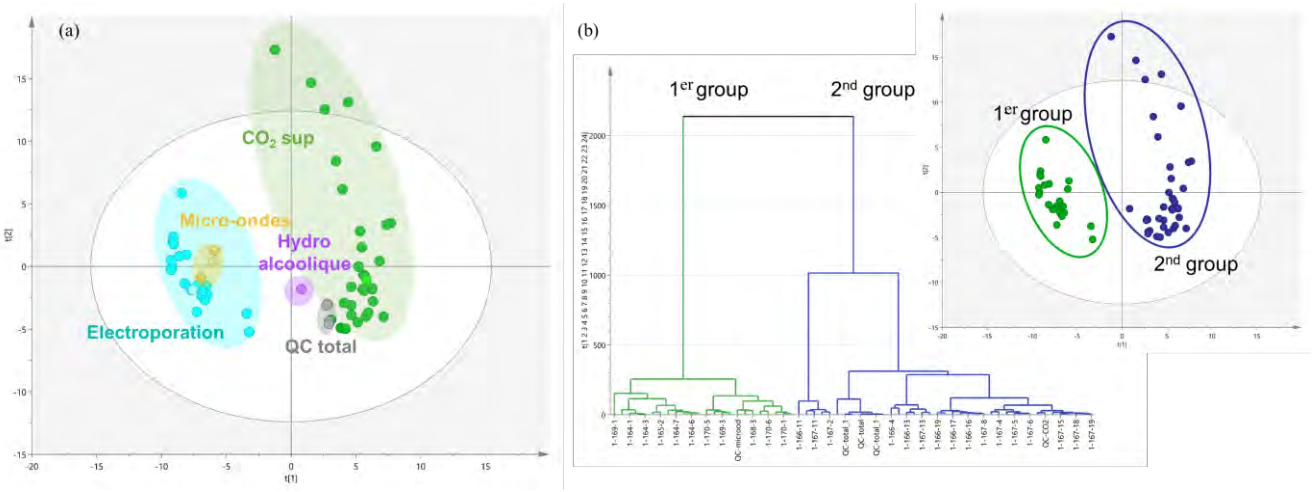


Figure 101. ACP des profils chimiques des différents extraits de violettes (a) et CAH avec mise en évidence de deux principaux groupes (b)

Deux principaux critères ont été pris en compte dans la sélection de la méthode d'extraction, le rendement d'extraction et le potentiel antioxydant par l'intermédiaire du dosage des polyphénols. Des tests DPPH ont été entrepris sur quelques échantillons sélectionnés par « model design » : un échantillon représentatif dans chaque groupe homogène observé en ACP est testé.

II.5.3 Rendement et potentiel antioxydant

II.5.3.1 Rendement d'extraction

Le rendement d'extraction a été calculé pour chaque extrait obtenu par les différentes méthodes employées (Figure 102).

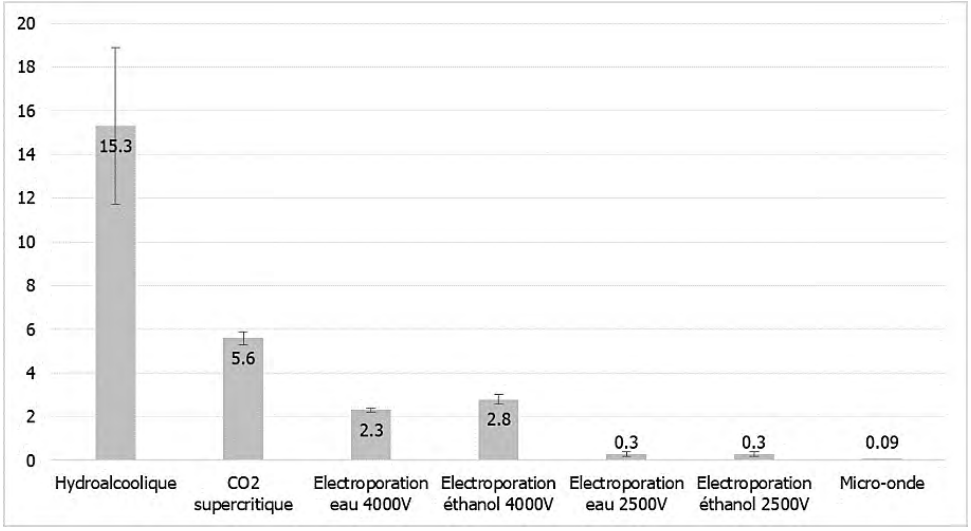


Figure 102. Rendements d'extraction des différentes méthodes employées

Ces résultats montrent que le rendement d'extraction est plus important dans le cas d'une matière végétale broyée que dans le cas de feuilles entières, congelées ou fraîches s'expliquant par la surface de contact matière-solvant plus efficiente dans le cadre d'une poudre.

Les extractions aqueuses et hydroalcooliques par électroporation sont relativement similaires puisque les rendements sont du même ordre de grandeur à savoir 0,3% pour les feuilles congelées et 2-3% pour les feuilles fraîches. Le rendement plus élevé dans le cas des feuilles fraîches est probablement dû en partie à la différence de champ électrique appliqué. En effet, il est 1,5 fois plus élevé que pour les feuilles congelées. Par ailleurs, la quantité de solvant était plus importante pour les feuilles fraîches avec l'ajout de 100 g de solvant contre 80 g pour les feuilles congelées, conduisant à des rapports solvant/matière respectifs de 3 et 1. En effet, la différence de volume entre feuilles congelées et fraîches n'a pas permis de mettre autant de matière, laissant dans le second cas plus d'espace au solvant pour assurer le remplissage intégral de la cuve.

L'extraction par micro-ondes est la technique la moins efficace avec un rendement de près 0,1% à partir de feuilles fraîches. La méthode exploitant l'eau résiduelle dans les feuilles, le rapport solvant/matière est nettement inférieur aux autres techniques, ce qui pourrait expliquer ce résultat. Cependant, cet extrait est le plus intéressant d'un point de vue applicatif puisqu'il est très peu coloré (jaune clair). Il pourrait donc être directement formulé sans purifications supplémentaires, au contraire des autres extraits qui présentent une coloration verte plutôt intense, en particulier les extraits obtenus par CO_2 supercritique, due à la présence de chlorophylle.

Au contraire, l'extraction hydroalcoolique semble être la plus efficiente avec un rendement moyen sur la poudre de 15% contre 5,5% pour le CO_2 supercritique. Dans ce cas le rapport solvant/matière égal à 10 était aussi le plus élevé, atténuant ainsi le phénomène de saturation limitant l'extraction des métabolites.

II.5.3.2 Dosage des polyphénols

Des dosages spécifiques des coumarines et flavonoïdes ont été réalisés par UHPLC-UV, en complément du dosage des polyphénols totaux par Folin-Ciocalteu et des flavonoïdes par la méthode $AlCl_3$. Les moyennes pour chaque méthode ont été calculées et les résultats sont présentés, Figure 103, par mg d'extrait brut. Par ailleurs, les résultats pour les polyphénols sont présentés en mg équivalent d'acide gallique (AG) et en mg équivalent de rutine (R) pour les flavonoïdes. Pour l'essai

hydroalcoolique, les résultats de la violette V109, identifiée comme *V. alba* ont été intégrés au détriment de ceux de la violette de Toulouse en raison du vieillissement de l'échantillon, utilisé comme référence dans de nombreux dosages et donc de la probable dégradation des coumarines sensibles à la lumière. Pour obtenir un résultat plus proche de la réalité, il a été décidé d'utiliser l'un de ses homologues.

On constate une disparité de réponses en fonction de la méthode de dosage choisie avec des valeurs plus faibles au niveau du dosage réalisée par UHPLC-UV, et au contraire des valeurs plus hautes avec le dosage réalisé par la méthode Folin-Ciocalteu. Des différences respectives de l'ordre de 10^2 et 10 sont observées pour les échantillons issus de l'extraction par électroporation entre le dosage UHPLC-UV et les dosages spectrophotométriques. Il est à noter que le réactif Folin-Ciocalteu n'est pas spécifique des polyphénols, mais réagit également avec les sucres ou encore les acides aminés. Des teneurs plus fortes que pour les autres méthodes sont donc justifiables.

Malgré la variété de réponses des différents dosages, les tendances de teneurs entre les échantillons sont similaires. En effet, les extraits issus de l'électroporation et des micro-ondes sont ceux qui présentent les teneurs les plus faibles, tandis que le CO_2 supercritique et l'extrait hydroalcoolique sont ceux qui semblent extraire les polyphénols de manière plus efficace. Pour ces deux derniers, peu de différences significatives sont constatées au niveau des dosages colorimétriques, au contraire du dosage UHPLC-UV pour lequel les teneurs en coumarines sont similaires, mais la teneur en flavonoïdes est nettement supérieure pour l'extrait hydroalcoolique.

(a)

Méthode	**Dosage LC-UV coumarines** Teneur dans l'extrait brut (mg/mg)	**Dosage LC-UV flavonoïdes** Teneur dans l'extrait brut (mg/mg)	**Polyphénols totaux** Teneur dans l'extrait brut mgAG*/mg	**Flavonoïdes totaux** Teneur dans l'extrait brut mgR**/mg
Micro-onde	$2{,}6 \times 10^{-4}$	$3{,}2 \times 10^{-4}$	$1{,}4 \times 10^{-2} \pm 0{,}8 \times 10^{-3}$	$0{,}2 \times 10^{-2} \pm 3{,}0 \times 10^{-3}$
Electroporation eau (C)	$4{,}6 \times 10^{-4} \pm 3{,}2 \times 10^{-5}$	$2{,}2 \times 10^{-3} \pm 1{,}2 \times 10^{-3}$	$1{,}7 \times 10^{-2} \pm 0{,}1 \times 10^{-3}$	$0{,}3 \times 10^{-2} \pm 1{,}6 \times 10^{-3}$
Electroporation eau (F)	$4{,}3 \times 10^{-4} \pm 8{,}9 \times 10^{-5}$	$8{,}8 \times 10^{-4} \pm 3{,}8 \times 10^{-4}$	$1{,}6 \times 10^{-2} \pm 0{,}4 \times 10^{-3}$	$0{,}2 \times 10^{-2} \pm 0{,}0$
Electroporation EtOH 30% (C)	$1{,}7 \times 10^{-4} \pm 5{,}4 \times 10^{-5}$	$1{,}4 \times 10^{-3} \pm 5{,}1 \times 10^{-4}$	$1{,}9 \times 10^{-2} \pm 0{,}1 \times 10^{-3}$	$0{,}3 \times 10^{-2} \pm 0{,}8 \times 10^{-3}$
Electroporation EtOH 30% (F)	$5{,}9 \times 10^{-4} \pm 2{,}5 \times 10^{-4}$	$6{,}3 \times 10^{-4} \pm 1{,}7 \times 10^{-4}$	$2{,}2 \times 10^{-2} \pm 0{,}5 \times 10^{-3}$	$0{,}8 \times 10^{-2} \pm 1{,}1 \times 10^{-3}$
CO_2 supercritique	$1{,}9 \times 10^{-2} \pm 2{,}6 \times 10^{-3}$	$8{,}4 \times 10^{-3} \pm 8{,}4 \times 10^{-4}$	$2{,}2 \times 10^{-2} \pm 0{,}2 \times 10^{-3}$	$2{,}2 \times 10^{-2} \pm 0{,}3 \times 10^{-3}$
Hydroalcoolique (V109)	$1{,}8 \times 10^{-2}$	$2{,}8 \times 10^{-2}$	$3{,}2 \times 10^{-2} \pm 1{,}7 \times 10^{-3}$	$1{,}4 \times 10^{-2} \pm 1{,}2 \times 10^{-3}$

C = feuilles congelées ; F= feuilles fraîches ; * en équivalent acide gallique (AG) ; ** en équivalent rutine (R)

(b)

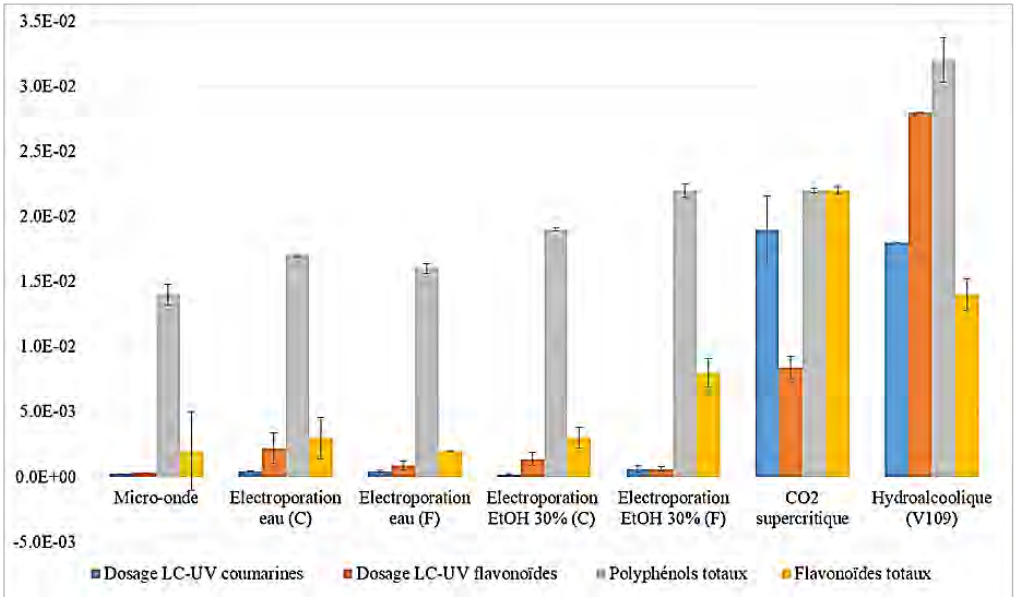


Figure 103. Résultats des dosages UHPLC-UV et spectrophotométriques des polyphénols dont coumarines et flavonoïdes des différents extraits (a) et histogramme associé (b)

II.5.3.3 Potentiel antioxydant

II.5.3.3.1 « Multivariate design »

Les extraits ont été testés par essais DPPH et RPE. Cependant, le nombre d'extraits étant conséquent, notamment pour les méthodes d'électroporation et de CO_2 supercritique, un « multivariate design », qui repose sur la sélection d'un échantillon représentatif de chaque groupe homogène, a été mis en place pour la sélection des échantillons (**Eriksson et al., 2013**). Cette notion repose sur le fait que l'échantillon sélectionné produit des résultats extrapolables à l'ensemble du groupe. Une CAH et une ACP réalisées sur les échantillons issus de l'électroporation (Figure 104 a) ont permis de mettre en évidence cinq groupes et de sélectionner cinq échantillons caractéristiques, tout comme pour le CO_2 supercritique (Figure 104 b).

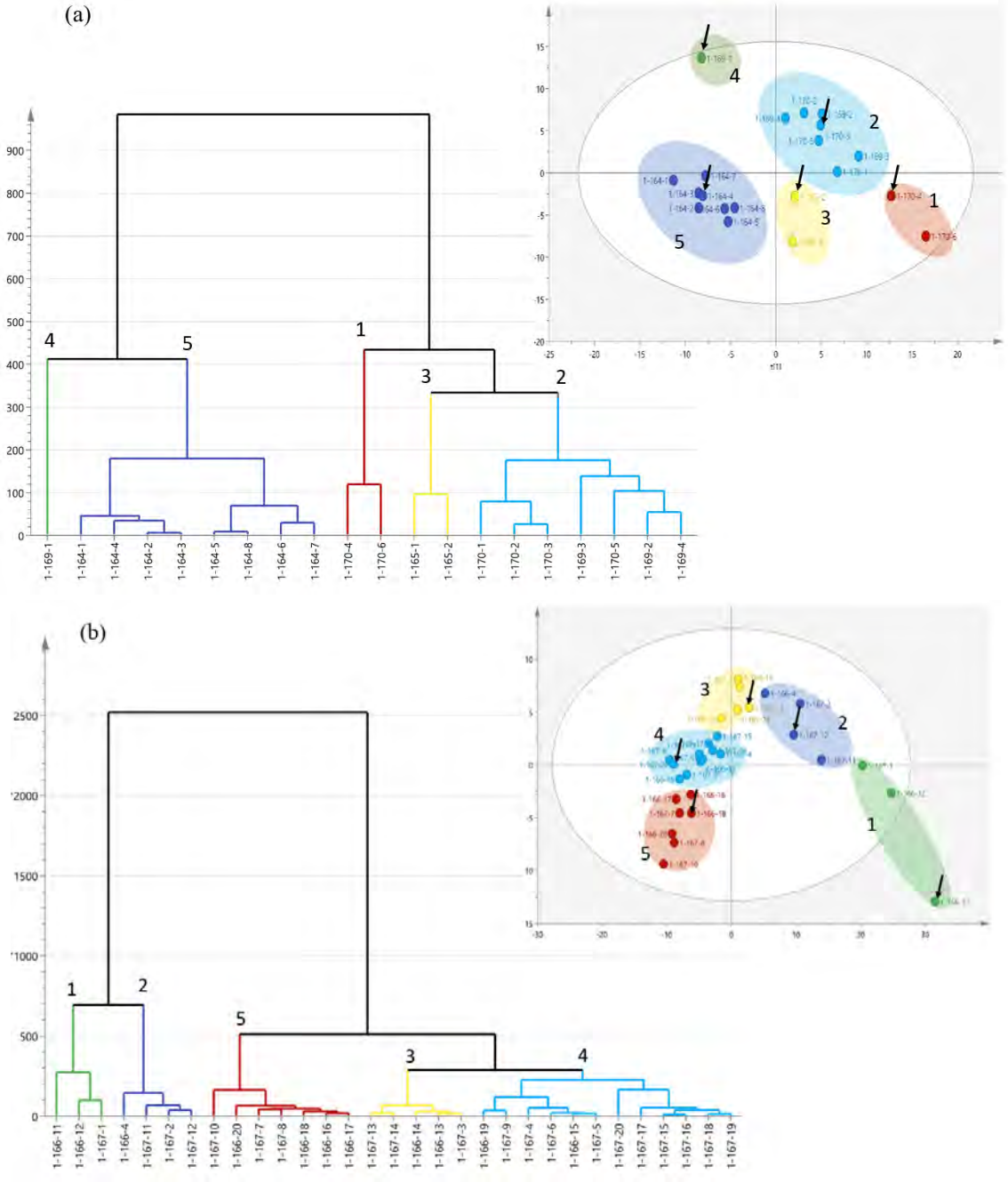


Figure 104. CAH et ACP des échantillons issus de l'électroporation (a) et du CO_2 supercritique (b) avec sélection des extraits à traiter par « multivariate design »

Pour la méthode d'électroporation, on a constaté que la séparation des extraits était essentiellement basée sur l'état des feuilles (fraîche ou congelée). En effet, les groupes 1, 2 et 4 comprennent les extraits issus de l'extraction de feuilles fraîches, tandis que les groupes 5 et 3 sont constitués des extraits issus de l'extraction des feuilles congelées (Figure 104 a). Dans les groupes 2 et 5, on retrouve les extraits aqueux ou avec EtOH à 30% ; tandis que dans les groupes 1 et 3, sont retrouvés les extraits traités avec des conditions plus fortes, c'est-à-dire avec EtOH à 50% ou un chauffage plus intense. Le groupe 4 constitué d'un seul échantillon représente le blanc H_2O pour les feuilles fraîches.

Pour le CO_2 supercritique, il est intéressant de noter que les groupes sont caractérisés par des teneurs en EtOH spécifiques : 1 et 2 sont constitués des extraits à 5 et 10% en EtOH, 3 de ceux à 15-20%, 4 de ceux à 25-45% et 5 de ceux à 35-50% (Figure 104 b).

II.5.3.3.2 Résultats des essais DPPH et RPE

Une dizaine d'échantillons a été sélectionnée : 5 issus du CO_2 supercritique, 5 de l'électroporation, 2 des micro-ondes et le brut de l'extraction hydroalcoolique. Les extraits étant de polarités variées, le diméthylsulfoxyde (DMSO) a été utilisé et validé par la droite d'étalonnage du Trolox linéaire entre 1 et 15 mg/mL avec une CI_{50} de 9,9 contre 7,0 avec l'éthanol (Figure 105).

C finales (mg/L)	% inhibition
1.33	15.12%
3.33	23.22%
6.67	39.65%
13.33	62.24%
33.33	76.49%

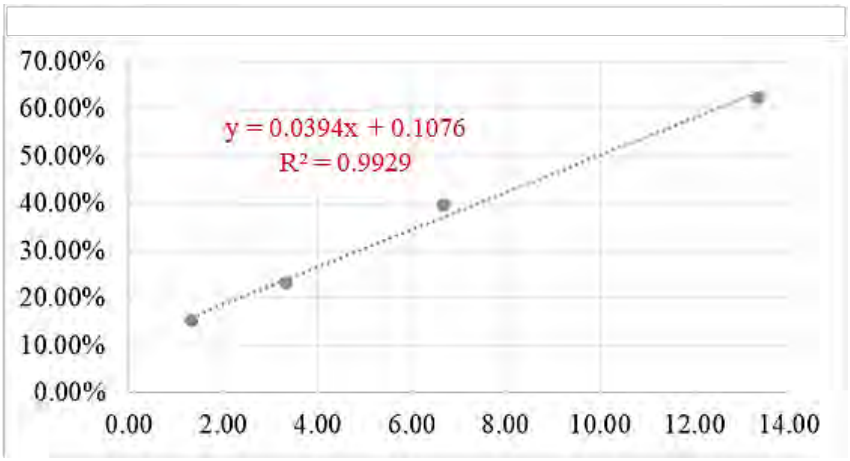


Figure 105. Résultats de la gamme d'étalonnage du Trolox dans le DMSO

Les résultats des essais DPPH et RPE sont présentés Figure 106. On constate que les CI_{50} des extraits les plus aqueux issus de l'électroporation et des micro-ondes sont très élevées. La présence de mucilage en forte concentration est suspectée dans ces extraits par leur forte viscosité, qui engendrerait ce résultat antioxydant (CI_{50} >1000). Les extraits obtenus par le CO_2 supercritique présentent une CI_{50} aux alentours de 500 mg/L du même ordre de grandeur que les extraits issus de l'extraction hydroalcoolique. Le second test RPE permet de discriminer l'extraction hydroalcoolique qui présente une intégration de l'ordre de 7, contre 10 pour les trois autres méthodes.

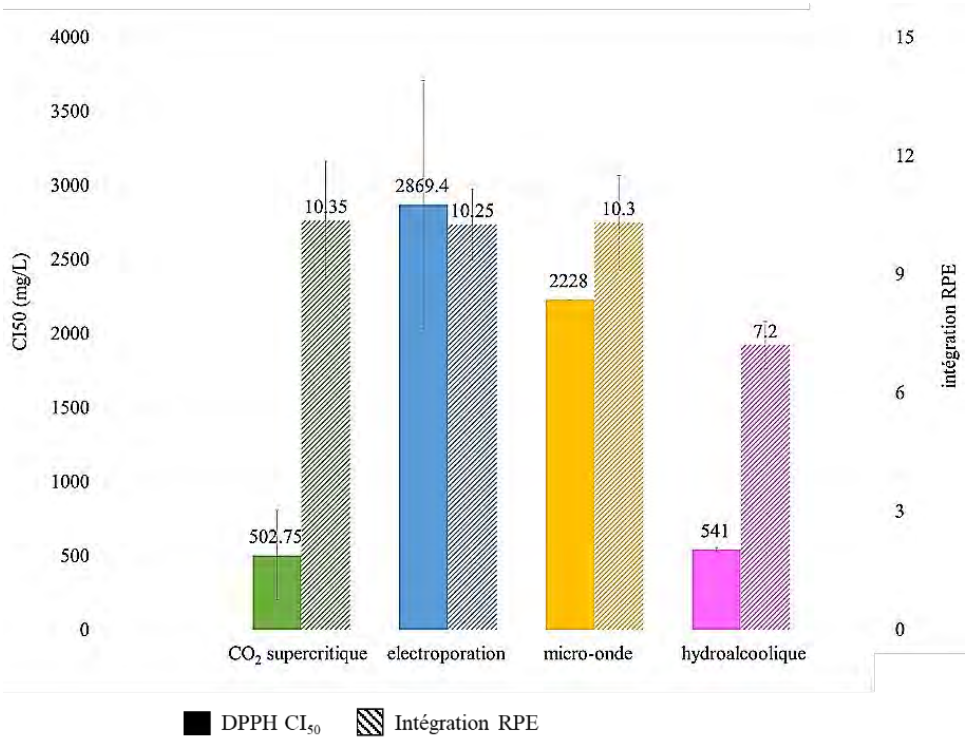


Figure 106. Valeurs moyennes des CI_{50} et intégration des signaux RPE pour chaque méthode d'extraction

Les dosages et les résultats des tests antioxydants ont ainsi permis de confirmer que l'extraction assistée par micro-ondes et l'électroporation n'étaient pas pertinentes du point de vue de l'extraction des polyphénols et des composés antioxydants, que ce soit en milieu aqueux ou hydroalcoolique. En revanche, les extractions par CO_2 supercritique et solvant hydroalcoolique sont les plus appropriées. Cependant, la coloration de la solution obtenue en CO_2 supercritique est nettement plus prononcée en raison de l'extraction plus importante des pigments, dont la chlorophylle, due à son caractère plus apolaire. La technique hydroalcoolique semble donc la méthode la mieux adaptée en termes d'extraction des composés polyphénoliques et de couleur de l'échantillon, paramètre important pour une future formulation. Par ailleurs, elle reste une technique facile à mettre en œuvre et qui ne nécessite pas l'utilisation d'installations particulières. De nombreux sous-traitants seront à même de proposer une extraction à l'échelle industrielle de ce type. Le principal inconvénient émane de la couleur de l'extrait qui ne correspond pas au cahier des charges de la cosmétique. Néanmoins, une décoloration sur charbon actif est réalisable et souvent mise en œuvre à l'échelle industrielle puisque les problèmes de coloration sont relativement fréquents lorsque l'on travaille sur de la matière végétale comme les feuilles. Toutefois, une étude d'optimisation de cette méthode a été préalablement proposée à l'échelle du laboratoire pour valider les conditions expérimentales les plus favorables et les plus économiques.

II.5.4 Optimisation de l'extraction hydroalcoolique

L'extraction hydroalcoolique a été révélée la plus pertinente dans notre recherche d'un extrait antioxydant. Une optimisation de l'extraction basée sur un système à reflux, comme proposé par le sous-traitant Berkem, a été réalisée à partir d'une matière sèche et broyée. Deux espèces ont été étudiées, la violette de Toulouse, notre plante référence, et la violette de Tourette, cultivée dans le Sud-Est et disponible en quantité plus importante, compatible avec une production industrielle.

Il a été décidé de valoriser la violette de Toulouse puisque même si elle présente un potentiel antioxydant inférieur aux *V. subs Rostratae*, celui-ci reste tout de même plus que satisfaisant et son

image de marque en tant que symbole de la ville rose apporte un vrai avantage économique. Une extraction industrielle par l'intermédiaire du partenaire Berkem a été entreprise afin de vérifier tout d'abord la correspondance avec les extraits obtenus à l'échelle du laboratoire, et pouvoir par la suite débuter des tests de formulation accompagnés de tests d'efficacité en lien avec la propriété antioxydante recherchée, et de stabilité. Les extraits étant à visée dermo-cosmétique, des tests préalables de cytotoxicité ont été effectués, afin de valider leur utilisation à ces fins.

II.5.4.1 Montage

Le montage a donc consisté en un système à reflux avec réfrigérants et chauffage au bain-marie avec des plaques chauffantes. Un système en série a été proposé pour effectuer six extractions en parallèle (Figure 107). Le temps d'extraction a été décompté à partir du reflux puis les solutions ont été filtrées et les extraits évaporés.

Figure 107. Système à reflux en série

II.5.4.2 Résultats d'optimisation

Deux principaux paramètres ont été étudiés, le rapport solvant/matière première (1/5, 1/7 et 1/10), et le temps de chauffage à partir du reflux (30 min et 1h). Deux paramètres ont été décisifs dans le choix du protocole qui sera adapté à l'échelle pilote : le rendement d'extraction et la CI_{50} par la méthode DPPH.

II.5.4.2.1 Rendement d'extraction

Les résultats présentés Figure 108 montrent peu de différences de rendement entre les deux espèces, à même rapport de solvant. Cependant, on observe une amélioration du rendement avec l'augmentation de rapport solvant/matière, résultat cohérent puisque plus l'on ajoute de solvant, meilleure est la surface de contact matière/molécule de solvant. Pour le rapport 1/5, on obtient en moyenne 12% de rendement contre 16% pour 1/7 et 18% pour 1/10, soit une augmentation de 25%. En revanche, l'augmentation du temps d'extraction ne semble pas conduire à de meilleurs rendements, impliquant que 30 minutes semblent suffisantes et permettraient par ailleurs d'éviter tous phénomènes de dégradation suite à une exposition trop longue.

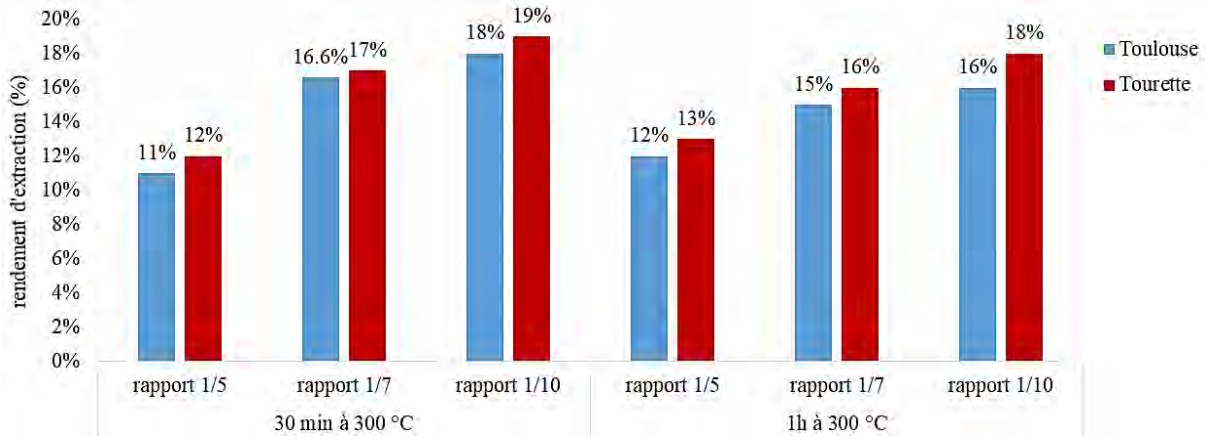


Figure 108. Rendements d'extraction aux différents ratios et temps d'extraction pour la violette de Toulouse et celle de Tourette

II.5.4.2.2 Potentiel antioxydant

Les données présentées Figure 109 confirment des résultats du laboratoire antérieurs à la thèse (**Valapimip, 2011**) démontrant que la violette de Toulouse possède un potentiel antioxydant supérieur à la violette de Tourette, puisque cette dernière présente une CI_{50} par la méthode DPPH environ trois fois supérieure à la violette de Toulouse. Par ailleurs, le potentiel antioxydant ne semble pas varier avec le rapport de solvant, indiquant que le rapport 1/5 est nettement suffisant pour l'extraction globale des composés antioxydants. L'ajout de solvant n'améliore donc pas la concentration en composés antioxydants et dans un souci de limitation de consommation de solvant pour l'échelle pilote, le rapport 1/5 sera préféré. Il est à noter que la violette de Toulouse semble la plus concentrée en polyphénols, d'après les résultats du dosage Folin-Ciocalteu.

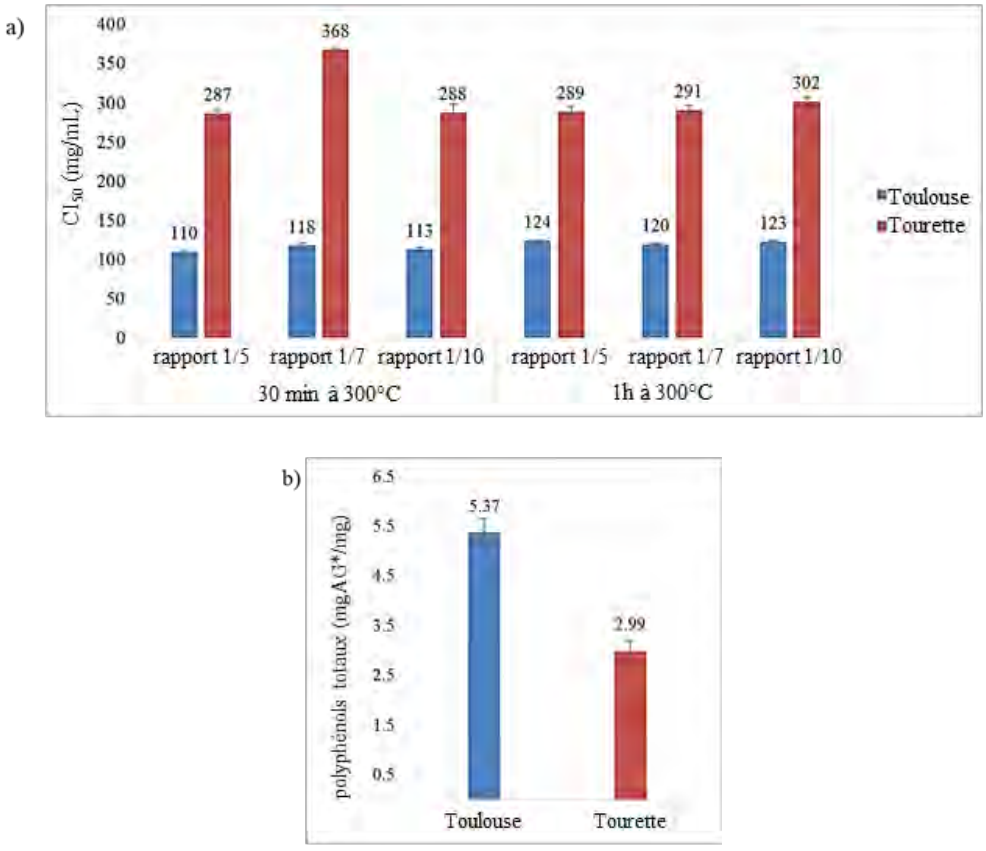


Figure 109. CI_{50} par la méthode DPPH aux différents ratios et temps d'extraction pour la violette de Toulouse et celle de Tourette (a) et teneurs en polyphénols totaux par le dosage Folin-Ciocalteu (b)

II.5.4.2.3 Profils UHPLC-HRMS/MS

La comparaison des profils chromatographiques (Figure 110) et l'interrogation des bases de données MS a mis en évidence la présence de flavonoïdes dans les deux types de violettes avec quelques divergences, dont deux flavonoïdes 25 et 41 dans la violette de Tourette absents de la violette de Toulouse (Tableau 33). En revanche, il a été constaté que la violette de Tourette était dépourvue de coumarines (1 et 2), présentes dans la violette de Toulouse et participant clairement au potentiel antioxydant, d'après les études précédentes. Ainsi, la différence d'activité constatée dans les analyses DPPH semble ici justifiée, d'autant plus que les flavonoïdes majoritaires de la violette de Tourette ne présentent pas de noyau catéchol, motif caractéristique de l'activité antioxydante.

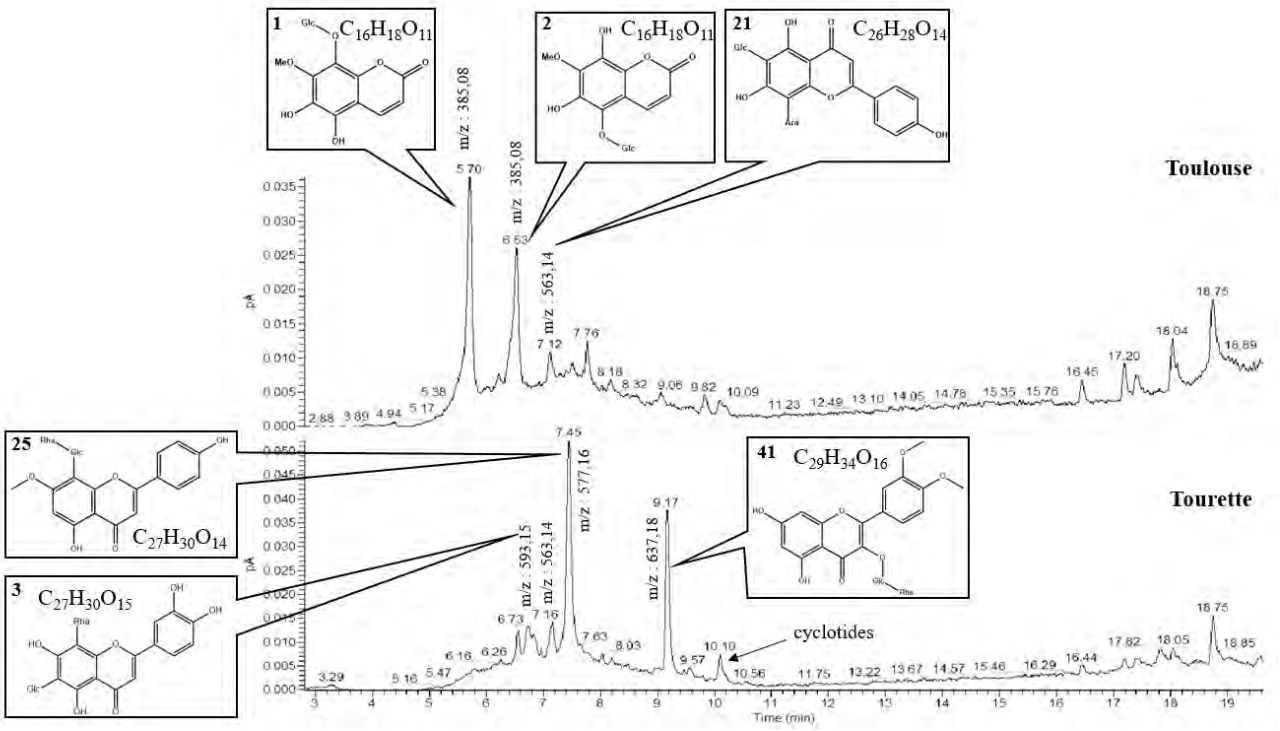


Figure 110. Comparaison des profils MS des violettes de Toulouse et de Tourette

Tableau 33. Composés majoritaires des violettes de Toulouse et de Tourette

Rang	RT (min)	[M-H]⁻	Formule brute	UV (nm)	Identification potentielle*	Classe chimique	Violette
1	5,70	385,0771	$C_{16}H_{18}O_{11}$	259 326	7-Methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside[a]	Coumarine	Toulouse
2	6,49	385,0768	$C_{16}H_{18}O_{11}$	258 326	7-Methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[a]	Coumarine	Toulouse
5	6,76	593,1500	$C_{27}H_{30}O_{15}$	268 337	6-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxy-8-*C*-α-L-rhamnopyranosylflavone[b]	Flavonoïde	Toulouse Tourette
21	7,13	563,1404	$C_{26}H_{28}O_{14}$	271 331	6-*C*-α -L-Arabinopyranosyl-8-*C*-β-D-glucopyranosyl-4',5,7-trihydroxyflavone[b]	Flavonoïde	Toulouse Tourette
25	7,43	577,1550	$C_{27}H_{30}O_{14}$	271 337	8-*C*-β-D-glucopyranosyl-4',5,7-trihydroxyflavone; 2''-*O*-α-L-rhamnopyranosyl[b]	Flavonoïde	Tourette
41	9,13	637,1763	$C_{29}H_{34}O_{16}$	267 332	3',4',5,7-Tetrahydroxyisoflavone; 3',7-Di-Me ether, 4',5-di-*O*-β-D-glucopyranoside[b]	Flavonoïde	Tourette

* Déterminés par fragmentation in silico ; [a] Confirmés par RMN ; [b] Identifications issues de la base de données *Viola*

II.5.5 Conclusion

L'extraction hydroalcoolique est la plus pertinente puisqu'elle assure un meilleur enrichissement des extraits en composés antioxydants par rapport à l'ensemble des autres techniques testées. Même si elle présente le principal inconvénient de l'obtention d'extraits verts en raison de la présence de chlorophylle, l'activité antioxydante, jugée par la teneur en coumarines et flavonoïdes, reste l'élément primordial pour départager les extraits. L'optimisation à l'échelle du laboratoire sur le modèle d'une extraction à reflux, a permis de rendre compte qu'un rapport matière/solvant de 1/5 pendant 30 min était suffisant pour extraire les composés antioxydants. Pour faire suite à ces travaux, des tests d'extraction à l'échelle pilote, ainsi que des tests de formulation en solution hydro-glycérinée parfaitement utilisable en l'état et de meilleure stabilité sans même l'ajout de conservateur, sont en cours chez le sous-traitant Berkem

Les violettes ayant été qualitativement décrites, tant au niveau de leur profil chimique que de leurs activités biologiques (antioxydante, antifongique et de défense), une étude semi-quantitative basée sur la combinaison des profils UHPLC-HRMS et CAD a ensuite été entreprise, de manière à obtenir une spéciation chimique des composés majoritaires chez les différentes espèces identifiées dans la collection.

II.6 SPECIATION SEMI-QUANTITATIVE DE LA COLLECTION

II.6.1 Contexte

Une méthode combinant les données qualitatives de l'UHPLC-MS aux données semi-quantitatives du détecteur Corona (CAD) a été mise en place afin d'obtenir une carte d'identité complète des composés majoritaires des espèces de violettes. Le détecteur Corona présente les principaux avantages d'être universel, de mesurer une large gamme d'analytes et d'avoir une réponse quasi-uniforme et indépendante de la structure chimique (**Górecki et al., 2006**). Ce detecteur mesure la charge d'un courant chargé d'analyte directement proportionnelle à la quantité dans l'extrait du composé élué par le système chromatographique. Des optimisations ont été faites afin de parvenir à des résultats pertinents, dont la majorité portant sur le nettoyage de la liste de pics UHPLC-MS. En effet, l'annotation des pics par interrogation de bases de données, par l'intermédiaire de la formule brute obtenue à partir de la masse exacte et des fragments MS/MS, est une étape sensible en raison de la nature de l'ionisation par électronébulisation (ESI) engendrant de nombreuses redondances sous la forme d'adduits, d'isotopes, de pertes de neutre ou encore de fragments formés en source. Par conséquent, la recherche de l'identification sans ces annotations préalables peut conduire à des erreurs. Néanmoins, cette dégénérescence des masses apporte des informations intéressantes puisqu'elle permet de déterminer avec certitude l'ion moléculaire et apporte des indications structurales. Le principe va donc reposer sur le regroupement des variables caractéristiques d'un même et unique composé et l'identification des liens entre ces masses, afin de pouvoir par la suite nettoyer la liste et conserver uniquement les variables pertinentes correspondant aux ions moléculaires.

II.6.2 Principe de traitement des donnees

II.6.2.1 Annotation et nettoyage des pics UHPLC-MS

Pour l'annotation des pics, le package R Mz-unity a été utilisé (https://github.com/nathaniel-mahieu/mz.unity) (**Mahieu et Patti, 2017**). Il s'agit d'un script sous le logiciel R permettant

d'automatiser ce travail d'annotation, en travaillant à partir de la matrice totale composée des pics détectés en modes PI et NI. La présence de relations entre toutes les variables de cette matrice, sans prise en compte du temps de rétention, est étudiée. Une fois les annotations effectuées, un travail d'élimination des pics redondants est entrepris afin de simplifier la liste brute obtenue à l'aide du logiciel MS-DIAL. Pour cela des règles ont été établies selon la nature de la relation entre les variables. De manière générale, l'aspect de polarité (ion moléculaire détecté en NI et PI) est celui qui permet de déterminer directement l'ion moléculaire. Par conséquent, le nettoyage commence par cette donnée et toutes les variables liées par une quelconque relation avec ces ions seront éliminées. En revanche, lorsque la polarité n'est pas obtenue, les autres relations doivent être scrupuleusement étudiées avec élimination des masses les plus hautes dans le cas d'adduits, d'isotopes et de dimères, et à l'inverse des masses les plus basses dans le cas de pertes de neutres.

II.6.2.2 Formation de la liste de pic du CAD

Les données du détecteur Corona sont extraites sous Excel à partir du logiciel Xcalibur. Le script R nommé « speaq » (https://github.com/beirnaert/speaq) (**Beirnaert et al., 2018**) est utilisé. Ce dernier repose sur la détection des pics, leur alignement et la construction de la matrice résultat (Figure 111).

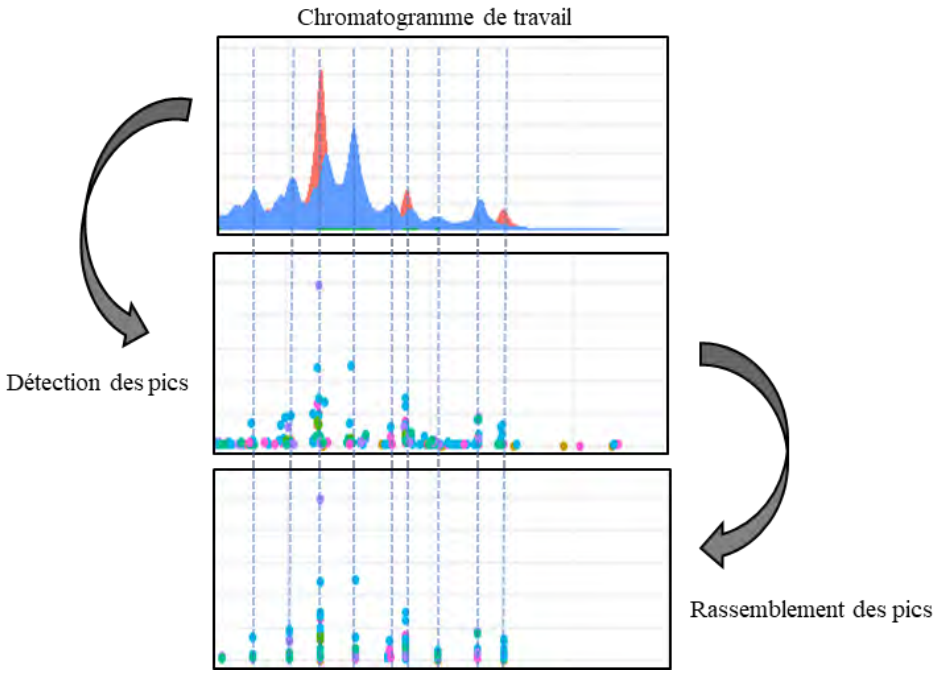


Figure 111. Explication schématique des différentes étapes de détection et regroupement des pics

Ainsi, en appliquant le script, une liste de pics représentatifs du chromatogramme visualisé avec Xcalibur est obtenue. En effet, uniquement les apex des pics observés sont retenus, éliminant ainsi l'ensemble des autres variables caractérisant le même pic dans une fenêtre de 0,1 minute environ.

II.6.2.3 Quantification avec le CAD

En étudiant les profils MS et CAD d'un échantillon, il a été constaté des réponses différentes selon le détecteur. Dans certains cas, le pic CAD est inférieur au pic en masse comme dans le cas n°1, Figure 112, entre les profils $[M-H]^-$ et CAD ; mais l'inverse est également constaté comme dans le cas n°2, Figure 112.

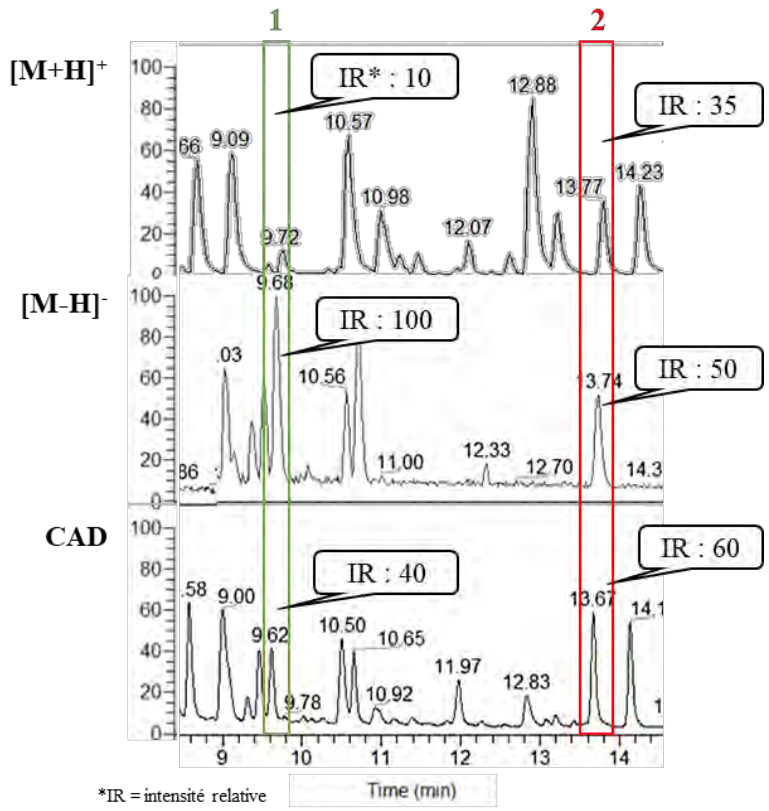
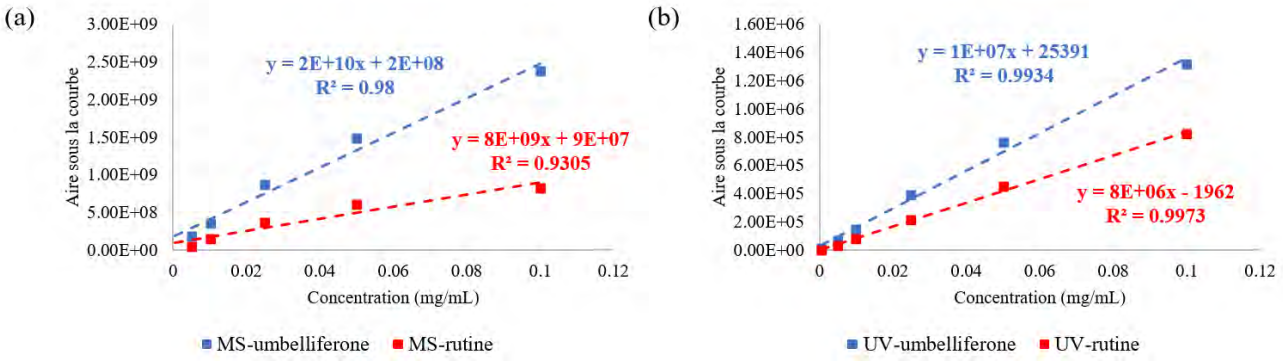


Figure 112 : Variabilité des réponses des détecteurs CAD et MS

Pour obtenir la spéciation semi-quantitative des échantillons analysés, deux références ont été utilisées : l'umbelliferone et la rutine. Les courbes de calibration des trois détecteurs (MS, DAD et CAD) ont été tracées (Figure 113). Il a été constaté que les réponses de la masse (a) et du DAD (b) n'étaient ni parfaitement linéaires ni spécifiques d'un chromophore. En effet, deux équations différentes ont été obtenues pour la rutine et l'umbelliferone. Au contraire, la réponse du CAD (c-d) est universelle puisqu'une même équation linéaire a été obtenue pour les deux références analysées dans les mêmes conditions et aux mêmes concentrations. L'analyse semi-quantitative est donc pertinente avec les données du CAD et ne nécessite l'usage que d'une seule référence, en étalon interne ou externe.



(c)

C (mg/mL)	Aire (umbelliferone)	Aire (rutine)
0,1	1,15	1,18
0,05	0,52	0,57
0,025	0,22	0,26
0,01	0,07	0,07
0,005	0,02	0,02

(d)

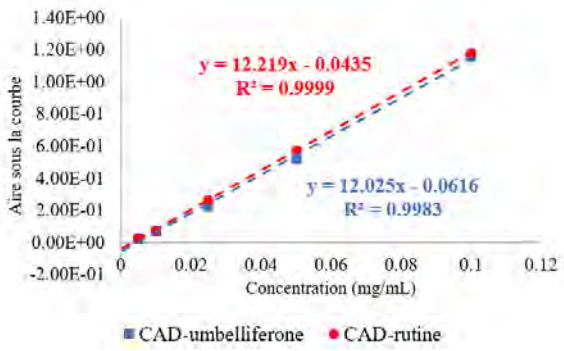


Figure 113. Courbes de calibration des signaux MS (a), des signaux UV (b), résultats d'intégration des signaux CAD (c) et droites de calibration associées (d) de la rutine et de l'umbelliferone

II.6.2.4 Combinaison des données UHPLC-MS et du CAD

La dernière étape consiste à combiner la matrice CAD obtenue par le script « speaq » avec la matrice UHPLC-MS nettoyée à l'aide des annotations obtenues par l'intermédiaire du logiciel MS-DIAL et du script « Mz-unity ». Pour cela, un troisième script R nommé « MS-combine » est utilisé (https://github.com/cran/MScombine) (**Calderón-Santiago et al., 2016**). Cet outil ayant été développé pour combiner les données UHPLC-MS en modes positif et négatif, il a fallu adapter le contenu des matrices et il a été décidé de paramétrer la matrice CAD comme des données positives en indiquant pour l'ensemble des variables une masse arbitraire de 501,007, et la matrice UHPLC-MS comme des données négatives en indiquant une masse de 498,993. Ainsi, entre les deux masses, une différence de 2,016 est observée correspondant à une différence de polarité $[M+H]^+$ et $[M-H]^-$. La liaison entre les variables prendra en compte le temps de rétention, avec une flexibilité fixée à 0,1, mais portera uniquement sur la corrélation de Pearson, c'est-à-dire sur leur tendance évolutive au sein des échantillons. Le résultat obtenu abouti à une liste mettant en regard des variables annotées CAD et MS avec un score de corrélation. Uniquement les corrélations supérieures à 0,5 entre les deux signaux sont conservées et étudiées plus en détail afin de faire correspondre un pic CAD avec un pic UHPLC-MS, et donc une masse précise correspondant à un composé identifié à l'aide des bases de données.

II.6.3 Optimisation du processus sur des standards

L'ensemble du processus d'annotation a tout d'abord été testé sur des mélanges de standards de produits naturels, injectés par pool de 10, de manière à mettre clairement en évidence l'intégralité des phénomènes de dégénérescence évoqués précédemment. La liste des standards est présentée en annexe 9 page 231.

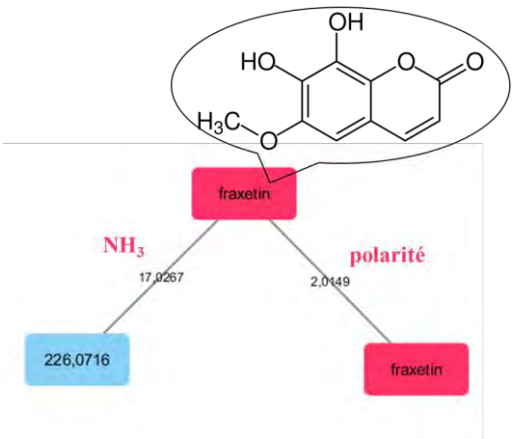
Il a été tout d'abord constaté que la relation de polarité était très marquée suggérant une redondance importante lors de l'analyse en modes NI et PI. C'est la relation primordiale qui permet de déterminer avec certitude l'ion moléculaire du composé. Les exemples du fraxetin de la classe des benzopyranoïdes (Figure 114 a) et du tryptophane, un acide aminé (Figure 114 b) ainsi que de l'acide rosmarinique (Figure 114 c) et de la phloridzine (Figure 114 d) illustrent particulièrement bien l'importance de prioriser cet aspect de polarité. Dans le premier cas, une différence de masse de 17,03 est observée avec l'ion moléculaire positif. Pour le tryptophane, cela est cohérent avec une perte de NH_3, ce composé étant un acide aminé présentant une fonction NH_2 et la masse de l'ion moléculaire étant la plus grande (188⇔**205**). Au contraire, le fraxetin ne présente aucun atome d'azote (N) dans sa structure et cette relation relèverait donc plutôt d'une addition, la masse de l'ion moléculaire étant la plus petite (**209**⇔226). Dans le second cas, une différence de 46 est observée avec l'ion moléculaire négatif. Pour

la phloridzine, cela correspondrait plutôt à la présence d'un adduit formiate puisque la masse de l'ion moléculaire est la plus petite (**435**⇔481**)**, tandis que pour l'acide rosmarinique, cela correspondrait plutôt à une perte de neutre, la masse de l'ion moléculaire étant la plus grande (**359**⇔313). Ainsi, à travers ces deux exemples, on remarque qu'une même différence de masse peut parfois être interprétée de manières différentes et il semble alors difficile de pouvoir déterminer des règles strictes pour l'élimination des redondances. C'est pourquoi la relation de polarité prime et toutes les variables gravitant autour sont éliminées, sans même se préoccuper de la nature du lien.

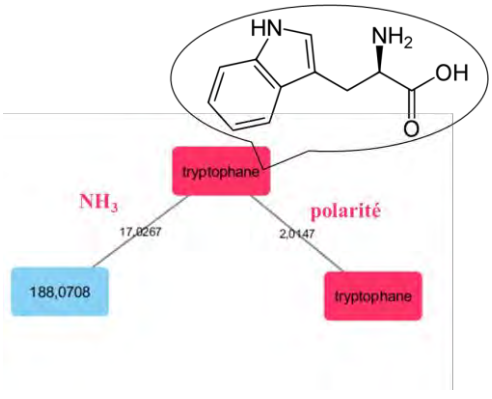
De la même manière, l'adduit sodium est très largement détecté en mode PI. Celui-ci permet de bien identifier l'ion moléculaire positif. La relation de dimère est régulièrement observée et particulièrement lors de l'utilisation du mode NI. En effet, sur les neuf relations dénombrées (donnés non montrées), plus de 75% d'entre elles interviennent en mode NI.

Les alcaloïdes et acides aminés, bien que parfois ionisables dans les deux modes, présentent une ionisation plus riche en mode PI. Ils sont caractérisés par des pertes de neutres et d'ions multichargés. En effet, l'ionisation repose sur un processus de compétition entre les fonctions ionisables de la structure. Globalement, ce sont majoritairement les atomes d'oxygène et d'azote les deux sites ionisables des molécules naturelles, en raison de la présence de leur doublet libre. Pour déterminer les sites favorables, on peut s'intéresser aux valeurs de constante d'acidité (pK_A) traduisant l'acidité ou la basicité des fonctions. Dans les cas énoncés Figure 115, l'on dispose des fonctions alcools (-OH ; 14 < pKa < 18 dans l'eau), amine (-NR2) et acide carboxylique (-COOH ; 0 < pKa < 5 dans l'eau). Dans le cas de la fonction amine, la basicité augmente globalement avec la classe de l'amine en raison de l'effet inductif donneur des substituants alkyles. Cependant, dans le cas des amines tertiaires, l'encombrement stérique entre également en compte et on obtient donc une baisse de la basicité. L'ordre de réactivité est donc R_2NH (~10,7) > RNH_2 (~10,6) > R_3N (9,8) > NH_3 (9,2). Lorsqu'un seul site est prépondérant, la molécule est chargée une seule fois (z = 1) et alors l'ion moléculaire correspondra à la masse exacte de la molécule soustraite de la masse de l'atome d'hydrogène. C'est ce que l'on peut observer pour l'ensemble des flavonoïdes, par exemple. Par ailleurs, si le site ionisable est intégré à une aromaticité, son doublet participe à sa stabilité par sa délocalisation et est donc moins disponible, comme pour un des atomes d'azote du tryptophane intégré au système Π (Figure 115 b). Lorsque plusieurs sites sont équivalents, des ions multichargés sont observés. Ainsi, on comprend bien que dans le cas du tryptophane, la fonction ionisée sera l'amine secondaire conduisant à un ion monochargé. En revanche, dans le cas de l'ajmaline (Figure 114 e), les deux fonctions amines semblent équivalentes, car toutes deux tertiaires, c'est pourquoi ces deux sites s'ioniseront conduisant à un ion multichargé (z = 2).

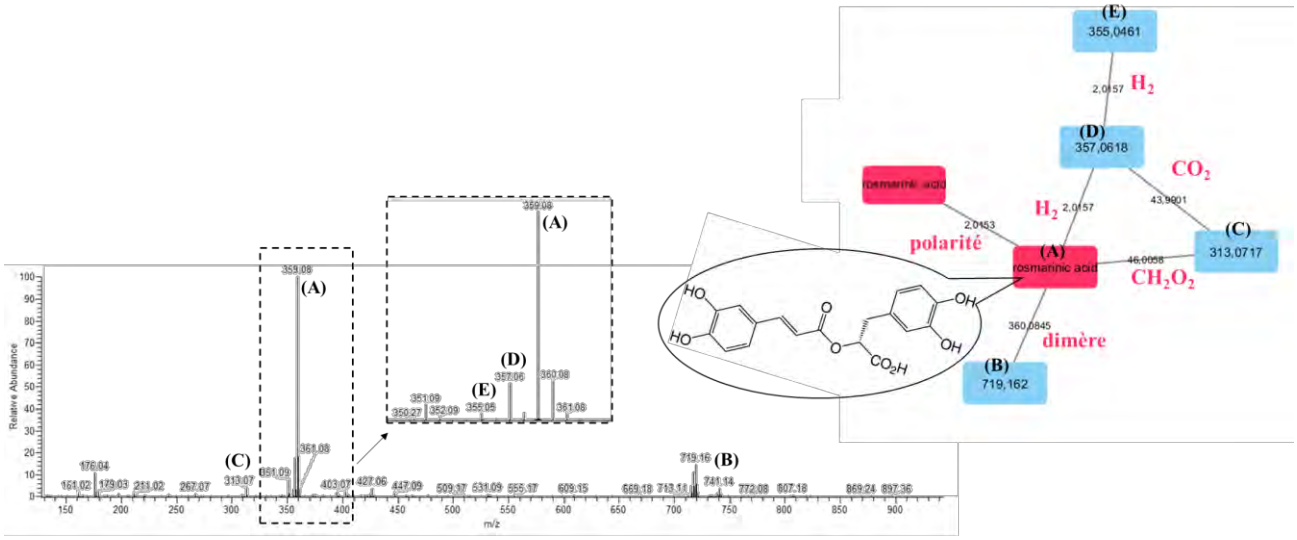
(a) Fraxetin (benzopyranoïde)



(b) Tryptophane (acide-aminé)



(c) Acide rosmarinique (aromatique)



(d) Phloridzine (chalcone)

(e) Adjmaline (alcaloïde)

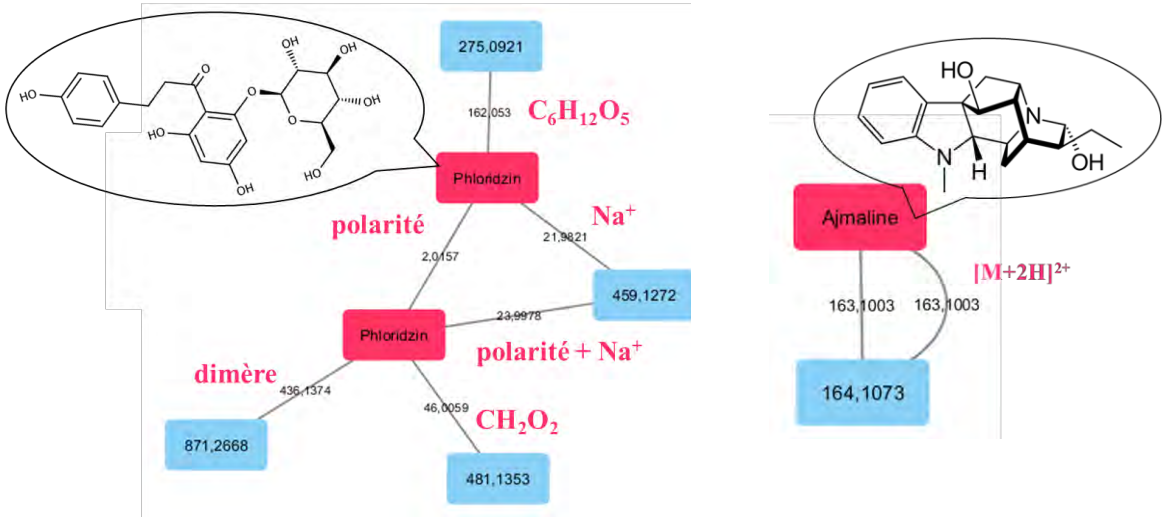


Figure 114. Exemples de dégénérescence sur quelques standards en ESI du fraxetin (a), du tryptophane (b), de l'acide rosmarinique (c), de la phloridzine (d) et de l'ajmaline (e).

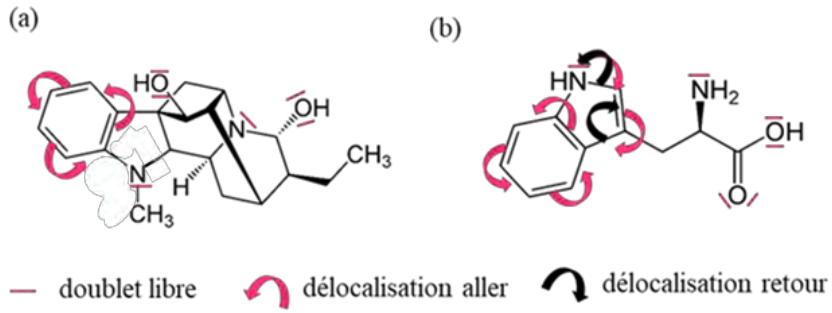


Figure 115. Sites ionisables sur de l'ajmaline (a) et du tryptophane (b) expliquant le phénomène d'ions multichargés

Les données UHPLC-MS ont conduit à la détection respective de 160 et 436 variables en modes NI et PI, conduisant à un total de 596 pics en combinant les données. Après application du script Mz-unity, 517 relations, caractérisées comme résumé dans le Tableau 34, sont obtenues.

Tableau 34. Relations mises en évidence par l'application du script MZ-unity

Relation	Nombre	Relation	Nombre
Isotopes	**190**		
▪ en pos (PI)	136	**Perte neutre PI (nlp)**	**89**
▪ en neg (NI)	54		
Polarité (pol)	**59**	**Perte neutre NI (nln)**	**19**
Adduit (z)	**30**	**Mer (pMer)**	**130**
Na^+	22	Dimère	31
K^+	0	Hétéromère	99
FA	8		

En éliminant les pics redondants ainsi que les pics constitutifs du blanc, la liste initiale est réduite de plus de 40% conduisant à 236 variables. L'utilisation d'une base de données des standards injectés et de bases de données plus généralistes a permis d'identifier 34 composés avec MS-DIAL, en se basant uniquement sur la formule brute, et 33 composés avec MS-FINDER en intégrant les informations de fragmentation. Au total, 39 composés sur les 43 ont été identifiés. Trois d'entre eux, l'umbelliferone, la caféine et la L-proline, ne présentaient pas de MS/MS d'où l'impossibilité de les identifier par MS-FINDER. Un travail manuel de déreplication a été mis en œuvre pour compléter cette liste et a permis de valider les identifications précédentes et d'ajouter deux autres composés : l'arbutine ($[M-H]^-$ = 271,0827) n'a pas été automatiquement identifié en raison de la détection unique de son adduit formiate (mz = 317, 0876) ; la quinine est un isomère de la quinidine et l'identification a été confirmée en travaillant directement à partir de l'injection des mixtes de standards. Les deux composés faisant partie de deux mélanges différents, leurs temps de rétention sur un gradient de 15 min ont été comparés et il a été observé que la quinidine était éluée avant la quinine. Les trois derniers composés, l'acide ellagique, l'hypericine et le phloroglucinol, n'ont pas été détectés en raison de leur intensité qui était inférieure au seuil de détection appliqué pour limiter au maximum le bruit de fond. Ce sont ainsi 40 standards sur les 43 qui ont été détectés et identifiés par déreplications automatique et manuelle.

Les données brutes du CAD obtenues à partir d'Xcalibur ont conduit à une liste de 7500 variables. L'application du script « speaq » a finalement conduit à une liste réduite de 60 variables. La liste nettoyée de 236 pics UHPLC-MS et la liste du CAD de 60 pics ont été combinées par l'intermédiaire du script « MS-combine » et un fichier de 173 relations CAD-MS avec 56 signaux CAD a été obtenu. Plusieurs masses correspondent à un unique signal CAD, c'est pourquoi les identifications préalablement obtenues ont été utilisées pour mettre en évidence les corrélations les plus pertinentes. La mise en regard des données a permis de faire corréler 38 signaux CAD avec des standards. Dans le cas de coélution, les standards sont corrélés au même signal CAD détecté. C'est le cas pour l'arecoline et le glutathion à 0,98 min et pour l'umbelliferone et l'acide rosmarinique à 4,19 min. Ainsi, les 40 standards identifiés ont été corrélés à leurs signaux CAD respectifs, validant ainsi la démarche proposée, malgré les limites évoquées, dont la présence de co-élution qui peut conduire à une perte d'informations. Une bonne résolution des pics et une séparation optimale sont donc les clés pour une bonne corrélation des données qualitatives et semi-quantitatives du détecteur de masse et du détecteur Corona.

II.6.4 Application aux especes de violettes

Une approche d'extraction plus efficiente et permettant de refléter au mieux la physiologie de la matière végétale a été mise au point. Elle repose sur le broyage, l'extraction et la purification quasi

simultanés des feuilles de violette, en respectant la chaîne du froid pour ne pas dégrader d'éventuels composés thermosensibles.

II.6.4.1 Extraction hydroalcoolique exhaustive et rapide

Différentes optimisations ont été effectuées sur la violette de Toulouse pour parvenir à un protocole final appliqué à quelques échantillons de la collection pour une étude semi-quantitative des espèces. De premières purifications faisant usage de la silice et du charbon actif ont été mises en œuvre, mais les résultats n'étaient pas satisfaisants en raison dans le premier cas, d'une élimination incomplète de la chlorophylle, et dans le second cas, d'une forte sensibilité du réactif et d'un manque de spécificité conduisant à l'adsorption de composés.

En effet, deux taches rouges à 366 nm sont observées pour l'échantillon non traité correspondant aux pigments, dont la chlorophylle, et l'une d'elles est toujours constaté pour l'échantillon traité à la silice confirmant la purification partielle suspectée par la couleur toujours verte du surnageant. En revanche, pour les échantillons traités par charbon actif, ces deux taches ne sont plus visibles et la couleur verte a totalement disparu. Ce composé présente donc une meilleure efficacité pour l'élimination de la chlorophylle. Cependant, un ajout excessif de charbon actif engendre l'adsorption de métabolites secondaires dont les coumarines majoritaires dans l'extrait, comme en témoigne la disparition progressive des tâches bleues en bas de plaque avec l'ajout croissant de charbon actif (Figure 116). Visuellement, cette rétention s'est traduite par le passage d'une solution jaune à une solution incolore.

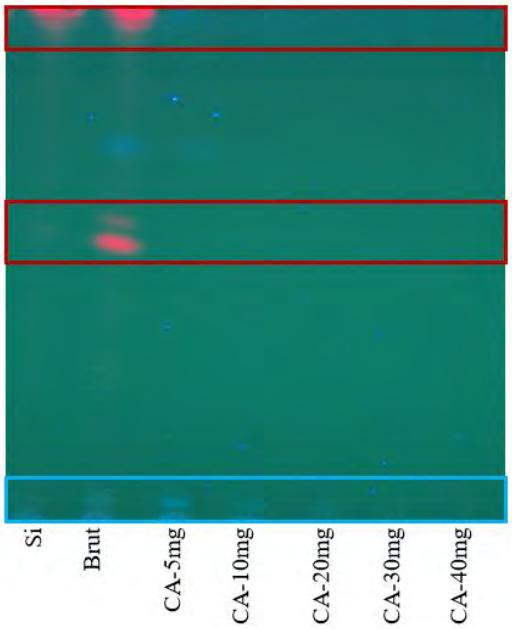


Figure 116. CCM révélée aux UV à 366 nm de différents extraits non traités (brut), traités par silice (Si) ou charbon actif (CA-quantité introduite)

Une méthode biphasique basée sur la méthode de M. Salem et de ses collaborateurs (**Salem et al., 2016**) a été mise en œuvre pour palier à ces problèmes. Celle-ci fait usage de deux mélanges de solvants : MTBE/MeOH et H_2O/MeOH. Une fois les processus de broyage et d'extraction réalisés, on observe deux phases. La première verte comprend principalement les pigments, dont la chlorophylle comme reflété par la couleur vert intense, les composés apolaires et les lipides de hauts temps de rétention. La deuxième phase présente une couleur jaune et est constituée des métabolites primaires et secondaires, polaires à semi-polaires de bas temps de rétention. Il a été noté par ailleurs un dépôt au fond du tube correspondant aux nervures et fibres de la feuille, reflétant la puissance du broyage réalisé (Figure 117).

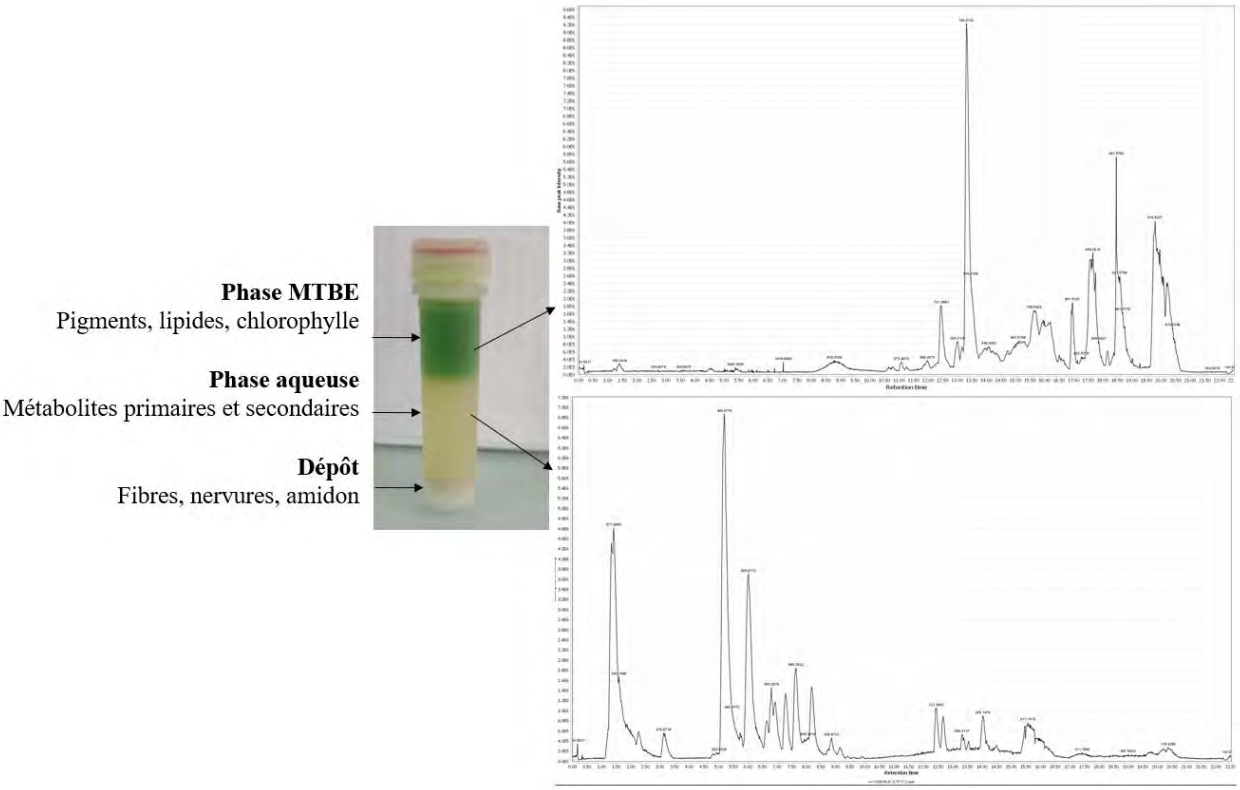


Figure 117. Résultat de l'extraction biphasique et chromatogrammes MS associés aux deux phases

Cette méthode présente le principal avantage d'être rapide et d'assurer l'analyse exhaustive de l'ensemble des constituants de la plante à partir d'une quantité limitée de matériel végétal. En pratique, le processus complet requiert en moyenne vingt minutes ce qui assure le traitement d'un nombre conséquent d'échantillons sur une journée, d'autant plus que les appareils présentent une capacité maximale de 24 échantillons. Toutes les phases aqueuses ont été injectées en UHPLC-CAD-MS et les données ont été traitées sur le modèle précédemment expliqué.

II.6.4.2 Liste de pics UHPLC-MS et annotation des variables

Les données UHPLC-MS ont été traitées en modes PI et NI par MS-DIAL avec une ligne de base respectivement fixée à 1.10^6 et 4.10^5. Une première annotation des pics a été obtenue en recherchant les pertes d'eau, les adduits chlorure, sodium, potassium et formiate, ainsi que les dimères. Les données brutes ont conduit à la détection de 350 pics dont 306 présentant des données MS/MS en mode NI et 179 dont 145 avec MS/MS en mode PI. Il est ainsi constaté que la majorité des composés sont ionisables en mode NI et peu en mode PI.

Après application du script Mz-unity et filtration par temps de rétention, un total de 308 relations caractérisées comme résumé dans le Tableau 35 a été obtenu. Il a été répertorié 75 annotations par le logiciel MS-DIAL.

Tableau 35. Relations mises en évidence par l'application du script MZ-unity

Relation	Nombre (MZ-unity)	Nombre (MS-DIAL)
Polarité	**46**	**-**
Isotopes	**23**	**35**
Adduit (z)	**21**	**19**
▪ Na	4	5
▪ K	0	1
▪ Cl^-	0	-
▪ Fa	17	13
Perte neutre NI (nln)	**141**	**5** (H_2O)
Perte neutre PI (nlp)	**33**	**6** (H_2O)
Mer (pMer)	**44**	**10**
▪ Dimère	3	10
▪ Hétéromère	41	-

En combinant les informations et en appliquant les règles d'annotation précédemment évoquées, les pics redondants ont été éliminés, conduisant à une réduction de 36% de la liste initiale. Une liste réduite de 339 variables est alors obtenue. L'utilisation d'une base de données du genre *Viola* et de bases de données plus généralistes a permis de proposer des identifications potentielles (score > 6) pour respectivement 78 et 250 composés.

II.6.4.3 Liste de pics quantitatifs du CAD

L'application du script « speaq » a permis d'obtenir une liste de 91 variables représentatives du signal CAD enregistré pour un échantillon QC (Figure 118).

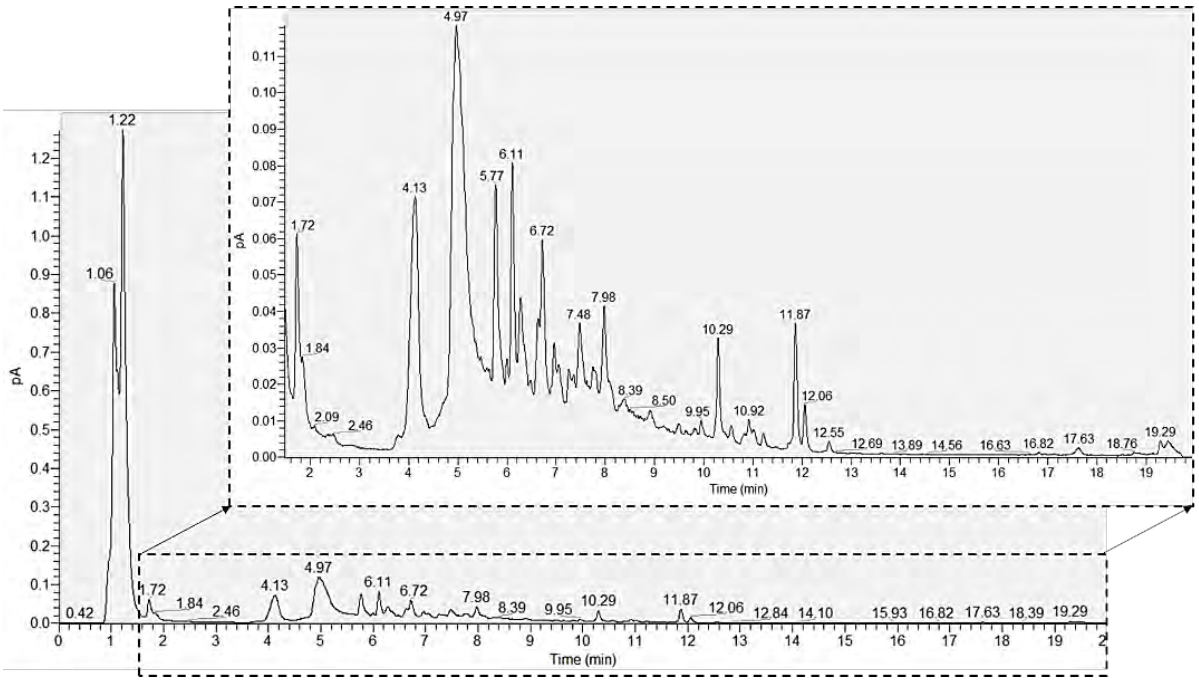


Figure 118. Signal CAD caractéristique d'un échantillon QC

II.6.4.4 Combinaison des données UHPLC-MS et CAD

La liste de 339 pics UHPLC-MS et celle de 91 pics du CAD ont été combinées par l'intermédiaire du script « MS-combine » et un fichier de 126 relations CAD-MS avec un score de corrélation supérieur à 0,5, a été obtenu. Comme plusieurs masses correspondent parfois à un signal CAD, le score de corrélation et les identifications préalablement obtenues ont été utilisées pour mettre en évidence les relations les plus pertinentes. Finalement, la combinaison des données a permis de faire corréler 29 signaux CAD avec 32 pics UHPLC-MS avec une identification (Tableau 36) et 3 non identifiés. En effet, il a parfois été constaté que pour un même pic CAD, différentes masses été corrélées selon l'espèce étudiée, multipliant les candidats pour un même pic CAD.

De manière générale, trois pics ont été observés en début de gradient (Figure 119), entre 0,9 et 1,3 min, correspondant à des composés très polaires, probablement de la famille des acides aminés et sucres type melibiose ($C_{12}H_{22}O_{11}$). La colonne utilisée ne permettant pas de bien les séparer, aucune correspondance CAD-MS n'a pu être clairement définie. Cependant, puisque ces trois pics sont récurrents dans tous les extraits et non négligeables en terme quantitatif, ils ont été retenus. Il est intéressant de noter que leurs concentrations varient quelque peu selon les espèces. Les *V. sp* Hearn en sont particulièrement riches puisqu'ils représentent près de 80% des échantillons contre en moyenne 55 à 60% pour les autres espèces. Ainsi, une analyse plus approfondie de cette composition, par une méthode analytique orthogonale de type HILIC (hydrophilic liquid chromatography) par exemple, pourrait permettre d'apporter des compléments d'information et des critères de différenciation supplémentaires entre ces groupes de violettes.

Les autres principaux composés quantitatifs déterminés dans cette étude appartiennent à la famille des polyphénols, principalement des flavonoïdes et des coumarines, ainsi que des cyclotides (Figure 119)

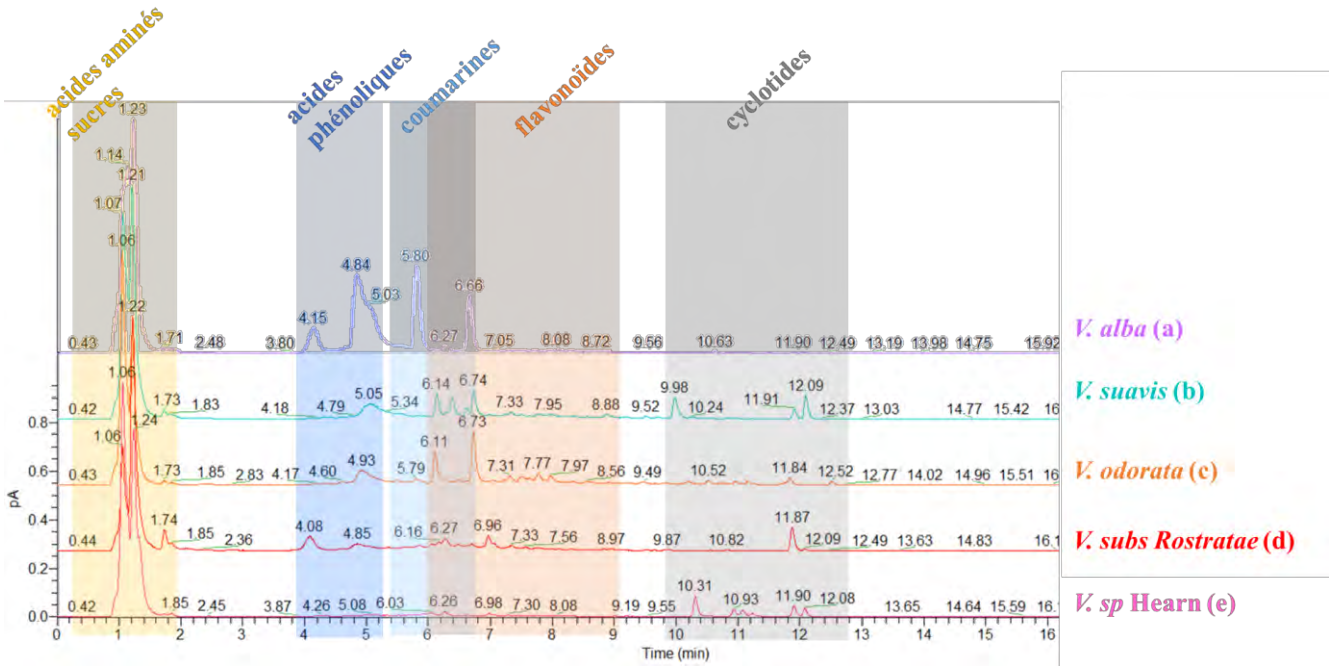


Figure 119. Profils CAD caractéristiques d'une violette *V. alba* (a), *V. suavis* (b), *V. odorata* (c), *V. subs Rostratae* (d) et *V. sp* Hearn (e)

Les coumarines sont particulièrement concentrées chez les *V. alba* et représentent près de 20% des extraits bruts, contre en moyenne moins de 10% pour les autres espèces (Figure 120, Coumarines). En

étudiant plus en détail cette composition, il a été constaté que les *V. alba* étaient les plus riches en coumarines 1 et 2 à hauteur totale de 95%. Les *V. suavis* et *V. odorata* sont plus riches en coumarine 46 à hauteurs respectives de 65 et 75%, tandis que pour les *V. subs Rostratae* et *V. sp* Hearn la coumarine 47 est majoritairement représentée à hauteurs respectives de 85 et 76%. Les coumarines en général ne sont donc pas spécifiques d'une seule espèce, mais la quasi-spécificité des coumarines 1 et 2 chez *V. alba* évoquée lors des études précédentes est ici confirmée.

Les flavonoïdes sont majoritaires chez *V. subs Rostratae* et *V. odorata* à hauteurs respectives de 15 et 13% de l'extrait brut, contre moins de 6% chez les autres espèces (Figure 120, Flavonoïdes). Une étude plus détaillée a mis en évidence des compositions variées. Chez les *V. suavis* et *V. odorata*, le kaempferol 17 représente près de la moitié de la teneur en flavonoïde, tandis que pour les *V. sp* Hearn, le kaempferol 52 est majoritairement exprimé. Par ailleurs, les flavones 25 et 4 caractérisent respectivement un tiers des flavonoïdes de *V. odorata* et *V. sp* Hearn. Les compositions des *V. alba* et *V. subs Rostratae* sont moins hétéroclites et le flavone 21 représente près d'un quart de leurs flavonoïdes. L'isoflavone 54 représente un tiers des flavonoïdes chez *V. alba* et le flavone 5 et le kaempferol 49 près d'un cinquième pour *V. subs Rostratae*. Par ailleurs, le flavonol 24 est essentiellement identifié chez *V. subs Rostratae*. Les flavonoïdes sont donc des composés ubiquitaires, mais sont présents en quantité variable selon les espèces. Comme évoqué précédemment, la capacité antioxydante n'est pas autant reflété par la concentration de ces composés dans les extraits que pas leur composition intrinsèque.

Les cyclotides sont prépondérants chez *V. sp* Hearn et représentent près de 15% des extraits bruts. Pour les autres espèces, excepté les *V. alba* caractérisées par une teneur de moins de 0,5%, ils sont exprimés à hauteur de 8 à 11% (Figure 120, Cyclotides). Une étude plus fine a permis de constater que la cycloviolacine O12 (35) est clairement majoritaire chez les *V. subs Rostratae*, atteignant près de 90% de la concentration totale en cyclotides. Les cycloviolacines O22 (43) et O12 (35) sont dominantes chez *V. suavis* représentant respectivement près de 65 et 15% de la teneur globale. Chez *V. odorata* ce sont les cycloviolacines O24 (44) et O12 (35) à hauteurs respectives de près de 65 et 25% de la teneur totale. Pour les *V.sp* Hearn, la composition en cyclotides est plus variée et six cycloviolacines (14, 26, 31, 35, 38 et 43) représentent plus de 5% de la teneur totale, dont 20% pour la cycloviolacine O4 (28) et près de 40% pour le cyclotide 14. Ainsi, les cyclotides sont des composés retrouvés chez toutes les espèces, mais une plus grande diversité a été constatée pour *V. alba* malgré sa faible richesse, et *V. sp* Hearn.

Le tannin 55 représente moins de 2% de l'extrait brut chez chaque espèce. Il est majoritaire chez *V. subs Rostratae* atteignant 1,8% contre 0,2% pour *V. alba* et *V. suavis*. Au contraire, les acides phénoliques 56 et 57 sont majoritairement représentés, à hauteur de 23% et 16%, chez *V. alba* et *V. suavis* contre 8 et 16% pour *V. odorata* et *V. subs Rostratae*. En revanche, ils sont très peu concentrés chez *V. sp* Hearn.

Les résultats semi-quantitatifs des pics majoritaires des profils CAD de chaque espèce a ainsi permis de mettre en évidence des points de divergences entre les espèces, non plus seulement en termes qualitatifs, mais également quantitatifs. Ces résultats sont en accord avec les précédentes études biologiques, notamment au niveau de l'activité antioxydante puisque les composés identifiés comme actifs apparaissent clairement prépondérants chez les *V. alba* (coumarines 1 et 2) et les *V. subs Rostratae* (flavonoïdes 5, 21, 24). Les activités phytosanitaires, principalement déterminées chez *V. odorata* et *V. sp* Hearn et portant sur l'hypothèse forte de l'activité des cyclotides, pourraient être liées à la cycloviolacine O24 (44) en forte concentration chez *V. odorata* ainsi qu'aux cyclotides 14, 28 et 31 spécifiques ou majoritaires chez *V. sp* Hearn.

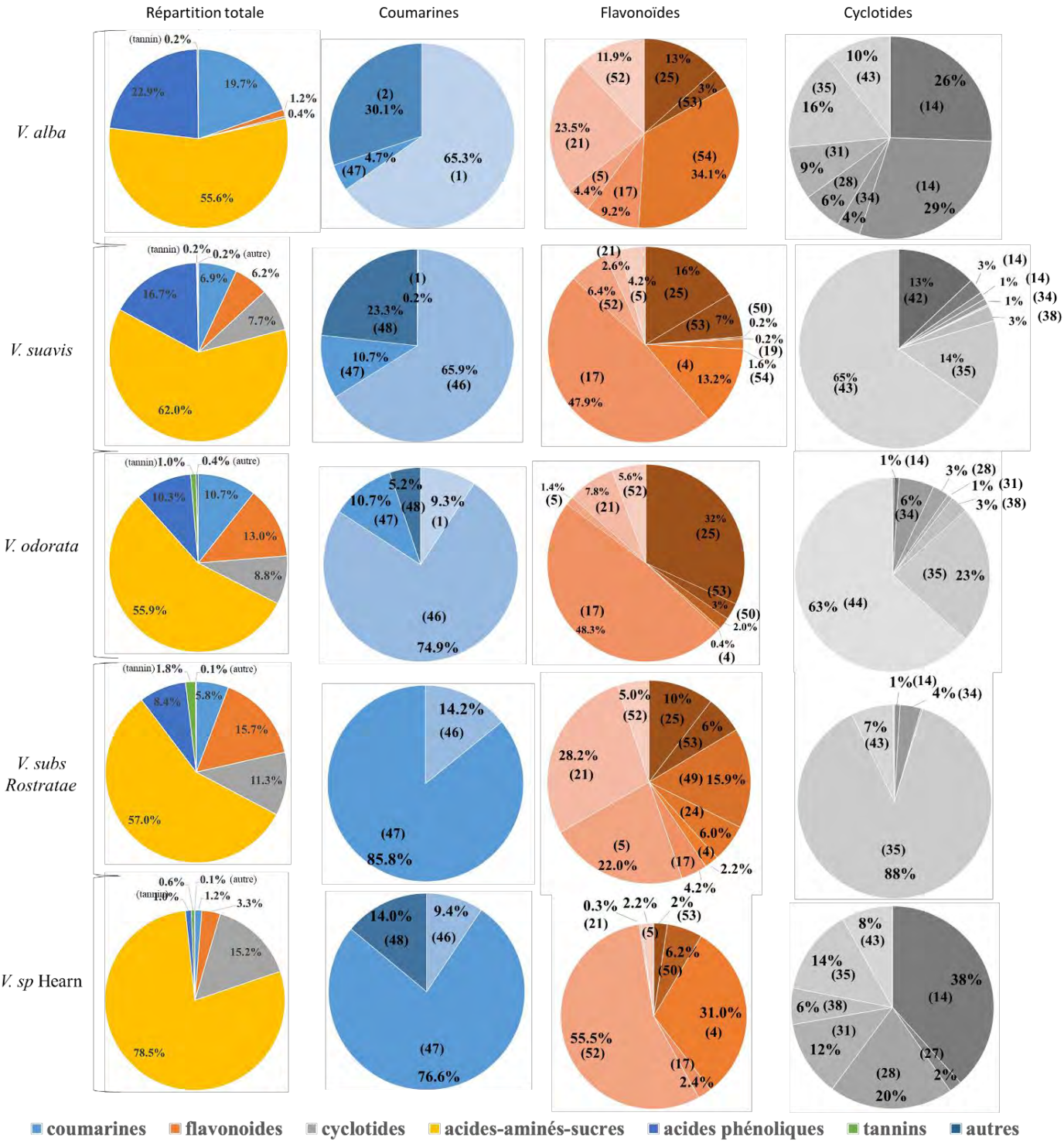


Figure 120. Teneurs en composés identifiés à partir des analyses semi-quantitatives par CAD-MS des parties aériennes des groupes génétiques de la collection

Tableau 36 : Identification des pics UHPLC corrélés aux pics CAD

N°	$[M-H]^-$ $[M+H]^+$	RT (min)	Formule brute	Corrélation	Identification potentielle*	Classe chimique
1	385,0767	5,78	$C_{16}H_{18}O_{11}$	0,84	7-Methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside[a,b]	Coumarine
2	385,0768	6,63	$C_{16}H_{18}O_{11}$	0,82	7-Methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[a,b]	Coumarine
4	755,2028	7,5	$C_{33}H_{40}O_{20}$	0,97	3,3',4',5,5',7-Hexahydroxyflavone; 3-*O*-[α -L-Rhamnopyranosyl-(1,2)-α -L-rhamnopyranoside], 3'-*O*-α -L-rhamnopyranoside[a]	Flavonoïde
5	593,1498	7,06	$C_{27}H_{30}O_{15}$	0,82	6-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxy-8-*C*-α-L-rhamnopyranosylflavone[a]	Flavonoïde
21	563,139	7,57	$C_{26}H_{28}O_{14}$	0,70	8-*C*-β-D-Glucopyranosyl-4',5,7-trihydroxyflavone; 6"-*O*-α-L-Arabinopyranosyl[a]	Flavonoïde
19	533,1288	7,98	$C_{25}H_{26}O_{13}$	0,82	6,8-Di-C-α-L-arabinopyranosylapigenina	Flavonoïde
24	609,1446	8,18	$C_{27}H_{30}O_{16}$	0,77	rutine[a]	Flavonoïde
25	577,1548	8,04	$C_{27}H_{30}O_{14}$	0,82	8-*C*-β-D-glucopyranosyl-4',5,7-trihydroxyflavone; 2"-*O*-α-L-rhamnopyranosyl[a]	Flavonoïde
14	1047,1316	10,4	$C_{133}H_{207}N_{37}O_{39}S_6$	0,97	cycloviolacine O9 / vignopeptide[a]	Cyclotide
14	1047,1318	10,66	$C_{133}H_{207}N_{37}O_{39}S_6$	0,50	cycloviolacine O9 / vignopeptide[a]	Cyclotide
34	1076,1498		$C_{138}H_{216}N_{36}O_{41}S_6$		cycloviolacine O8	Cyclotide
27	1041,8	10,56	$C_{133}H_{207}N_{37}O_{38}S_6$	0,66	cycloviolacine O13[a]	Cyclotide
28	1051,8031	10,95	$C_{134}H_{209}N_{37}O_{39}S_6$	0,65	cycloviolacine O4[a]	Cyclotide
31	1029,4644	11,15	$C_{132}H_{208}N_{34}O_{39}S_6$	0,96	cycloviolacine O18[a]	Cyclotide
35	1446,0808	11,97	$C_{117}H_{179}N_{35}O_{39}S_6$	0,92	cycloviolacine O12[a]	Cyclotide
38	1044,1392	11,29	$C_{132}H_{208}N_{36}O_{40}S_6$	0,67	cycloviolacine H1[a]	Cyclotide
42	1017,1013	10,06	$C_{131}H_{197}N_{33}O_{39}S_6$	0,82	cycloviolacine O16[a]	Cyclotide
43	1453,0886	12,17	$C_{118}H_{181}N_{35}O_{39}S_6$	0,58	cycloviolacine O22[a]	Cyclotide

44	1016,3838	12,64	$C_{126}H_{179}N_{35}O_{42}S_6$	0,89	cycloviolacine O24[a]	Cyclotide
45	755,2029	6,06	$C_{33}H_{40}O_{20}$	0,60	8-*C*-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxyflavone; 2"-*O*-[α-L-Rhamnopyranosyl-(1→6)-β-D-glucopyranosyl][a]	Flavonoïde
46	339,0714	6,12	$C_{15}H_{16}O_9$	0,77	6,7-Dihydroxy-2H-1-benzopyran-2-one; 6-*O*-β-D-Glucopyranoside[a]	Coumarine
47	355,0663	6,28	$C_{15}H_{16}O_{10}$	0,91	6,7,8-Trihydroxy-2H-1-benzopyran-2-one; 7-*O*-β-D-Glucopyranoside[a]	Coumarine
48	369,0819	6,63	$C_{16}H_{18}O_{10}$	0,82	6,7,8-Trihydroxy-2H-1-benzopyran-2-one; 6-Me ether, 8-*O*-β-D-glucopyranoside[a]	Coumarine
49	901,2601	6,77	$C_{39}H_{50}O_{24}$	0,82	Kaempferol 3,7-diglycosides; 3-*O*-[β-D-Glucopyranosyl-(1→2)-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside], 7-*O*-α-L-rhamnopyranoside[a]	Flavonoïde
17	739,2076	6,78	$C_{33}H_{40}O_{19}$	0,72	8-*C*-β-D-Glucopyranosyl-4',5,7-trihydroxyflavone; 2"-*O*-α-L-Rhamnopyranosyl, 4'-*O*-β-D-glucopyranoside[a]	Flavonoïde
50	463,087	7,73	$C_{21}H_{20}O_{12}$	0,84	3',4',5,5',7-Pentahydroxyflavone; 3'-*O*-β-D-Glucopyranoside[a]	Flavonoïde
51	919,2488	7,83	$C_{42}H_{48}O_{23}$	0,90	8-*C*-β-D-Glucopyranosyl-4',5,7-trihydroxyflavone; 6"-*O*-(3,4-Dimethoxybenzoyl), 2"-*O*-[β-D-glucopyranosyl-(1→6)-β-D-galactopyranosyl][a]	Flavonoïde
52	593,1499	8,45	$C_{27}H_{30}O_{15}$	0,57	Kaempferol 3-glycosides; 3-*O*-[β-D-Glucopyranosyl-(1→4)-α-L-rhamnopyranoside][a]	Flavonoïde
53	577,1546	8,88	$C_{27}H_{30}O_{14}$	0,73	4',5,7-Trihydroxyflavone; 5-*O*-[α-L-Rhamnopyranosyl-(1→4)-β-D-glucopyranoside] [a]	Flavonoïde
54	643,2224	8,18	$C_{36}H_{36}O_{11}$	0,77	4'-*O*-Methylneobavaisoflavone 7-*O*-(2"-p-coumaroylglucoside[c]	Flavonoïde
55	919,2488	7,83	$C_{42}H_{48}O_{23}$	0,90	8-*C*-β-D-Glucopyranosyl-4',5,7-trihydroxyflavone; 6"-*O*-(3,4-Dimethoxybenzoyl), 2"-*O*-[β-D-glucopyranosyl-(1→6)-β-D-galactopyranosyl]	Tannin
56	315,0716	4,94	$C_{13}H_{16}O_9$		2,5-Dihydroxybenzoic acid; 2-*O*-β-D-Glucopyranoside[a]	Acide phénolique
57	243,0504	4,19	$C_{10}H_{12}O_7$		1-*O*-Galloylglycerol[c]	Acide phénolique

[a] Déterminés par fragmentation *in silico* avec la base de données *Viola* ; [b] Confirmés par RMN *;* [c] Déterminés par fragmentation *in silico* avec les bases de données MS-FINDER

II.6.5 Conclusion

L'étude semi-quantitative des espèces de violettes précédemment identifiées a été entreprise en combinant les données du détecteur Corona et de la masse. Le détecteur Corona présente une réponse universelle indépendante du squelette chimique (chromophores, sites ionisables, …) et est ainsi adapté à l'étude comparative quantitative des extraits. Une des principales contraintes pour la stratégie proposée émane du phénomène de dégénérescence, conduisant à la formation d'un nombre conséquent de pics redondants représentatif d'un unique ion moléculaire. Une méthode d'élimination de ces pics a ainsi été proposée afin de réduire la liste de masse bruitée et parvenir à des corrélations pertinentes.

L'application aux violettes de la méthode de traitement optimisée sur des standards a permis de rendre compte de la variabilité quantitative de leur composition. Les cinq espèces présentent toutes une concentration en sucres et acides aminés (pics d'injection) très forte, atteignant près de 80% de l'extrait brut chez *V. sp* Hearn. Ces dernières sont les plus riches en cyclotides, en particulier les cycloviolacines suivi des *V. subs Rostratae* dont la composition reste néanmoins moins hétéroclite puisqu'une unique cycloviolacine représente près de 90% de la concentration totale. Les *V. subs Rostratae* sont les plus riches en flavonoïdes dont la composition est très variée avec l'identification de neuf flavonoïdes. Les *V. odorata* sont également concentrés en ces polyphénols en particulier deux dérivés de flavone (25) et de kaempferol (17) représentant près de 80% de la teneur totale. Les *V. alba* sont les plus pourvues en coumarines en particulier deux isomères quasi-spécifiques représentant 95% de la teneur totale. Deux autres coumarines ont été majoritairement identifiées à hauteur en moyenne de 70% chez *V. suavis* et *V. odorata* (46) et 80% chez *V. subs Rostratae* et *V. sp* Hearn (47).

La méthode mise en œuvre a ainsi permis d'obtenir une spéciation semi-quantitative des pics majoritaires observés en CAD avec une indexation par l'intermédiaire de la fragmentation. Néanmoins, ils représentent entre 65 et 75% du signal CAD total. Ce résultat pourrait être amélioré en optimisant la séparation et la résolution des pics, en limitant la taille des raccordements, permettant d'obtenir une meilleure corrélation entre un seul pic en masse et son équivalent en CAD.

PARTIE III
CONCLUSION GENERALE

L'étude phytochimique des violettes réalisée au cours de ce travail avait pour principal objectif le criblage d'activités biologiques, en particulier antioxydantes, pour une valorisation industrielle d'un extrait à visée cosmétique.

L'analyse d'une collection nationale mal caractérisée au niveau botanique a tout d'abord conduit à un travail de classification et d'identification des violettes à disposition. En effet, l'intérêt premier portant sur la violette de Toulouse, en raison de son image de marque au sein de la région Occitanie, l'éventualité de la découverte d'une autre violette aux vertus plus intéressantes n'a pas été écartée.
Le couplage des méthodes basées sur la botanique, la génétique et les composés phytochimiques, aussi bien des fleurs que des feuilles, a permis de classer 96% de la collection et d'attribuer des noms d'espèce aux différentes violettes. Il a ainsi été mis en évidence qu'à partir de l'identification génétique de quelques violettes, le seul profilage chimique des métabolites secondaires non volatils des feuilles permet d'indexer les plants par projection à l'aide d'outils statistiques. Ces résultats reflètent donc une bonne complémentarité des deux approches, meilleure qu'avec l'aspect volatil dont la sensibilité, la variabilité et la disponibilité ne permettent pas l'acquisition rapide de données à grande échelle. Néanmoins, la combinaison de l'ensemble des résultats a permis d'établir la présence de profils chimiques volatils des fleurs et non volatils des feuilles spécifiques pour une espèce, avec l'identification de biomarqueurs chimiques. Ces derniers sont globalement des dérivés d'ionones et des terpènes pour les composés volatils, et des polyphénols types coumarines et flavonoïdes ainsi que des cyclotides pour les non volatils. L'apport de données botaniques a permis d'étayer les identifications faites et d'apporter des éléments de différenciation complémentaires basés sur les aspects morphologiques et anatomiques des organes floraux et foliaires des violettes.
La limite de la méthode employée émane de la non-identification des variétés ou sous-espèces, suspectées au sein des plants. Néanmoins, la dénomination des espèces était suffisante pour cette étude, mais des analyses complémentaires sont à envisager dans le cadre de l'approfondissement des génotypes de la collection en séquençant quelques régions de l'ADN chloroplastique (cpDNA) ou le clonage des produits PCR pour étudier plus en détails les hybrides potentiels mis en évidence par la présence d'un polymorphisme important des sites nucléotidiques de la région des ITS. Il serait également intéressant de récolter des espèces de jardin botanique identifiées afin de compléter les séquences références de la base de données utilisée puisque toutes les violettes connues n'y sont pas référencées, ce qui explique que certaines violettes de la collection n'aient pas pu être identifiées. Par ailleurs, d'autres profils chimiques auraient pu être étudiés, comme par exemple les profils de concrètes qui auraient pu apportés des critères de différenciation supplémentaires.

La diversité des profils chimiques des violettes tend à conforter la variété de leur potentiel biologique. Deux catégories d'activité ont été étudiées : l'activité antioxydante largement exploitée aujourd'hui en cosmétique, et l'activité phytosanitaire portant sur des aspects antifongique et elliciteur de défense. En effet, l'approche phytosanitaire s'axe aujourd'hui principalement sur le développement de produits naturels présentant une activité antimicrobienne directement sur les pathogènes, ou stimulant le système immunitaire de la plante en déclenchant des mécanismes de défense.

En ce qui concerne le potentiel phytosanitaire, le criblage de la collection a permis de mettre en évidence deux principales espèces d'intérêt, les *V. sp* Hearn et les *V. odorata*. Les premières violettes sont celles qui ont présenté l'activité antifongique la plus large avec une action sur quatre des cinq souches testées, dont une pseudo-spécificité sur l'une d'elle, la souche *Fusarium graminearum*. Les secondes ont présenté une activité significative sur seulement trois des souches. Les analyses métabolomiques entreprises n'ont pas permis d'établir des corrélations structures-activités satisfaisantes en raison de la diversité de réponse entre espèces et au sein des espèces. Néanmoins, l'étude plus détaillée de la réponse de trois *V. sp* Hearn sur *Fusarium* s'est avérée concevable et les principaux

composés révélés par les analyses statistiques et les données UHPLC-HRMS comme présentant le plus fort potentiel antifongique ont été les cyclotides, et plus particulièrement les cycloviolacines. Un approfondissement de ces résultats par fractionnement d'extraits de violettes issus de *V. sp* Hearn et *V. odorata* a alors permis d'établir que ce sont les fractions enrichies en cyclotides qui présentent un effet antifongique, validant ainsi les hypothèses statistiques en terme de famille chimique active. Ces deux ensembles de violettes présentent par ailleurs les réserves en cyclotides les plus importantes, résultat cohérent avec les conclusions établies. Cependant, les résultats d'inhibition n'étant pas aussi significatifs qu'attendu, ces tests ont mis en lumière l'impact éventuel des mucilages, sources de carbone, participant au développement des champignons. La présence de ces polysaccharides suspectée par l'aspect très visqueux des solutions, ayant d'ailleurs compliqué leur fractionnement, est due à l'utilisation d'un solvant aqueux. Ainsi, suivant la balance source de carbone/cyclotides, les résultats pourraient être sous-estimés. L'usage d'un autre protocole faisant intervenir uniquement des solvants organiques pourrait donc être proposé pour l'obtention de réponses plus franches.

Ces deux ensembles, *V. odorata* et *V. sp* Hearn se sont également démarqués dans l'étude de l'activité inductrice de défense de la plante, représentant respectivement 38% et 23% des violettes révélées les plus actives. Différentes études métabolomiques ont été menées sur les résultats biologiques obtenus et aucune relation espèce-activité n'a pu être clairement mise en évidence. En effet, certaines violettes des autres espèces ont engendré une réponde de défense positive et des disparités de réponses biologiques ont été observées au sein d'une même espèce. Or, il a été démontré des différences chimiques claires entre les différentes espèces, ce qui ne permet donc pas de conclure clairement sur les composés responsables de cette activité. Des études de fractionnement complémentaires seraient donc à envisager pour cibler le type de composés responsables. Une hypothèse est posée sur l'action des cyclotides particulièrement riches chez les deux ensembles de violettes les plus actives. Il serait alors nécessaire de procéder à l'enrichissement allant jusqu'à la purification des cyclotides pour valider l'hytpothèse portée sur leur rôle dans ces activités phytosanitaires.

Le potentiel antioxydant, aspect lié à une valorisation industrielle, a tout d'abord été évalué en détail sur la violette de Toulouse. Cette étude a conduit à l'identification, par UHPLC-HRMS et à l'aide de modèles statistiques, des classes chimiques impliquées dans cette activité et appartenant aux polyphénols : les coumarines-*O*-glycosylées et les flavonoïdes-*C*-glycosylés. La majorité des structures a été identifiée par fragmentation *in silico* à l'aide de bases de données généralistes ou préférentiellement basées sur le genre *Viola*. Trois composés de la classe des coumarines ont néanmoins été structurellement déterminés par RMN 1D et 2D et deux inconnus des bases de données ont pu être *de novo* dérépliqués à partir d'un réseau moléculaire. Suite à une validation de l'activité par TLC-MS, ce sont finalement quatre coumarines et trois flavonoïdes dont seulement deux connus dans le genre *Viola*, qui ont été révélés antioxydants.

La comparaison du potentiel antioxydant des violettes de la collection a été basée sur l'extraction hydroalcoolique avec de l'éthanol 80%. Deux principaux groupes de violettes se sont démarqués, les *V. alba* dont fait partie la violette de Toulouse, et la sous-section des *V. subs Rostratae* à laquelle *V. labradorica*, *V. grypoceras* et *V. mirabilis* appartiennent qui ont présenté un potentiel antioxydant supérieur à la violette de Toulouse. En accord avec les résultats de cette dernière, ce sont des coumarines et des flavonoïdes glycosylés qui ont été mis en relief par les analyses statistiques. Il est également ressorti de cette étude et des différents dosages effectués, que les coumarines sont des composés plutôt spécifiques des *V. alba* présentant une teneur de l'ordre de 0,05% dans la poudre, environ au moins cinq fois plus que les autres violettes de la collection. A l'inverse, les flavonoïdes sont retrouvés chez toutes les violettes en proportion variées et en moyenne, ce sont les *V. subs Rostratae* les plus concentrées avec une teneur de l'ordre de 1,5 à 1,8% dans la poudre. Il s'est avéré que la teneur en flavonoïdes n'était pas

directement liée au potentiel antioxydant puisque trois violettes ont démontré une teneur plus forte que les *V. subs Rostratae*, malgré leur faible potentiel antioxydant. Ces résultats sont clairement à mettre en relation avec la structure des flavonoïdes dont trois groupes structuraux particuliers confèrent ce potentiel, notamment le noyau catéchol relié à la structure flavonoïde.

Dans ces travaux, l'activité s'est essentiellement centrée sur le potentiel antiradicalaire en particulier, c'est-à-dire sur l'action de nos extraits sur des radicaux libres, un des aspects de l'activité antioxydante. Cependant, ce potentiel antioxydant englobe également d'autres aspects notamment en lien avec des intéraction au niveau des voies de singalisation rédox comme l'inhibition d'enzymes oxydantes ou encore l'action sur des enzymes antiralaires. Pour étudier plus en profondeur ce potentiel antioxydant des violettes, il serait donc indispensable d'effectuer des tests sur cellules pour vérifier l'effet des extraits sur des cellules soumis à un stress oxydant, révélant ainsi la biodisponibilité des molécules non retransmis en tests *in vitro* (**López-alarcón et Denicola, 2013**).

Une spéciation chimique semi-quantitative des espèces de la collection a été entreprise. Pour cela, un nouveau protocole adapté à l'analyse de masse a été proposé. Il propose le broyage, l'extraction et la purification en chlorophylle quasi simultanés. L'optimisation sur des standards a permis de mettre au point la méthode de traitement tournée vers le nettoyage de la liste de pics, étoffée par la dégénérescence inhérente à l'ESI. L'application aux espèces de violettes a permis de mettre en évidence des différenciations quantitatives entre les espèces. Il a ainsi été constaté que les *V. alba* étaient les plus riches en coumarines, les *V. subs Rostratae* en flavonoïdes et les *V. sp* Hearn en cyclotides. Des études plus détaillées ont permis de rendre compte de la variabilité de la composition en ces composés plus ou moins hétéroclites selon les espèces. Par ailleurs, ces résultats semi-quantitatifs confirment le fort potentiel antioxydant des *V. alba* et *V. subs Rostratae* puisqu'elles sont les plus riches en composés actifs et apportent des informations additionnelles sur les éventuels cyclotides qui pourraient être responsables de l'activité antifongique et inducteur de défense chez la plante. Des essais complémentaires sont cependant nécessaires, notamment par l'isolement de ces cyclotides pour tester les produits purs et vérifier l'éventuelle spécificité de certains. Cette spéciation semi-quantitative pourrait être améliorée en optimisant la séparation chromatographique et en proposant également une quantification absolue par l'usage d'un étalon interne.

Dans l'optique d'une valorisation industrielle d'un extrait enrichi en composés antioxydants, quatre méthodes d'extraction ont été comparées : l'électroporation, les micro-ondes, le CO_2 supercritique et l'extraction hydroalcoolique. Le point commun à toutes ces méthodes émane de leur respect des préceptes de la chimie verte. Les études du rendement d'extraction et des teneurs en polyphénols par dosages spécifiques des coumarines et flavonoïdes par UHPLC-UV, des polyphénols totaux par la méthode Folin-Ciocalteu et des flavonoïdes par la méthode $AlCl_3$ ont conduit à sélectionner l'extraction hydroalcoolique qui assure un meilleur rendement et une extraction significative des composés d'intérêt. Son principal inconvénient reste la coloration verte de l'extrait issue de la chlorophylle, caractéristique partagée par toutes les méthodes excepté les micro-ondes qui conduisent à l'obtention d'un extrait potentiellement directement formulable, mais nettement moins enrichi.

Un transfert pilote est actuellement en cours chez le sous-traitant Berkem sur les saisons 2015 et 2016 de la violette de Toulouse et la saison 2015 de la violette de Tourette. La comparaison avec les essais du laboratoire a reflété des caractéristiques équivalentes à savoir, une teneur en polyphénols supérieure pour les violettes de Toulouse, environ deux fois plus concentrées que la violette de Tourette et l'obtention d'une pâte marron collante après concentration à sec à température modérée ne dépassant pas 50 °C. Aucune différence de mise en œuvre n'a été constatée entre les deux espèces de violettes, mais leurs caractéristiques chimiques et biologiques varient avec une meilleure activité antioxydante

pour la violette de Toulouse. Une formulation des extraits à 10% a été investiguée en milieu hydroglycériné assurant une meilleure stabilité sans l'ajout de conservateur et la faisabilité a été avérée. Cependant, l'obtention d'une formulation à l'aspect caramélisé indique une oxydation potentielle de certains composés, notamment les sucres libres et mucilages détectés lors des études en laboratoire. Une purification préalable de ces polysaccharides semble ainsi nécessaire avant la mise en œuvre de la formulation, par précipitation à l'éthanol par exemple. Des études complémentaires sont par conséquent en cours pour répondre à cette problématique qui n'avait pas été relevée lors des essais à l'échelle du laboratoire.

En vue d'une production industrielle des violettes, des études de résistance et de culture des violettes sont des étapes indispensables. La recherche de producteurs est également une étape clé pour l'approvisionnement en matières premières, premier maillon de la valorisation industrielle. Cette recherche est actuellement en cours dans l'étude économique de la valorisation de la violette de Toulouse choisie pour son fort potentiel antiradicalaire.

Les approches holistiques mises en œuvre dans ce travail ont permis de cribler l'ensemble d'une collection et de mettre en évidence deux espèces d'intérêt pour leur potentiel antioxydant et de proposer des ouvertures pour l'aspect phytosanitaire. Cette méthode est applicable à des problématiques similaires pouvant toucher d'autres domaines.

.

PARTIE IV
PARTIE EXPERIMENTALE

IV.1 CONDITIONNEMENT DE LA MATIERE PREMIERE

IV.1.1 MATIERE VEGETALE A DISPOSITION

Les violettes étudiées dans ce mémoire proviennent de la collection nationale de violettes comprenant une centaine de plants mis à disposition par les serres municipales de Toulouse. Deux références, *V. alba* et *V. odorata*, détenues par le Museum d'Histoire Naturelle de Toulouse, ont également été intégrées aux analyses afin de disposer d'un point d'ancrage pour comparer les violettes de la collection mal identifiées, à des échantillons botaniquement validés

IV.1.2 RECOLTE ET SECHAGE DES PLANTES

La récolte du matériel végétal a été réalisée à l'aide de ciseaux et de gants pour éviter toute contamination humaine. Les feuilles ont été rapidement placées dans du papier aluminium avant introduction dans l'azote liquide pour stopper toute activité enzymatique et figer le métabolisme. Pour les expériences requérant l'usage d'une matière non séchée, les échantillons ont été placés au congélateur à -80 °C, sinon ils ont été séchés au lyophilisateur (Labconco Lyph Loc 6) pendant 24 h puis conservés dans un endroit sec à l'abri de la lumière.

IV.2 ETUDE BOTANIQUE

IV.2.1 MONTAGE DES COUPES

IV.2.1.1 Solution d'hydrate de chloral

La solution d'hydrate de chloral est préparée en dissolvant 80 g d'hydrate de chloral dans 20 mL d'eau, d'après les protocoles de la Pharmacopée Européenne (édition en vigueur).

IV.2.1.2 Préparation des coupes

Les techniques utilisées sont celles décrites dans la Pharmacopée Européenne (édition en vigueur), rubrique 2.8.23. Une à deux gouttes de solution d'hydrate de chloral est placée sur une lame et un échantillon de matériel végétal y est introduit. L'ensemble est ensuite recouvert d'une lamelle puis délicatement chauffé jusqu'à ébullition, de manière à éliminer les bulles d'air facilitant par la suite l'examen microscopique (Figure 121). Après chauffage, la face inférieure noircie par la fumée est nettoyée, de même que l'excès de réactif sur la face supérieure et les bords afin de ne pas contaminer les objectifs et les endommager. En effet, l'hydrate de chloral est très corrosif. La lame est laissée à refroidir quelques minutes avant l'examen pour préserver l'objectif.

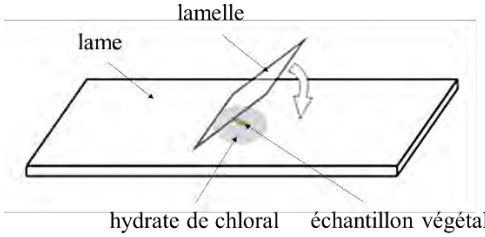


Figure 121. Schéma d'un montage pour analyse micrographique d'un matériel végétal

IV.2.1.3 Examen microscopique

Un étalonnage du microscope optique doit dans un premier temps être effectué à l'aide d'un micromètre, de manière à pouvoir apporter un aspect dimensionnel aux divers éléments observés. Tous les éléments cellulaires, c'est-à-dire les vaisseaux, les tissus de revêtement (épiderme), les tissus conjonctifs (parenchyme), mais encore les tissus sécréteurs (poils, cellules, poches, …) sont à identifier après chauffage, excepté les grains d'amidon qui peuvent être dégradés, voire détruits, par le chauffage.

IV.3 METHODE D'ETUDE GENETIQUE

IV.3.1 PROTOCOLES D'AMPLIFICATION, PURIFICATION ET SEQUENÇAGE DE L'ADN

Pour l'analyse génétique des violettes de la collection, la région des ITS a été intégralement étudiée en incluant la partie non codante 5,8S. Ces fragments d'ADN choisis pour leur fort polymorphisme, ont été amplifiés directement à partir d'un morceau de feuille congelé, par PCR avec le kit *KAPA3G Plant PCR Kit* de Kapabiosystems. Pour cela, plusieurs primers ont été testés mais ce sont finalement un primer universel de White (**White et al. 1990**) qui a été sélectionné comme primer sens (→) de séquence GGAAGTAAAAGTCGTAACAAGG, et un primer antisens (←) spécialement conçu de séquence TCCTCCGCTTATTGATATGC. L'amplification a été suivie d'une purification avant envoi au séquençage pour obtenir la séquence qui pourra être comparée aux bases de données existantes.

IV.3.1.1 Solution de préparation

Cette préparation se composait de 21,6 µL d'eau nucléase free, 25 µL de tampon, 1,5 µL de chaque amorce à 10 µM et 0,4 µL d'enzyme à 2,5 $U.\mu L^{-1}$. Un fragment de feuille encore congelé a été ajouté dans le tube. Toutes ces étapes ont été réalisées en conditions stériles, sous hotte, et après un cycle UV de 20 minutes pour éliminer tout risque de contamination avec d'autres échantillons utilisés par le laboratoire. En effet, les primers utilisés étant universels, ils sont capables de se lier à tout fragment d'ADN présent dans la solution à analyser. C'est pourquoi des précautions ont été nécessaires pour limiter toute contamination du matériel utilisé, ainsi que de l'environnement de préparation.

IV.3.1.2 Cycle PCR et électrophorèse

La PCR était constituée de 35 cycles, chacun comprenant un enchainement d'étapes à des températures fixes telles que : la dénaturation de l'ADN à 95 °C pendant 30 sec, l'hybridation à 62 °C pendant 20 sec et la polymérisation à 72 °C pendant 30 sec. Après réaction, 1 µL de bleu de dépôt 6X a été ajouté à 5 µL de produit PCR. Ce mélange a été déposé sur gel d'agarose à 1% et une migration de 30 minutes à 100 V a été effectuée afin de vérifier l'amplification des fragments d'ADN. En effet, l'électrophorèse sur gel d'aragose en champ homogène assure la séparation des ADN linaires de faible taille (<50 kb) par tamisage moléculaire, puisque les molécules entre une centaine de paires de bases et une vingtaine de kilobases présentant une mobilité sensiblement égale à l'inverse du logarithme de leur taille. La présence d'une unique bande vers 600 pb indiquera la bonne réalisation de la PCR, tandis que la présence de plusieurs bandes pourra indiquer une contamination ou la présence d'hybrides, couramment observés chez les *Viola*. La révélation du gel est ensuite faite au bromure d'éthidium (BET) et aux ultraviolets (UV).

IV.3.1.3 Purification

La purification des séquences amplifiées a été faite grâce au kit *Wizard® SV Gel et PCR Clean-Up System* de Promega, conçu pour extraire et purifier des fragments d'ADN de l'ordre de 100 et 10 k pb, à partir de gel d'aragose standard ou directement à partir de l'amplification PCR. Pour des fragments de l'ordre de 500 pb, le taux de récupération est d'environ 90%. Cette étape est réalisée dans le but d'éliminer les nucléotides ainsi que les primers introduits en excès dans le milieu réactionnel. Le produit de la PCR (45 µL) a été déposé sur une petite colonne de purification avec une quantité équivalente de solution de liaison membranaire (isothiocyanate de guanidine à 4,5 M et acétate de potassium à 0,5 M). Après incubation à température ambiante pendant 1 min, la colonne est centrifugée à 14 000 rpm pendant 1 min puis le liquide est jeté. La colonne est ensuite lavée deux fois avec 700 µL de solution de lavage (acétate de potassium à 10 mM, éthylène diamine tétra-acétique [EDTA] à 16,7 µM et éthanol 80%) avec centrifugation pendant 5 minutes à 14 000 rpm. Ce système de purification repose sur la capacité de l'ADN à se lier aux membranes de silice en présence de sels chaotropiques détruisant la structure tridimensionnelle des macromolécules biologiques. Enfin, 50 µL d'eau nucléase free ont été ajoutés puis après 1 min à température ambiante, une centrifugation à 14 000 rpm pendant 1 min est réalisée et l'ADN récupéré est stocké à 4 °C.

IV.3.1.4 Séquençage

Pour chaque échantillon, 5 µL de produit d'amplification purifié ont été envoyés à séquencer avec 5 µL de solution d'amorce à 5 µM. La société GATC Biotechn a réalisé le séquençage par la méthode de Sanger permettant de déterminer l'ordre des nucléotides dans un brin d'ADN (**Sanger et Coulson, 1975**). La méthode est basée sur la synthèse d'un brin d'ADN complémentaire du brin matrice à l'aide de l'ADN polymérase. Celle-ci assure l'incorporation de nucléotides libres (notés dNTP) présents dans le milieu réactionnel. Elle repose sur la présence de dNTP modifiés dans le milieu (notés ddNTP) et conduisant à l'interruption de l'élongation en empêchant la formation de la liaison avec le nucléotide suivant (**Prober et al., 1987**). Il y a donc compétition entre les deux types de nucléotides et les conditions de l'expérience sont telles que les ddNTP s'incorporent à toutes les positions possibles. Ces derniers sont porteurs d'un marqueur fluorescent distinct assurant ensuite l'identification des nucléotides (Figure 122). Les fragments d'ADN obtenus sont soumis à une électrophorèse capillaire et les fragments migrent alors à différentes vitesses selon leur taille.

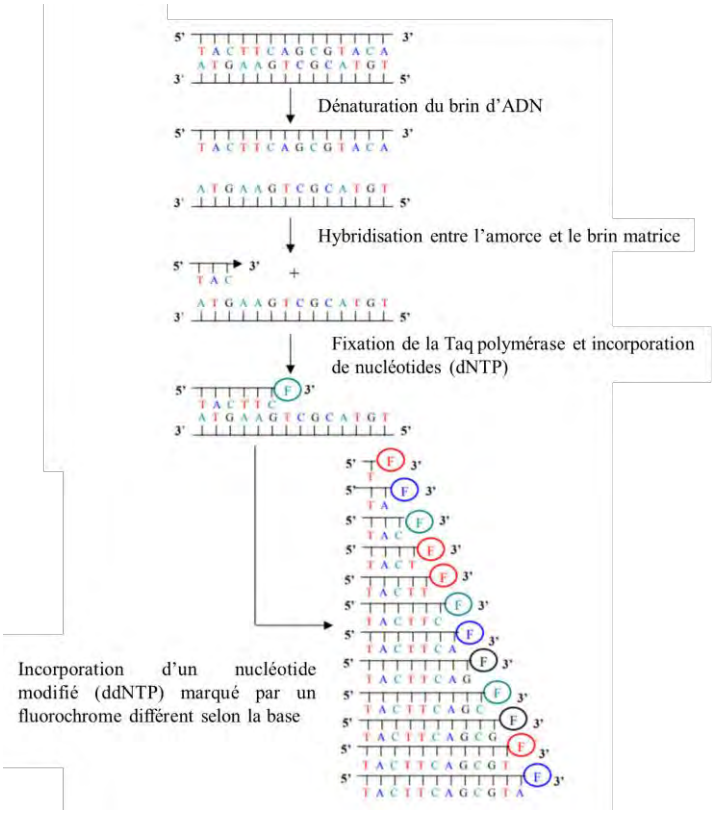


Figure 122. Principe de la méthode de Sanger

IV.3.2 PROTOCOLE D'ANALYSE PHYLOGENETIQUE

La construction des arbres phylogénétiques a été réalisée à l'aide des logiciels CLC Main Workbench software version 7.8.1 (Qiagen Aarhus A/S) et PAUP*4.0 (**Swofford, 2001**). Un nettoyage préliminaire des séquences a été réalisé afin d'éliminer dans un premier temps les parties de séquences ambiguës, notamment au niveau de l'accroche des primers, et dans un deuxième temps afin d'harmoniser la taille des séquences. En effet, cela peut avoir un impact sur la qualité de l'alignement, et donc inévitablement sur les valeurs de bootstrap par la variabilité des sites comparables.

IV.3.2.1 Analyse de similarité de séquences

Pour l'analyse de similarité de séquences, une méthode de distance a été utilisée avec l'algorithme Neighbor Joining basée sur le principe du minimum d'évolution, impliquant que la longueur totale de l'arbre doit être la plus courte possible. L'indice Kimura 80 a été utilisé et une analyse bootstrap mise en œuvre sur 1000 réplicas.

IV.3.2.2 Arbre phylogénétique

Pour la construction de l'arbre phylogénétique, la méthode du maximum de vraisemblance et celle de parcimonie ont été mises en œuvre. Les arbres de parcimonie maximum ont été construits à l'aide du logiciel PAUP*4.0. Des recherches heuristiques ont été menées avec les options MulTrees et ACCTRAN ainsi que l'algorithme de permutation de branche TBR. L'arbre de maximum de vraisemblance a été réalisé avec le logiciel CLC en intégrant uniquement quelques séquences expérimentales caractéristiques de chaque groupe précédemment identifié par l'analyse de distance, et ceci pour une meilleure lisibilité de l'arbre. Les valeurs bootstrap ont été calculées à partir de 1000 réplicas afin de définir la robustesse des branches. L'enracinement des arbres a été effectué par intégration d'un groupe extérieur composé d'échantillons de référence de la base de données GenBank : *Viola tricolor* du genre *Viola*, et plus généralement du genre *Rinorea* de la famille des Violacées et du genre *Malpighia* de l'ordre des Malphighiales.

Pour la construction de l'arbre combinant les données génétiques et chimiques, les données UHPLC-MS ont été binarisées en fonction de la moyenne de l'intensité des pics en indiquant une valeur de «1» pour les intensités supérieures à la moyenne et «0» dans le cas inverse.

IV.4 METHODE D'ANALYSE METABOLOMIQUE ET RESEAU MOLECULAIRE

Le cheminement de l'analyse métabolomique (Figure 123) va tout d'abord consister à visualiser les données avec MZmine 2 version 2.29, afin de déterminer la ligne de base permettant d'éliminer une partie du bruit de fond et d'identifier la présence de complexes. Par la suite, ces mêmes données sont introduites dans le logiciel MS-DIAL qui va conduire en sortie et après fixation de certains paramètres, dont la tolérance pour la MS1 et la MS2 ainsi que la valeur du bruit de fond pour considérer la présence d'un pic, à une liste de variables. Celle-ci pourra être indexée par l'identification d'adduits qui auront été renseignés selon le mode d'ionisation utilisé. L'utilisation de bases de données est possible directement par MS-DIAL et l'identification reposera alors sur la formule brute. L'ensemble de cette liste de pics, caractérisée par des temps de rétention et des masses, pourra être intégrée dans le logiciel MS-FINDER pour interrogation d'une plus large gamme de bases

de données locales et/ou maison, adaptées au type d'extrait analysé. Ces correspondances sont basées sur la comparaison des fragments des composés obtenus expérimentalement avec les fragments des références des bases de données obtenus *in silico*. Des scores de similarité permettent ensuite de juger de la pertinence de l'hypothèse structurale proposée. Les données pourront ensuite être statistiquement traitées avec le logiciel SIMCA version 14 (Umerics AB) pour mettre en relation des activités biologiques et des composés potentiellement actifs.

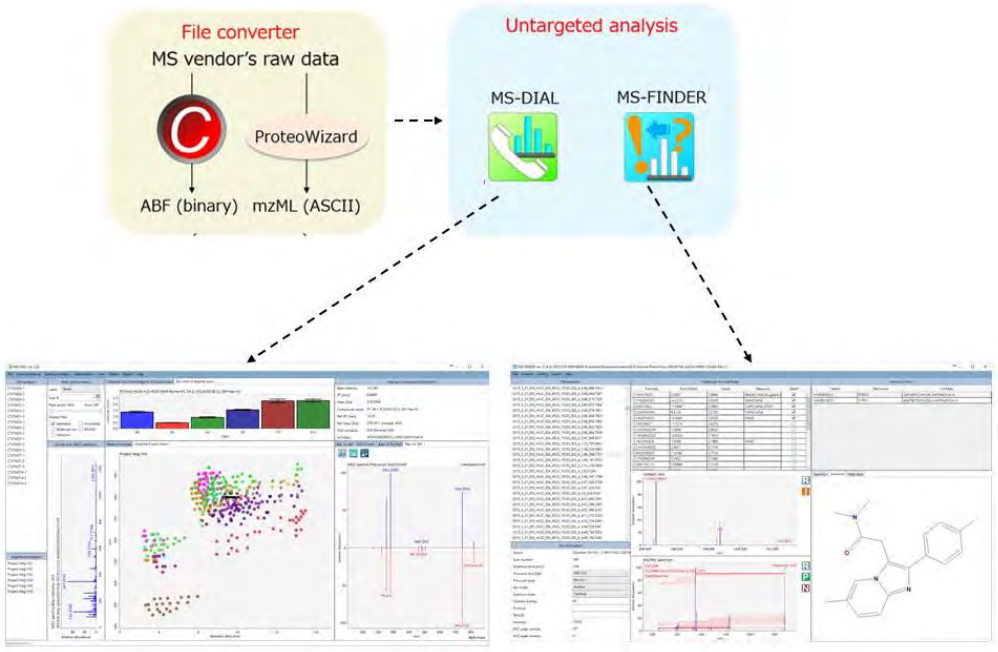


Figure 123. Cheminement de l'analyse métabolomique

IV.4.1 CONVERSION ET ALIGNEMENT DES DONNEES

Les profils chromatographiques obtenus en GC-MS et UHPLC-MS ont été convertis en .abf à l'aide du convertisseur Abf Converter et l'ensemble des profils ont été alignés avec MS-DIAL en mode centroïde. Les paramètres de base ont été optimisés et fixés comme suit : la tolérance pour MS1 à 0,01 Da, pour MS2 à 0,2 Da, la charge maximale des isotopes à 2. La méthode de « smoothing » est celle fixée par défaut à savoir « Linear weighted moving average » et la valeur du « sigma window » affectant la résolution de la déconvolution des pics a été fixée à 0,5 afin de ne pas intégrer trop de bruit de fond en appliquant une valeur trop basse (< 0,3), ni réduire la résolution par une valeur trop haute (> 0,7).

Dans le cas d'acquisition des profils en modes négatif et positif, l'utilisation du script R MS-combine (**Calderón-Santiago et al., 2016)** a permis de combiner les deux listes de pics en éliminant les composés redondants, c'est-à-dire en conservant un seul ion moléculaire s'il est détecté dans les deux modes.

IV.4.2 IDENTIFICATION DES COMPOSES

L'identification des composés a ensuite été réalisée sur la base des fragmentations à l'aide des bases de données NIST en GC-MS et de produits naturels de MS-FINDER (UNPD, KNApSAcK, PubChem, NANPDB, PlantCyc et ChEBI) et des bases de données plus ciblées sur le genre *Viola* à l'aide du DNP en UHPLC-MS. De même que pour MS-DIAL, les tolérances de MS1 et MS2 ont été

respectivement fixées à 0,01 et 0,2. Les règles de Lewis et de Senior ont été appliquées avec une tolérance au niveau du profil isotopique de 20% et la recherche de formules brutes incluant uniquement les éléments chimiques O et N. Une analyse batch est réalisée et les résultats ont ensuite été exportés au format csv.

IV.4.3 ANALYSE STATISTIQUE

L'ensemble des analyses statistiques a été réalisé à l'aide des logiciels XLSTAT version 2018.1 (Addinsoft, New-York, USA) et SIMCA version 15 (Umerics, Umeå, Sweden). Différentes méthodes ont été appliquées telles que l'analyse de variance, l'analyse en composantes principales ainsi que des régressions linéaires afin de mettre en évidence d'éventuelles relations entre les réponses biologiques observées et la composition des extraits analysés, de manière à cibler plus rapidement les molécules présentant un potentiel biologique au sein d'un extrait naturel complexe.

IV.4.3.1 ANOVA

La première étape a été d'effectuer une analyse de variance appelée ANOVA afin de vérifier l'existence d'une relation entre les activités biologiques mesurées et l'appartenance à des groupes, principalement les espèces concernées dans ces travaux. Ainsi, elle va permettre de mettre en évidence d'éventuelles différences significatives d'activité entre les différentes espèces. Une comparaison des profils chimiques par ACP et régression linéaire seraient alors plus que pertinentes pour aller plus loin dans l'analyse. La principale contrainte pour pouvoir appliquer une ANOVA est la normalité des données. Ainsi, il est important de la vérifier au préalable sur les données brutes ou encore, plus pertinent, sur les résidus centrés réduits issus de l'ANOVA qui doivent suivre une loi normale, c'est-à-dire que 95% des résidus doivent appartenir à l'intervalle [-1,96, 1,96]. Des transformations des données brutes, telle que la transformation logarithmique, peuvent être opérées pour entrer dans ce cadre.

IV.4.3.2 ACP et régression linéaire

Avant interprétation des données à l'aide du logiciel SIMCA, différents traitements peuvent être réalisés sur les données, dont des transformations logarithmiques et le choix de l'échelle de travail. Pour cela la position des QC (pool des échantillons analysés) est étudiée et le modèle les plaçant le plus proche du centre est sélectionné. En effet, ces QC contiennent la globalité des composés des échantillons qui eux n'en contiennent qu'une partie et devraient donc s'organiser tout autour. Une ACP est ensuite réalisée en première intention afin de visualiser l'organisation globale des données et détecter la présence éventuelle d'outliers. Ensuite, des analyses supervisées PLS et OPLS sont entreprises pour pouvoir effectuer des études de prédiction et identifier les métabolites qui sembleraient les plus impliqués dans les différentes activités étudiées. Ainsi différents modèles peuvent être construits selon les études menées. Par exemple, dans le cas de l'analyse de l'activité antioxydante, l'intégration de la valeur de l'activité mesurée en paramètre Y va permettre, à l'aide de l'analyse des coefficients de la régression linéaire, de mettre en évidence les composés jugés statistiquement actifs (OPLSr). Dans le cas de la classification des violettes c'est l'identification des espèces en paramètre Y qui va permettre d'étudier la présence de groupes en lien avec l'identification génétique (OPLS-DA). Pour chaque modèle, une validation est effectuée afin d'attester de sa fiabilité à travers les tests de permutations et le diagnostic CV-ANOVA intégrant le paramètre nommé *p*-value. Si ce dernier est inférieur à 0,05 le modèle sera considéré valide et donc prédictif et son analyse permettra de tirer des conclusions pertinentes.

IV.4.4 RESEAU MOLECULAIRE

Les réseaux moléculaires (« molecular networking ») sont une approche bio-informatique qui permet d'interpréter les données de spectres de masse qui peuvent être très étendues selon la complexité de l'extrait analysé. Cet outil a été développé par les groupes de Pieter C. Dorrestein et de Nuno Bandeira de l'université de Californie, de manière à aider l'interprétation des données provenant d'un mélange complexe comme le sont les extraits d'origine naturelle (**J. Y. Yang et al., 2013**).

C'est donc un outil relativement puissant notamment lors de l'analyse sans connaissance préalable des extraits. Deux plateformes ont été mises en place : GNPS (https://gnps.ucsd.edu) et MetamapR (**Grapov et al., 2015**). Un logiciel de visualisation, Cytoscape (**Shannon et al., 2003**) est utilisé de manière à améliorer l'apparence des réseaux obtenus sur les plateformes.

IV.5 METHODES D'ETUDES PHYTOCHIMIQUES

IV.5.1 METHODES D'ETUDE DES COMPOSES VOLATILS DES FLEURS

L'analyse des métabolites volatils des plantes s'est fortement accentuée avec le développement de la chromatographie gazeuse associée à un spectromètre de masse. Les nombreuses études réalisées ont montré le lien entre ces émissions de composés volatils et l'environnement de la plante. Ils sont également très utilisés en tant que biomarqueurs pour étudier les différences de profils chimiotaxonomiques entre différentes espèces (**Niederbacher et al., 2015**).

IV.5.1.1 Extraction par espace de tête et microextraction en phase solide

La technique de l'espace de tête (**Fernandez et al., 2007**) couplée à la microextraction en phase solide (SPME) (**Papet et al., 2010** ; **Burgot et Pellerin, 2003**) est une méthode d'extraction et de concentration de molécules organiques sans solvant à partir d'une matrice complexe liquide ou solide. L'échantillon est introduit dans un flacon hermétique surmonté d'un septum. Les composés volatils contenus dans la matrice diffusent vers la phase gazeuse constituant l'espace de tête jusqu'à l'atteinte d'un équilibre, avant prélèvement par adsorption sur une fibre de silice fondue revêtue d'un polymère et analyse en GC-MS. L'équilibre matrice/espace de tête est régi par un coefficient K_h défini comme suit, où C est la concentration de l'analyte soit dans la matrice (M), soit dans l'espace de tête (G) : $K_h = \frac{C_M}{C_G}$. Cet équilibre est influencé par la température, l'agitation et le temps de chauffage.

La désorption des composés est généralement réalisée par voie thermique. Plusieurs fibres de polarités différentes (**Fernandez et al., 2007**) sont actuellement disponibles (Tableau 37). Elles reposent sur l'existence de trois types de revêtement (**Wardencki at al., 2004**):

- les polymères pseudo-liquides, comme le polydiméthylsiloxane (PDMS) très visqueux et le polyacrylate (PA). Ils extraient les composés par partage entre la matrice et le polymère ;
- les polymères poreux comme le divinylbenzène (DVB) associé au PDMS ou au Carbowax®. Ils extraient les composés par un mécanisme d'adsorption.
- Les polymères très poreux comme le Carboxen® (CAR).

Tableau 37. Fibres disponibles pour les analyses SPME

Nature de la phase stationnaire	Sigle	Polarité	Composés extraits
Polydiméthylsiloxane	PDMS	Apolaire	Volatils Semi-volatils
Polyacrylate	PA	Polaire	Semi-volatils
Polydiméthylsiloxane/Divinylbenzène	PDMS/DVB	Bipolaire	Volatils polaires
Carboxen/Divinylbenzène	CAR/PDMS	Bipolaire	Volatils
Divinylbenzène/Carboxen/Polydiméthylsiloxane	DVB/CAR/PDMS	Bipolaire	Volatils Semi-volatils

La SPME est une méthode couramment utilisée de nos jours pour remplacer des analyses qui nécessitent des extractions longues et couteuses, avec une forte consommation de solvant. Par ailleurs, elle présente de nombreux avantages (**Burgot et Pellerin, 2003**) :

- un temps de préparation des échantillons réduit ;
- une automatisation possible ;
- peu d'interférences avec les composés non volatils de la matrice ;
- un système respectueux de l'environnement par l'absence de solvant ;
- la réutilisation des fibres ;
- une sélectivité des composés extraits par le choix du revêtement polymérique.

Cependant, elle présente un manque de reproductibilité qui peut varier jusqu'à 10%, mais qui reste meilleure que d'autres techniques, comme l'extraction liquide-liquide ou solide-liquide (**Burgot et Pellerin, 2003**). De plus l'analyse quantitative portant sur l'espace de tête est fortement influencée par les conditions expérimentales mises en œuvre. Enfin, les fibres restent fragiles et peuvent être endommagées par des composés de hautes masses moléculaires telles que les protéines qui s'adsorbent de manière irréversible.

IV.5.1.1.1 Standardisation de la méthode d'extraction HS-SPME

L'extraction des composés volatils a été réalisée par la méthode de l'espace de tête couplée à une microextraction en phase solide, à l'aide d'une fibre Supelco DVB-CAR-PDMS. Avant utilisation, la fibre a été conditionnée selon les recommandations du fabricant. Un protocole précis a été mis en place afin d'analyser les violettes dans des conditions similaires. Pour cela, il a été décidé de laisser les pots de violettes dehors, puis de les placer en intérieur à température ambiante une heure avant le début des manipulations. Dans un environnement contrôlé, les pots ou les fleurs, selon la configuration étudiée, ont été confinés sous cloche ou en flacon (Figure 124) afin de procéder à l'émission des composés volatils, qui ont été par la suite piégés sur la fibre.

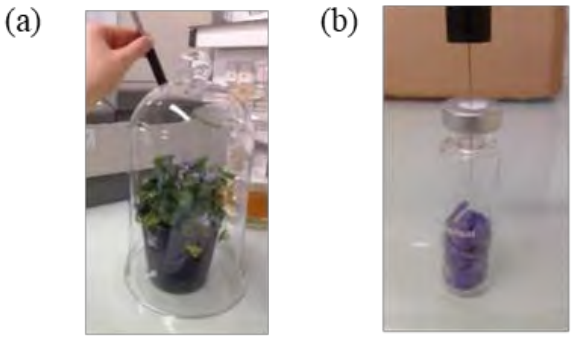


Figure 124. Technique de piégeage pour l'analyse de la plante entière (a) et des fleurs coupées (b)

IV.5.1.1.2 Optimisation de l'extraction HS-SPME par un plan d'expérience

Pour optimiser la méthode de piégeage et d'adsorption des composés volatils émis par les fleurs fraîches à maturité, un plan d'expérience de Doehlert du second degré a été mis en œuvre. Cette méthode a été utilisée car le plan de Doehlert est adapté à l'optimisation de paramètres, dans le cas du suivi d'une équation du second degré, ce qui a été supposé dès le départ. C'est donc un plan de modélisation qui va permettre d'obtenir une relation mathématique liant la réponse observée, aux facteurs étudiés et ainsi assurer le tracé d'une surface de réponse.

Dans cette étude, deux paramètres ont été étudiés, à savoir le temps d'équilibre de la matière végétale pour l'émission des composés volatils, et le temps d'exposition de la fibre pour l'adsorption des composés émis. Cette optimisation a été réalisée sur la violette de Toulouse pour des raisons d'approvisionnement. Un domaine d'étude a été préalablement défini à partir de la littérature afin d'établir les bornes des paramètres étudiés.

IV.5.1.2 Méthode d'analyse sensorielle

Les composés volatils issus de matrices végétales, et particulièrement de fleurs, peuvent posséder une odeur et contribuer à l'empreinte olfactive du parfum de la fleur. C'est pourquoi une analyse sensorielle peut être réalisée en parallèle de la séparation chromatographique afin de pouvoir déterminer l'impact olfactif des différentes molécules identifiées dans le mélange.

IV.5.1.2.1 Le « Champ des odeurs »®

Le « Champ des odeurs »® a été créé par Jean-Noël Jaubert en 1983 et IAPSENTIC suite à un programme de recherche sur la relation structure-activité des molécules odorantes. La mise en application de ce concept a été concrète en 1984 avec la naissance de l'analyse olfactive. La description olfactive par la méthodologie du « Champ des odeurs »® ne consiste pas à évoquer des souvenirs dont l'information serait bien trop variable, mais bien de comparer à des éléments référents connus. Il est alors important de différencier l'évocation de connotations personnelles et la description qui est plus objective car n'intégrant pas de dimensions psychosociales. La description demande donc l'établissement d'un langage commun par l'intermédiaire de référents. La représentation du « Champ des odeurs »® a débuté sur le postulat que deux molécules qui ne déclenchent pas la même évocation, avaient une odeur différente (**Rouadjia, 2009**). Les résultats des travaux ont permis la mise en évidence d'un espace multidimensionnel avec six premiers groupes de substances odorantes ne partageant aucune évocation commune et appelé « pôles » (Figure 125 a). Les autres substances ont été placées dans la structure sous forme de secteurs de transition appelés « dominantes » (Figure 125 b), jusqu'à l'obtention du référentiel total comportant 45 référents (Figure 125 c). En haut, on retrouve les notes qualifiées de légères à faible durabilité (pinène), vers l'avant les notes un peu pointues, fusantes (citral) et vers le bas les notes lourdes à durabilité élevée (pyrazine). A gauche, sont retrouvées les notes aminées et à droite les notes soufrées.

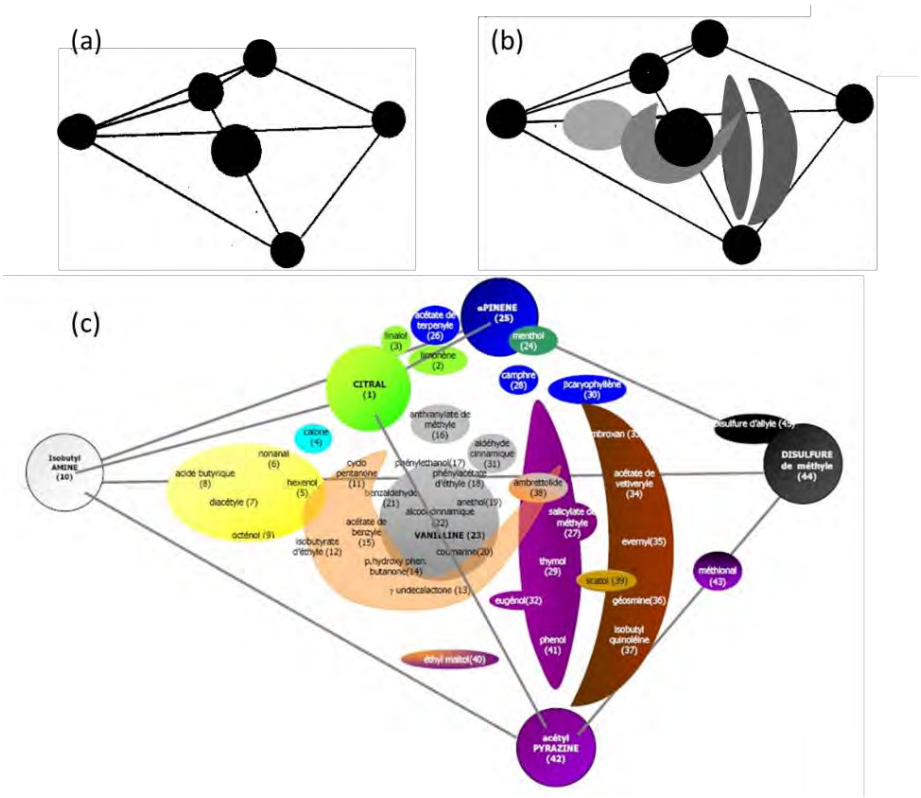


Figure 125. Construction du « Champ des odeurs »®

Cette approche demande donc une formation et un entraînement régulier et est utilisée depuis 1983 en laboratoires d'analyses sensorielles R&D. Sur la base de cet espace des odeurs, l'analyse olfactive permettra de caractériser les qualités odorantes des extraits ou des molécules directement séparées par GC, afin d'obtenir des profils olfactifs objectifs qui pourront non seulement être comparés dans le temps, mais également mis en miroir des profils chromatographiques afin d'identifier les molécules odorantes à l'aide de bases de données. Un composé odorant est caractérisé par trois paramètres : sa concentration, sa note olfactive ainsi que son seuil de perception, c'est-à-dire la concentration minimum à laquelle le composé est détecté. Il est à noter que l'intensité du pic et son impact olfactif ne sont pas liés. En effet, un composé mineur peut avoir une forte activité odorante et inversement un composé majoritaire n'en avoir aucune. De nombreuses applications touchent actuellement le domaine alimentaire et le domaine de la parfumerie et des huiles essentielles.

IV.5.1.2.2 Analyse des effluves odorantes par GC/O

La chromatographie gazeuse couplée à l'analyse olfactométrique (GC/O) (**Fernandez et al., 2009**) repose sur une séparation sur colonne des composés volatils avec une analyse simultanée de leur impact odorant. Pour ce faire, le nez est généralement employé en tant que détecteur sensible et sélectif. Pour mettre en place cette analyse, un système de division est introduit en fin de colonne chromatographique dans le but de scinder le flux en proportion égale, dont une partie sera dirigée vers le détecteur physique et l'autre vers celui sensoriel. Des sensations de dessèchement des parois nasales causées par le gaz vecteur de la GC ont rapidement été ressenties. Ainsi, de l'air humide (humidité relative : 50-75%) est à présent utilisé en tant que gaz vecteur délivrant les composés odorants (Figure 126).

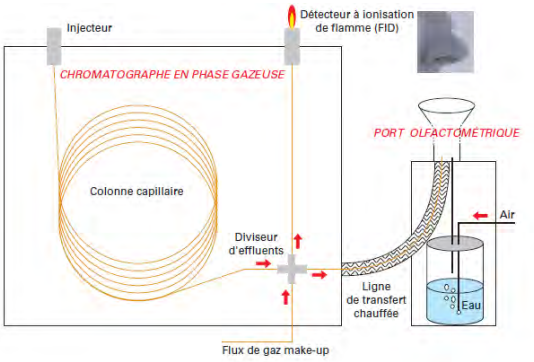


Figure 126 : Schéma du couplage GC/O (Fernandez et al., 2009)

Les effluves odorantes ont été analysées au niveau des ports olfactométriques (ODP). Ces derniers sont reliés à l'ordinateur et les olfactogrammes sont enregistrés à l'aide du logiciel Record (Agilent). Cette analyse permet de déterminer la concentration de l'odeur, son intensité et sa nature, pour chaque pic élué. La mesure a fait appel à un jury local de trois personnes, dont deux ayant une formation dans ce domaine. A l'aide d'une télécommande, l'enregistrement de l'odeur détectée a été réalisé avec appréciation de l'intensité de l'odeur sur une échelle de 1 à 4, de la durée de l'odeur et de sa qualification par enregistrement vocal.

IV.5.1.3 Méthode d'analyse par chromatographie en phase gazeuse

La chromatographie gazeuse (GC) a été introduite en 1952 (**Bartle et Myers, 2002**) et est très utilisée depuis dans de nombreux domaines et reste la méthode de référence pour l'analyse des composés volatils. Elle sépare les molécules selon leur volatilité et leur affinité avec la phase stationnaire généralement apolaire comme le polydimethylsiloxane et sa longueur est généralement comprise entre 25 et 30 m pour obtenir un bon compromis entre la qualité de la séparation et le temps d'analyse (**Sangster et al., 2006**). Deux principaux modes d'injection sont utilisés. Le mode split s'applique pour prévenir la saturation de la colonne, mais pouvant créer une discrimination au niveau des métabolites faiblement concentrés. Au contraire, le mode splitless assure l'injection de l'intégralité de l'échantillon et est adapté à l'analyse de traces.

Pendant longtemps, le système de détection utilisé était le détecteur à ionisation de flamme (FID) pour sa bonne stabilité et sa haute sensibilité, avec une limite de détection aux alentours du picogramme voire du nanogramme. Cependant, le manque de données spectrales d'intérêt pour l'identification des composés a conduit à l'utilisation plus importante du spectromètre de masse. Les composés élués sont ionisés selon un mode d'ionisation doux par impact chimique (CI) ou dur par impact électronique (EI). L'identification peut s'effectuer à l'aide de composés de référence ou par l'intermédiaire de bases de données (**Milman et Zhurkovich, 2016**) dont les principales sont regroupées dans le Tableau 38**.**

Tableau 38 : Bases de données les plus utilisées pour l'identification en GC-MS

Base de données	Nb de spectres	Site internet
Wiley	775 500	http://eu.wiley.com/WileyCDA/ WileyTitle/productCd-1119171016. html
NIST 14	276 248	http://nistmassspeclibrary.com/
Golm Metabolome Database	3 500	http://gmd.mpimp-golm.mpg.de/
Fiehn Library	2 200	http://fiehnlab.ucdavis.edu/Metabolite- Library-2007/

Les applications de cette technologie regroupent trois principaux domaines : la pétrochimie (**Blomberg et al., 1997**), l'environnement et l'agroalimentaire comme le travail réalisé sur les composés volatils des truffes (**Costa et al., 2015**) et les arômes et parfums comme pour l'étude des roses (**Özel, Göğüş, et Lewis, 2006**).

IV.5.1.3.1 Méthode analytique GC-O-MS

La fibre DVB-CAR-PDMS est thermiquement désorbée au niveau de l'injecteur de la GC à 250 °C pendant 5 min. L'analyse GC a été réalisée sur un appareil Agilent 7890B, avec une colonne apolaire HP 5 MS (30 m x 0,25 µm id x 0,25 µm, Agilent, Sante Clara, USA) en mode splitless, en couplage à un spectromètre de masse Agilent Technologies 5977B MSD. Les conditions d'analyse ont été :

a) Pour la GC

- Température de l'injecteur : 250 °C
- Pression de l'injecteur : 14,6 psi
- Flux gaz vecteur (hélium) : 1 mL/min
- Pression dans la colonne : 3,6 psi
- Gradient de température de la colonne :
  - 60 °C pendant 10 min
  - 130 °C à 4 °C/min
  - 150 °C à 2 °C/min
  - 250 °C à 25 °C/min (maintien 10 min)

b) Pour la masse

- Température de la masse : 250 °C
- Énergie d'ionisation (EI) : 70 eV
- Plage de balayage : 50 à 550 Da

c) Pour les sorties olfactives (« Olfactory Detector Port »)

- Température des lignes de transfert : 220 °C

Dans le cas de l'analyse des standards en condition liquide, la même méthode a été utilisée, mais cette fois-ci en mode split de manière à ne pas saturer le détecteur. Ce split a été fixé à 150.

IV.5.1.3.2 Méthode d'analyse métabolomique des données GC-MS

IV.5.1.3.2.1 Traitement des données GC-MS

Les données GC-MS ont été analysées avec MS-DIAL version 2.74 avec les paramètres suivants : la fenêtre d'analyse était comprise entre 4 à 40 min, les masses détectées entre 50 et 600 Da et la ligne de base a été fixée à 500. L'alignement des profils a été réalisé sur la base du profil de la violette de Toulouse, avec une tolérance sur le temps de rétention de 0,075 min.

IV.5.1.3.2.2 Identification des composés

L'identification des composés a été basée dans un premier temps sur la comparaison des spectres de masse expérimentaux avec la base de données NIST 14. L'injection de standards Sigma Alrich a ensuite été réalisée afin de comparer les temps de rétention, en utilisant la même méthode d'analyse. Pour cela, deux mélanges ont été préparés en additionnant deux gouttes de chaque standard. Deux gouttes de ces mélanges ont été déposées dans un flacon et de l'EtOH a été ajouté de manière à diluer la mixture avant injection en GC-MS.

IV.5.2 METHODES D'ETUDE DES METABOLITES SECONDAIRES NON VOLATILS DES FEUILLES

IV.5.2.1 Broyage des feuilles

Les feuilles lyophilisées ont été broyées à l'aide d'un moulin à café (SEB) ou par l'intermédiaire d'un broyeur à billes Beadbug de 230V (Benchmark) avec des tubes de 2 mL comprenant des billes de zirconium de diamètre 1 mm. Dans ce cas, deux cycles de 60 min avec passage dans la glace 30 secondes ont été réalisés pour un broyage satisfaisant.

IV.5.2.2 Méthodes d'extraction

Les méthodes d'extraction traditionnelles font généralement appel à l'utilisation de solvants. Aujourd'hui, des approches plus en lien avec la chimie verte sont appliquées à savoir des techniques classiques telle que l'utilisation de solvant hydroalcoolique qui peuvent convenir au domaine cosmétique, mais aussi des techniques plus innovantes telles que l'extraction supercritique qui est en plein essor depuis quelques années en raison des réglementations environnementales plus strictes.

Différentes méthodes d'extraction ont été mises en œuvre à l'échelle du laboratoire dans le but de comparer la teneur en molécules antioxydantes des extraits obtenus. Ces méthodes ont été sélectionnées selon deux critères : leur impact environnemental (elles devaient faire partie des procédés verts) et leur faisabilité industrielle dans une optique future de valorisation. Au total, quatre méthodes ont été étudiées à savoir, l'extraction par solvant hydroalcoolique, par CO_2 supercritique, par micro-ondes et par électroporation.

IV.5.2.2.1 Extraction hydroalcoolique classique

Les extractions solide-liquide (**Nejia, 2013**) sont des opérations de transfert de matière entre la matrice solide et le solvant d'extraction liquide. C'est une étape essentielle dans le procédé d'analyse et de production d'extraits naturels pour divers domaines industriels comme la pharmacie ou la cosmétique.

IV.5.2.2.1.1 Protocole pour l'étude détaillée de la violette de Toulouse

L'extraction a été réalisée avec une solution éthanolique à 80% compatible avec l'industrie cosmétique. Une quantité de 10 g de poudre de parties aériennes de violette de Toulouse a été placée dans un erlenmeyer, et environ dix fois plus, soit 100 mL, de solvant ont été ajoutés. L'erlenmeyer bouché pour éviter l'évaporation du solvant a ensuite été placé aux ultrasons pendant 30 minutes, de manière à favoriser et accélérer l'extraction. Après filtration sur papier, une nouvelle extraction a été effectuée sur le résidu. Les deux filtrats ont ensuite été combinés puis séchés à l'aide d'un évaporateur rotatif ainsi qu'un concentrateur centrifuge sous vide (SpeedVac) jusqu'à obtenir un extrait sec (sous forme de pâte) conservé au congélateur.

IV.5.2.2.1.2 Protocole pour l'analyse de la collection de violettes

L'extraction a été réalisée avec une solution éthanolique à 80%. Une quantité de 80 mg de parties aériennes lyophilisées a été broyée au broyeur à billes et environ dix fois plus de solvant, soit 800 µL, a été ajouté. Le tube a ensuite été placé aux ultrasons pendant 10 minutes, suivi d'une centrifugation à 14 000 rpm pendant 2 min. Le surnageant a été mis de côté et une nouvelle extraction a été effectuée sur le résidu. Les deux filtrats ont ensuite été combinés puis une élimination de la chlorophylle a été réalisée

à l'aide de silice C_{18}. Une masse quatre fois supérieure à la masse végétale initialement apportée a été ajoutée au surnageant, suivi d'un passage au vortex et d'une centrifugation à 14 000 rpm pendant 2 min. Le nouveau surnageant est récolté puis séché jusqu'à obtenir un extrait sec conservé au congélateur.

IV.5.2.2.1.3 Protocole pour l'extraction des cyclotides

L'hypothèse formulée sur l'activité phytosanitaire des violettes portant sur leur teneur en cyclotides, il a été décidé d'extraire l'ensemble de la collection avec un solvant plus apolaire que dans toutes les expériences réalisées précédemment, sur la base des recommandations de Yeshak et ses collaborateurs (**Yeshak et al., 2012**). Une extraction avec une solution aqueuse de méthanol à 50%, assistée par ultrasons a été préférée à la macération exposée dans ces travaux pour des raisons de traitement haut débit de la collection nationale.

IV.5.2.2.2 Extraction hydroalcoolique optimisée

L'étude d'une collection entière de violettes a conduit à optimiser l'extraction, afin de traiter rapidement et de manière exhaustive un maximum d'échantillons. Le principe repose sur le broyage, l'extraction et l'élimination des pigments quasi simultanés. L'appareil utilisé est un broyeur-homogénéisateur (Bead ruptor elite, OMNI international). Environ 100 mg de feuilles fraîches stockées à - 80 °C ont été introduits dans des tubes de 2 mL comprenant une quarantaine de billes de céramique de 1,4 mm de diamètre (environ 0,61g). Est ensuite directement additionné le premier solvant (700 µL), une solution de méthy tert-butyl éther et de méthanol en proportion volumique 3 :1. Un broyage a été effectué à 6 m/s pendant 1 min puis le tube a été placé aux ultrasons pendant 10 min. Est ensuite ajoutée une quantité équivalente d'une solution aqueuse de méthanol en proportion volumique 3 :1 contenant trois standards : la corticostérone et le ^{13}C sorbitol à 0,2 µg/mL ainsi que l'ampicilline à 0,1 µg/mL. Le tube est repassé au broyeur à bille pendant 1 min, suivi d'une centrifugation pendant 5 min à 12 500 rpm et 4 °C. Les deux phases observées sont distribuées dans deux eppendorf conservés au congélateur.

IV.5.2.2.3 Extraction par électroporation

IV.5.2.2.3.1 Principe de l'électroporation

L'électroporation est une technique d'extraction qui consiste à exposer des cellules vivantes à des impulsions électriques courtes, mais intenses entre deux électrodes. Ainsi, une déstabilisation de la membrane sera induite par des modifications de sa polarisation et par sa perméabilisation qui assurera la diffusion des molécules vers l'extérieur de la cellule (Figure 127) (**Ferrari et al., 2010**). En plus de pouvoir conduire à l'extraction de molécules, l'électroporation peut aussi permettre l'insertion de molécules à l'intérieur des cellules. Par exemple, dans le cas d'une électroporation réversible avec une perméabilisation de la membrane et non une rupture, de l'ADN exogène peut être introduit permettant d'étudier par exemple les propriétés et les fonctions d'un gène (**Saldaña et al., 2014**).

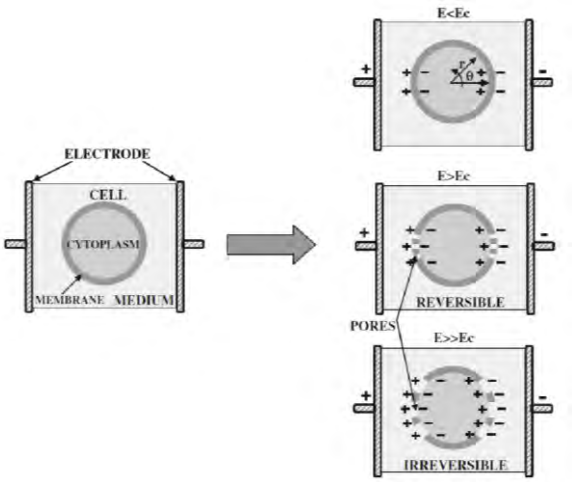


Figure 127. Principe de l'extraction assistée par champ électrique

IV.5.2.2.3.2 Protocole d'extraction mis en œuvre

L'extraction par électroporation a été réalisée par l'intermédiaire de la société Leroy Biotech (www.leroybiotech.com) sur des feuilles de violette de Toulouse surgelées et fraiches. Environ respectivement 70 ou 30 g de parties aériennes surgelées ou fraîches ont été placés entre deux électrodes carrées de 50 mm et écartées de 50 mm, avec respectivement 80 ou 100 g de solvant, à savoir de l'eau ou une solution éthanolique. Deux paramètres ont été étudiés : la valeur du champ électrique appliqué entre les électrodes et le temps d'exposition. La réponse observée était la variation de la conductivité avant et après traitement. Afin de rester homogène entre les différents échantillons, une macération de 2 min a été définie par les opérateurs et respectée pour l'ensemble des analyses. L'extrait, feuilles et solvant, est récupéré dans un flacon puis conservé à 4°C. Un blanc a été réalisé pour chaque type de solvant, c'est-à-dire que les feuilles ont été placées en présence de solvant entre les électrodes pendant 2 min sans application d'un champ électrique. Les différentes conditions opératoires définies par les opérateurs sont regroupées dans le Tableau 39. Il a été vérifié qu'elle n'engendrait pas un champ supérieur à 35 A, la limite du générateur étant de 40 A. Le temps de traitement est caractérisé par la durée d'impulsion et le nombre de répétitions. Une impulsion se compose de deux motifs inversés de même durée et séparés de 50 ms (Figure 128). Ainsi, un traitement de 100 impulsions avec un motif de 1 ms conduira à un temps de traitement de 200 ms.

Tableau 39. Conditions opératoires pour les extractions par électroporation

Champ électrique (V)	Temps d'exposition (ms)
2500	100
2500	200
2500	400
4000	100
4000	200
4000	300
4000	400

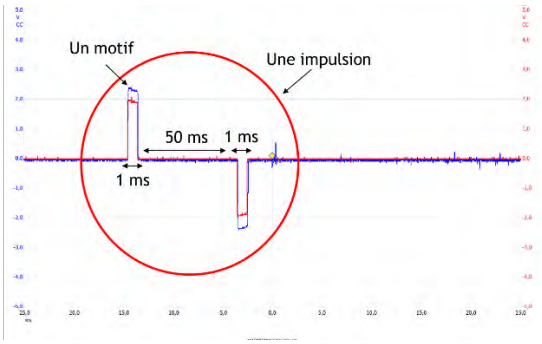


Figure 128. Caractéristique du signal enregistré

IV.5.2.2.4 Extraction par CO_2 supercritique

IV.5.2.2.4.1 Principe de l'extraction supercritique

L'extraction par solvant supercritique (**Nejia, 2013**) est une technique d'extraction plus respectueuse de l'environnement et qui s'inscrit dans le concept de chimie verte. Elle a d'ores et déjà été appliquée avec succès sur une large gamme de matière végétale pour l'obtention de divers composés.

Le principe de cette extraction repose sur la variation du pouvoir solvant du CO_2 en fonction des conditions opératoires de température et de pression appliquées, assurant ainsi une extraction sélective des composés en fonction de leur nature chimique. Il est à noter que dans les conditions supercritiques, le CO_2 s'avère être un bon solvant des molécules peu polaires et apolaires. Les principaux avantages de cette technique sont : une bonne diffusivité en raison de la faible viscosité qui va ainsi permettre une meilleure pénétration dans les matrices et ainsi augmenter le transfert de masse permettant de réduire le temps d'extraction en comparaison à des méthodes traditionnelles ; un bon pouvoir solvant modifiable selon les conditions appliquées assurant une extraction sélective des composés ; une consommation de matière végétale réduite et des conditions opératoires douces qui permettent l'extraction de composés thermosensibles. En revanche, le CO_2 supercritique présentant une faible polarité, son utilisation n'est

pas recommandée pour l'extraction de molécules très polaires. L'alternative est d'utiliser un co-solvant polaire tel que l'eau ou l'éthanol par exemple **(Benaissi, 2013**).
L'extraction supercritique touche de nombreux domaines et est adaptée à l'obtention de molécules antioxydantes comme les polyphénols et est conforme aux exigences de formulations cosmétiques par l'absence de traces de solvants (**Liza et al., 2012**).

IV.5.2.2.4.2 Protocole d'extraction mis en œuvre

L'extraction au CO_2 supercritique a été réalisée à la plateforme de chimie, au laboratoire SPCMIB sur l'appareil SFC-PICLAB analytic md-prep station pic solution. Les feuilles lyophilisées de violette de Toulouse ont été broyées puis une masse exacte (environ 5 mg) de feuilles a été placée dans le réacteur de 10 mL. Deux paramètres ont été étudiés à savoir la température (40 et 50 °C) et la pression (100 et 150 bars). Des billes ont été ajoutées au mélange afin de faciliter l'extraction et du co-solvant, l'éthanol, a été introduit entre 5 et 50% par pas de 5%. Chaque palier a été tenu 20 min. Trois essais ont été réalisés donnant chacun 10 fractions conservées au congélateur.

IV.5.2.2.5 Extraction assistée par micro-ondes

IV.5.2.2.5.1 Principe des micro-ondes

L'extraction assistée par micro-ondes (**Anizon et al., 2003**) implique une interaction directe entre le rayonnement électromagnétique de l'appareil et la matière végétale. Si le solvant absorbe la totalité de l'énergie, le transfert de chaleur s'effectuera du solvant vers la matière végétale par conduction. Le principe est donc similaire à une extraction classique, mais les micro-ondes permettent un échauffement rapide et contrôlable limitant tous risques de dégradation par contact. Dans le cas d'un solvant transparent aux micro-ondes, l'énergie sera absorbée par la matière végétale ce qui entrainera une élévation rapide de la température qui provoquera une augmentation de la pression interne des cellules des structures et finalement leur rupture avec relargage des métabolites sans dégradation. L'extraction peut également s'effectuer sans apport de solvant en utilisant par exemple l'eau de constitution du végétal (**Mandal et al., 2007**). En pratique, pour utiliser cette technique, il faut que le matériel puisse s'échauffer sous l'action du rayonnement. Le principal paramètre à prendre en compte est alors la teneur en eau : si celle-ci est supérieure à 20% en masse, l'échauffement sera facile tandis que pour une teneur inférieure, la mise en œuvre est complexifiée.

IV.5.2.2.5.2 Protocole d'extraction mis en œuvre

L'extraction par micro-onde a été réalisée à Avignon avec l'aide du Dr Farid Chemat. Une masse exacte de feuilles fraîches (154 g ou 1025 g) de violette de Toulouse a été introduite dans le réacteur de l'appareil EOS-GR milestone. Une puissance équivalente à la masse introduite a été fixée (154 W ou 1025 W). Le dispositif était en mode statique conduisant à un chauffage réparti, mais non optimisé et pouvant provoquer la présence de point chaud. Le régime permanent est atteint quand la température du réacteur avoisine les 93 °C et l'extraction est maintenue jusqu'à obtenir un rendement de 60% de la masse initialement apportée. Les solutions obtenues sont ensuite conservées au congélateur.

IV.5.2.2.6 Fractionnement par SPE

Certains extraits bruts ont été fractionnés sur cartouche SPE C_{18} ec (Spe-Pak® C_{18} cartridge, Waters, Milford, Ma, USA) à l'aide d'une pompe de 1 bar (KNF Neuberger, Freiburg, Allemagne) et d'un Vac elut SPS 24 (Varian), de manière à pouvoir identifier des fractions d'intérêt simplifiées, plus faciles à

travailler par la suite. Les cartouches SPE utilisées sont des cartouches C_{18} de 1 et 10 g dans lesquelles on peut déposer respectivement 100 mg et 1 g d'extrait, généralement dissous dans un minimum de solvant de constitution proche du premier solvant d'élution. Les solvants d'élution au nombre de cinq ont été choisis de polarité décroissante comme suit (Tableau 40):

Tableau 40. Constitution des solvants utilisés lors des fractionnements SPE sur cartouche C_{18}

Solvant	Pour les extraits bruts en général	Pour les cyclotides en particulier
1	eau/méthanol (95:5)	bicarbonate d'ammonium (NH_4HCO_3 à 0,1M et pH 8),
2	eau/méthanol (75:25)	eau/éthanol (80:20)
3	eau/méthanol (50:50)	eau/éthanol (50:50)
4	eau/méthanol (25:75)	eau/éthanol (20:80)
5	méthanol (100)	

Après activation avec un volume de colonne de méthanol et équilibrage avec un volume de colonne du premier solvant, l'extrait est déposé et trois volumes de colonne de chaque fraction ont été élués. Les solutions obtenues ont ensuite été évaporées et le résidu sec obtenu pesé puis resolubilisé à 2 mg/mL pour analyse UHPLC-MS. Ce fractionnement permet de purifier grossièrement l'extrait brut en séparant les molécules très polaires éluées en grande partie avec les premiers solvants, des molécules plus apolaires retenues et essentiellement éluées avec les derniers solvants. Un rinçage avec une solution 50:50 de dichlorométhane/méthanol permet ensuite de nettoyer la cartouche et de récupérer les composés retenus.

IV.5.2.3 Méthodes d'analyse chimique

Une fois les extraits obtenus, ces derniers sont caractérisés à l'aide d'appareils chromatographiques, spectroscopiques et spectrométriques. Une déréplication des métabolites connus est aujourd'hui indispensable pour un gain de temps. En effet, elle permet d'éviter un long processus d'isolement de molécules déjà connues ou d'intérêt limité pour l'étude en cours. Elle met en œuvre deux techniques, la UHPLC-UV-DAD et la UHPLC-MS/MS qui permettent l'obtention d'informations structurelles et spectrales indispensables à l'identification (**Wolfender et al., 2006**). Pour les molécules inconnues, ces techniques couplées à des bases de données permettent d'obtenir des hypothèses de structures très pertinentes avec l'utilisation de la haute résolution qui assure l'obtention d'une masse relayant la composition élémentaire de la molécule, soit une masse exacte. Cependant, une purification sera inévitable pour une identification plus précise par l'intermédiaire de la résonance magnétique nucléaire mono et bidimensionnelle.

IV.5.2.3.1 La chromatographie liquide couplée à la spectrométrie de masse

La chromatographie liquide couplée à un spectromètre de masse (LC-MS) est devenue la technique analytique la plus utilisée pour les mélanges complexes d'extraits naturels (**Wolfender et al., 2015**) et est plus adaptée à l'analyse des composés non volatils ou thermodégradables (**Caude et Jardy, 2016**). Des informations structurales peuvent être obtenues grâce à un couplage à un spectromètre de masse : un ion précurseur est sélectionné puis fragmenté par dissociation induite par collision et les fragments formés sont ainsi analysés en basse ou haute résolution donnant des indications sur la structure de la molécule. Cette scission, nommée fragmentation, peut intervenir à différents endroits de la molécule conduisant ainsi à plusieurs fragments pour une seule structure (**de Hoffmann et Stroobant, 2005**).

C'est cette technique qui a été mise en œuvre au cours de cette thèse au moyen d'un Orbitrap, assurant l'obtention d'une masse exacte, avec une interface électropsray (ESI), une source à pression

atmosphérique. C'est une technique qui a été développée par Fenn (**Fenn et al., 1990**) pour laquelle un important champ électrique, obtenu par différence de potentiel entre 3 et 6kV est appliqué sur un liquide traversant un capillaire. Une accumulation de charges à la surface du liquide est alors provoquée conduisant à la formation de gouttelettes chargées, phénomène nommé nébulisation. Cet appareil hybride est constitué d'une trappe ionique linéaire (LTQ, haute sensibilité) et d'un piège d'ion électrostatique appelé trappe Orbitale à transformée de Fourrier (Orbitrap XL^{TM}, haute résolution et précision en masse). Plusieurs paramètres sont à optimiser afin d'obtenir une bonne séparation des pics à savoir, la longueur de la colonne, la température, le débit du gaz vecteur et la nature de la phase stationnaire.

Le détecteur Corona (Charged Aerosol Detector - CAD) a été employé afin d'obtenir une information quantitative des extraits analysés (**Ligor et al., 2013**). En effet cette technologie de détection est universelle et offre des réponses constantes indépendantes de la structure chimique, pour des résultats quantitatifs exacts en chromatographie en phase liquide. Un flux de gaz positivement chargé entre en collision avec les particules d'analytes. La charge est alors transférée vers les particules. Cette charge est ensuite mesurée par un électromètre générant un signal directement proportionnel à la quantité d'analytes dans l'échantillon.

IV.5.2.3.1.1 Méthodes analytiques UHPLC-CAD-HRMS/MS

Les analyses UHPLC-HRMS/MS ont été réalisées sur un instrument UHPLC-LTQ Orbitrap XL (Ultimate 3000, Thermo Fisher Scientific, Hemel Hempstead, Royaume-Uni) disponible au laboratoire Pharma-Dev. Celui-ci est piloté à l'aide de trois logiciels : Xcalibur version 2.1 (Thermo Fisher Scientific Inc.), Chromelon pour la partie UHPLC et Thermo Tune plus version 2.5.5 pour la partie MS. Deux colonnes Acquity UPLC BEH C_{18} (100/150 x 2,1 mm i.d., 1,7 um, Waters, MA, USA) équipée d'une colonne de garde ont été utilisées pour l'ensemble des analyses de ce projet avec un gradient de 15 ou de 25 min faisant usage de l'eau et d'acétonitrile (ACN) avec 0,1% d'acide formique comme solvant d'élution (Tableau 41). Un détecteur d'aérosols chargés (CAD) a été connecté pour apporter une dimension quantitative aux analyses. Le débit a été fixé à 0,3 mL/min sans CAD et à 0,45 ml/min avec. La température de la colonne a été réglée à 40 °C et le volume d'injection a été fixé à 2 µl sans CAD et 5 µL avec, car l'échantillon est alors distribué entre ce détecteur et la masse. La détection aux ultraviolets (UV) a été réalisée par un détecteur à barrettes de diodes (DAD) de 210 à 400 nm. La détection de masse a été réalisée en utilisant une source d'électronébulisation (ESI) dans les modes d'ionisation positif (PI) et négatif (NI). La température du capillaire a été fixée à 300 °C et la tension de l'électrospray à 4,2 kV en positif et 3 kV en négatif. La plage de balayage de masse était de 100 à 2000 Da. Chaque balayage a été suivi par une fragmentation des trois ions les plus intenses par dissociation induite par collision (CID). L'ensemble des échantillons a été préparé à 2 mg/mL dans le solvant le plus adapté et conservé à 10 °C dans le passeur d'échantillons. L'aiguille est systématiquement rincée avec une solution aqueuse méthanolique 50% entre deux injections.

Tableau 41. Gradients de température mis en œuvre pour les analyses UHPLC-MS

Gradient de 15 min		**Gradient de 25 min**	
Temps (min)	Ratio H_2O*/ACN*	Temps	Ratio H_2O*/ACN*
0	95:5	0	95:5
0,5	95:5	0,5	95:5
12	5:95	22	5:95
15	5:95	25	5:95
15,5	95:5	25,5	95:5
19	95:5	29	95:5

* avec 0,1% d'acide formique

IV.5.2.3.1.2 Méthodes analytiques UHPLC-UV

Des purifications des fractions antioxydantes riches en polyphénols (coumarines et flavonoïdes) ont été entreprises en UPLC-UV à partir d'échantillons concentrés à environ 25 mg/mL et filtrés avant injection. Un système d'autopurification Waters incluant un détecteur UV/vis et un système de pilote basé sur le logiciel MassLynx a été utilisé. Différentes colonnes ont été préalablement testées sur un système UPLC PDA SQD (Waters) piloté par le logiciel Empower dont l'Acquity BEH Shield RP18 (100 x 2,1 mm i.d., 1,7 µm, Waters, MA, USA) qui a présenté, à conditions équivalentes (injection de 5 µL de l'échantillon à 1 mg/mL, débit de 0,5 mL/min, température de colonne de 40 °C et même gradient) une séparation plus fine pour la fraction riche en flavonoïdes et la colonne XBridge prep C_{18} OBD (150 x 19 mm i.d., 5 µm, Waters, MA, USA) pour la fraction riche en coumarines. Les phases mobiles étaient de l'eau et de l'acétonitrile acidifiés à 0,1% en acide formique. Pour le passage en autopurification, le volume d'injection de l'échantillon, préalablement filtré à 0,45 µm, a été fixé à 250 µL avec un débit de 20 mL/min et une température de colonne de 40 °C. La détection UV a été réalisée à 325 nm pour les coumarines et 360 nm pour les flavonoïdes, avec une amplitude de 3 nm. Les fractions ont été automatiquement collectées en remplissant 30% du tube avec un maximum de 30 tubes par injection. Les deux gradients utilisés pour la purification de la fraction riche en coumarines Fr2 et en flavonoïdes Fr3 sont présentés dans le Tableau 42.

Tableau 42 : Gradients mis en œuvre dans la purification des coumarines et des flavonoïdes par UHPLC-UV

Fraction riche en coumarines		**Fraction riche en flavonoïdes**	
Temps (min)	Ratio H_2O/ACN	Temps (min)	Ratio H_2O/ACN
0	95:5	0	93:7
15	75:25	6	88:12
16	50:50	7	86:14
18	50:50	12	86:14
18,5	95:5	14	50:50
25		15	50:50
		15,5	93:7
		18	93:7

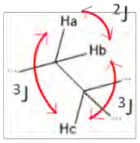
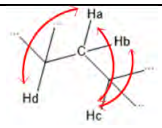
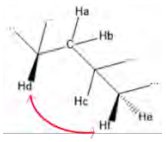
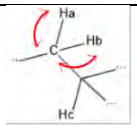
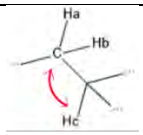
IV.5.2.3.2 Le spectre UV-visible

La spectrométrie UV-visible est basée sur la capacité de certaines molécules d'absorber certaines longueurs d'onde du spectre UV-visible (200-800 nm). La trace UV des composés étant caractéristique des classes chimiques, l'étude de ce paramètre a permis d'apporter des informations complémentaires sur le type de composés chimiques en présence, pouvant permettre de valider les hypothèses de structure déterminées par les bases de données.

IV.5.2.3.3 La résonance magnétique nucléaire

. Des informations globales seront obtenues par l'intermédiaire de la résonnance magnétique nucléaire (RMN) monodimensionnelle ^{1}H et ^{13}C et des informations complémentaires indispensables à l'étude structurale complète seront obtenues par RMN bidimensionnelle (Tableau 43) tels que les couplages $^{13}C/^{1}H$ (HMBC et HMQC) et $^{1}H/^{1}H$ (COSY, TOCSY, ROESY et NOESY).

Tableau 43. Principes des expériences de RMN bidimensionnelles

Méthode	Couplage	Principe	
COSY Correlated spectroscopy	$^{1}H/^{1}H$	Permet d'associer les protons non équivalents de couplage ^{2}J et ^{3}J	
TOSCY Total correlation spectroscopy	$^{1}H/^{1}H$	Permet d'associer les protons couplés mutuellement de façon scalaire dans un même système de spin	
ROESY Rotational overhauser enhancement spectroscopy NOESY Nuclear overhauser enhancement spectroscopy	$^{1}H/^{1}H$ $^{1}H/^{1}H$	Permettent de mettre en évidence la proximité de certains protons dans l'espace par interactions dipolaires. Complémentaires la NOESY est surtout utilisée pour des molécules de hauts poids moléculaires tandis que la ROESY est adaptée aux petites molécules	
HSQC Heteronuclear multiple quantum coherence	$^{13}C/^{1}H$	Permet d'associer les carbones protonés aux hydrogènes qu'il porte (couplage ^{1}J)	
HMBC Heteronuclear multiple bond connectivity	$^{13}C/^{1}H$	Permet d'associer les carbones aux hydrogènes de couplage longue distance ^{2}J, ^{3}J voire plus	

La RMN a été utilisée pour l'analyse structurale des composés purifiés. Les principaux noyaux étudiés ont été le proton ^{1}H et le carbone ^{13}C. Les échantillons ont été dissous dans du DMSO deutéré pour être certain de dissoudre l'intégralité des métabolites et introduits dans des tubes analytiques de 5 mm de diamètre. Les analyses ont été réalisées sur un spectromètre Bruker Advance 500 MHz équipé d'une cryosonde Prodigy pour des raisons de sensibilités plus accrues. Le logiciel Topspin™ (Brucker) assure le contrôle de l'appareil. Des enregistrements RMN mono et bidimensionnelles ont été entrepris. Les déplacements chimiques (δ) sont exprimés en ppm et corrigés par rapport à la référence DMSO (δ_H = 2,49 ppm et δ_C = 39,5 ppm) et les constantes de couplage sont exprimées en Hz. Le traitement des spectres est réalisé à l'aide du logiciel MNova™ (version 12.0.0, Mestrelab Research).

IV.5.2.4 Méthodes d'analyse métabolomique des données UHPLC-HRMS

IV.5.2.4.1 Pour la violette de Toulouse

Les profils UHPLC-MS des fractions de la violette de Toulouse ont été traités avec MS-DIAL version 2.54 comme suit : la fenêtre d'analyse était comprise entre 0 et 12,5 min, les masses détectées entre 100 et 1500 Da et la ligne de base optimale a été fixée à 2.10^4. Les profils ont ensuite été alignés sur celui d'un QC (aliquot de tous les extraits) avec les tolérances respectives sur le temps de rétention et les masses de 0,1 min et 0,025 Da.

IV.5.2.4.2 Pour la collection de violettes

Les données UHPLC-MS de la collection ont été examinées avec MS-DIAL version 2.74 comme suit : la fenêtre d'analyse était comprise entre 0,5 et 14 min, les masses détectées entre 100 et 1500 Da et la ligne de base respectivement fixée à 4.10^5 et $1,6.10^5$ pour NI et PI. Les profils ont ensuite été alignés sur

celui d'un QC avec les tolérances respectives sur le temps de rétention et les masses de 0,1 min et 0,025 Da.

IV.5.2.4.3 Pour les essais d'extraction

Les données UHPLC-MS des extraits obtenus par les différentes extractions testées ont été étudiées uniquement en mode NI avec MS-DIAL version 2.56 comme suit : la fenêtre d'analyse était comprise entre 0,5 et 14 min, les masses détectées entre 100 et 1500 Da et la ligne de base fixée à 9.10^4. Les profils ont ensuite été alignés sur celui d'un QCtot avec les tolérances respectives sur le temps de rétention et les masses de 0,1 min et 0,025 Da.

IV.5.2.4.4 Identification des métabolites

Les résultats de l'analyse de masse, c'est-à-dire la masse exacte de chaque composé ainsi que leur spectre de masse expérimental, ont été nécessaires au processus d'identification. En effet, à l'aide du logiciel MS-FINDER, la formule brute a pu être déterminée par l'intermédiaire de la masse exacte, tandis que des hypothèses de structure ont pu être obtenues par comparaison des spectres de masse expérimentaux à ceux de références contenues dans les différentes bases de données de produits naturels à disposition, tels que le dictionnaire des produits naturels (DNP), KNApSAcK et PlantCyc. La création de bases de données spécifiques du genre *Viola* a permis d'apporter des identifications, bien que potentielles, plus pertinentes que celles obtenues par des bases de données généralistes.

IV.5.2.5 Méthodes de construction des réseaux moléculaires

Le fichier texte exporté depuis MS-DIAL est nettoyé en éliminant les adduits identifiés avant importation dans MetamapR (version 1.4.0). Le réseau de similarité spectrale est créé avec un maximum de dix connexions entre les nœuds, un « cosine score » fixé à 0,3 et un temps de rétention filtré à 1,5 min ont été établis pour accepter une connexion entre deux nœuds. Le réseau de corrélation est créé à l'aide du modèle Spearman et d'une valeur *p* relative au taux de fausses découvertes (false discovery rate – FDR) définie à 5.10^{-6}. La liste des connexions a ensuite été téléchargée et traitée avec Cytoscape 2.8.3. Un fichier d'attributs contenant toutes les informations disponibles, en particulier les valeurs m/z, coefficients OPLS et classes chimiques des variables ont été importées pour améliorer la visualisation du réseau. L'intensité moyenne de chaque pic est ajoutée pour spécifier la proportion relative de chaque composé par espèce, en utilisant le plugin «Color Nodes».

IV.5.2.6 Méthodes de dosage des polyphénols et cyclotides

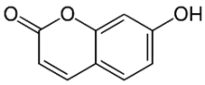
Le dosage des polyphénols des divers extraits de violette a était réalisé dans un premier temps par UHPLC-UV en se basant sur les spectres UV caractéristiques des deux classes chimiques d'intérêt : les coumarines et les flavonoïdes. A l'aide du logiciel Thermo Xcalibur version 3.0.63 et en se basant sur le spectre d'un QC (pool des extraits), une méthode de détection automatique des pics correspondant à une de ces classes a été réalisée en indiquant les temps de rétention. Elle a ensuite été automatiquement appliquée à l'ensemble des extraits analysés en UHPLC-MS afin de déterminer les aires de chaque pic d'intérêt. Des droites de calibration ont été réalisées à partir de l'analyse de six concentrations de deux témoins : l'umbelliferone pour les coumarines et la rutine pour les flavonoïdes, afin de pouvoir déterminer la concentration de chaque composé d'intérêt à partir de l'aire du pic.

Une méthode plus globale a été mise en œuvre sur quelques échantillons par la méthode Folin-Ciocalteu ainsi qu'un dosage colorimétrique des flavonoïdes par la méthode $AlCl_3$. Enfin, la teneur en cyclotides a été déterminée par la méthode de l'acide bicinchoninique.

IV.5.2.6.1 Dosage UHPLC-UV des coumarines

La droite d'étalonnage de l'umbelliferone est présentée Figure 129. Le coefficient de régression égal à 0,9998 atteste bien de la linéarité entre la concentration et l'aire sous le pic dans le domaine d'étude.

C (mg/mL)	Aire
0,0005	19683,6
0,0050	225802
0,0090	465148,4
0,0250	1248843,4
0,0500	2387282,6
0,1000	4877043



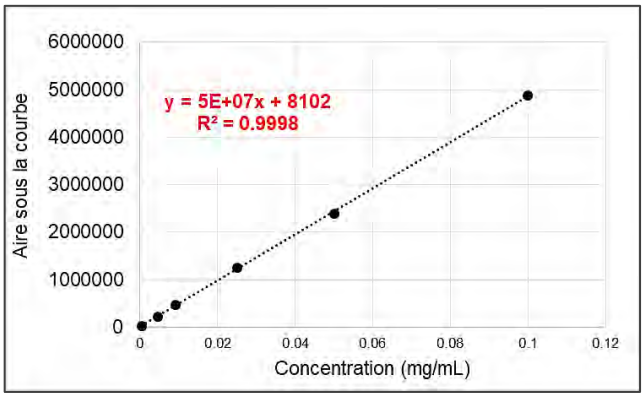
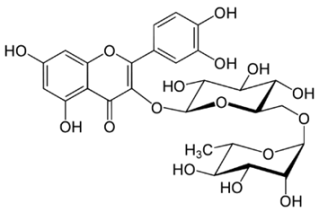


Figure 129. Droite de calibration UHPLC-UV de l'umbelliferone

IV.5.2.6.2 Dosage LC-UV des flavonoïdes

La droite d'étalonnage de la rutine est présentée Figure 130. Le coefficient de régression égal à 1 atteste bien de la linéarité entre la concentration et l'aire sous le pic dans le domaine d'étude.

C (mg/mL)	Aire
0,0005	3719
0,0050	28513
0,0090	56854
0,0250	158353
0,0500	312869
0,1000	629518



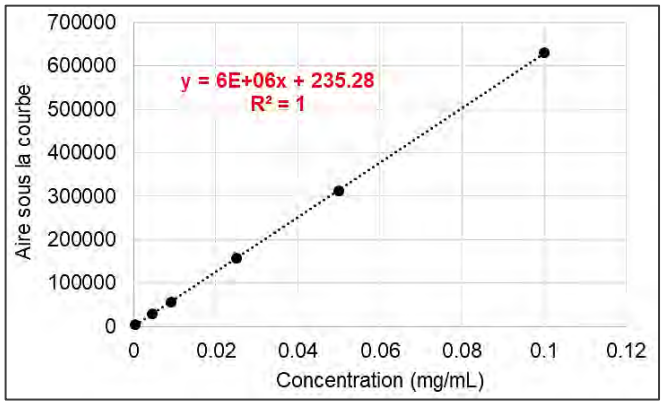


Figure 130. Droite de calibration UHPLC-UV de la rutine

Les résultats des calibrations permettent d'effectuer un dosage spécifique des coumarines et des flavonoïdes, basé sur les aires des pics et sur les spectres UV pour l'identification des classes chimiques (Figure 131). Les coumarines sont caractérisées par une bande d'absorption vers 325 nm et les flavonoïdes vers 360 nm, avec une certaine variabilité autour de ces valeurs moyennes selon la substitution des aglycones.

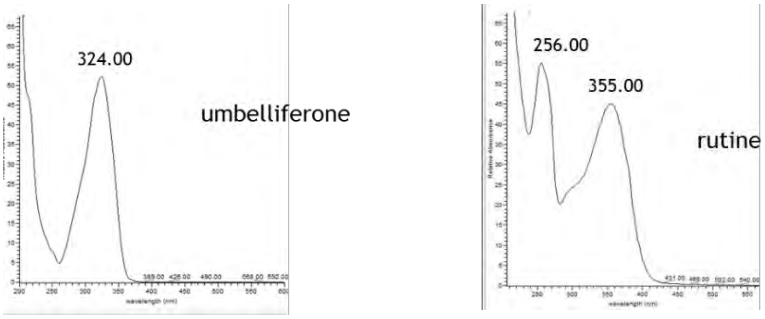


Figure 131. Spectres UV caractéristiques des coumarines (umbelliferone) et des flavonoïdes (rutine)

IV.5.2.6.3 Dosage Folin-Ciocalteu des polyphénols totaux

Les polyphénols totaux ont été déterminés par la méthode Folin-Ciocalteu adaptable en plaque 96 puits, sur le modèle du protocole établi par Zhang et ses collègues (**Zhang et al., 2006**). Ce dosage repose sur la réduction du réactif Folin-Ciocalteu, constitué d'acide phosphotungstique ($H_3PW_{12}O_{40}$) et d'acide phosphomolybdique ($H_3PMo_{12}O_{40}$), en milieu alcalin par les groupements OH des polyphénols conduisant à la formation d'oxydes bleus de tungstène (W_8O_{23}) et molybdène (Mo_8O_{23}). Ainsi, 20 µL d'extrait à 1 mg/mL ou de blanc ou de standard, ont été introduits dans chaque puit de la plaque, puis ont été ajoutés 100 µL de réactif Folin-Ciocalteu. La plaque a été mise sous agitation modérée avant de la laisser au repos 5 min. Ont ensuite été additionnés 80 µL d'une solution de carbonate de sodium (Na_2CO_3) à 7,5%, puis le mélange a été de nouveau homogénéisé sous agitation. La plaque est ensuite incubée dans le noir et une cinétique d'absorbance à température ambiante a été réalisée sur 2h10 à la longueur d'onde de 750 nm. Une gamme d'acide gallique (de 0 à 50 µg/mL) a été effectuée et les résultats du dosage sont présentés en mg équivalent d'acide gallique /mg d'extrait brut (Figure 132).

C (µg/mL)	Absorbance
50	2,97
25	1,65
12,5	0,93
6,25	0,52
3,125	0,31
1,5625	0,19
0,78125	0,14
0	0,09

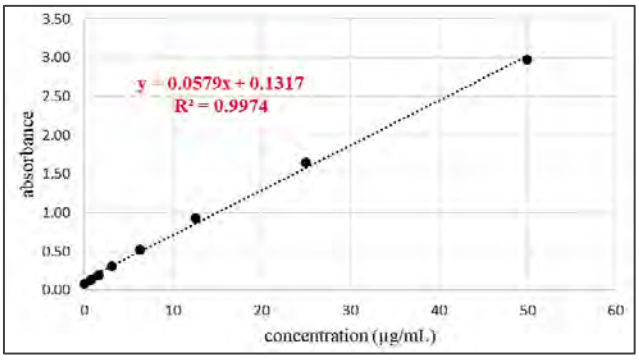


Figure 132. Droite de calibration de l'acide gallique

IV.5.2.6.4 Dosage $AlCl_3$ des flavonoïdes

Les flavonoïdes totaux ont été déterminés par la méthode $AlCl_3$. Ce dosage repose sur la formation de complexes stables de couleur jaune entre les fonctions hydroxyles des flavonoïdes et le chlorure d'aluminium ($AlCl_3$). Ainsi, 100 µL d'extrait à 0,2 mg/mL ou de blanc ou de standard, ont été introduits dans chaque puit de la plaque, puis ont été ajoutés 100 µL de solution d'$AlCl_3$ à 2%. La plaque a été mise sous agitation modérée avant de la laisser au repos 30 min à température ambiante. La mesure de l'absorbance a ensuite été prise à 430 nm. Une gamme de rutine (de 0 à 50 µg/mL) a été effectuée et les résultats du dosage sont présentés en mg équivalent de rutine /mg d'extrait brut (Figure 133).

C (µg/mL)	Absorbance
50	0,79
25	0,45
12,5	0,28
6,25	0,18
3,125	0,14
1,5625	0,11
0,78125	0,10
0	0,09

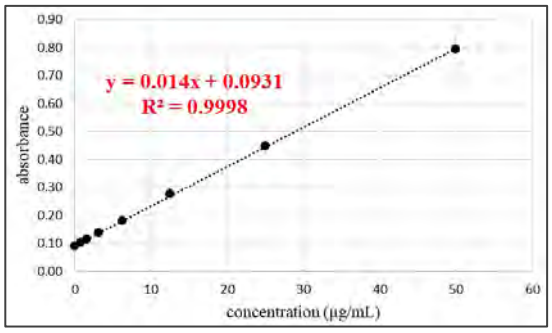


Figure 133. Droite de calibration de la rutine

IV.5.2.6.5 Dosage BCA des cyclotides

La méthode BCA (pour BiCinchoninic acid Assay) est un dosage colorimétrique des protéines, basé sur l'acide bicinchoninique. En milieu alcalin, les protéines réduisent alors l'ion cuivrique Cu(II) en Cu(I). L'acide bicinchoninique forme un complexe pourpre avec Cu(I) d'absorption maximale à 562 nm. Ce dosage est utilisé dans la quantification des cyclotides mais n'est pas spécifique car il assure également le dosage des acides aminés. Une gamme de la protéine BSA (albumine de sérum bovin) a été réalisée entre 0 et 100 µg/mL à partir d'une solution mère à 0,1% dans une solution tampon de NaCl à 0,15 M (Figure 134). Ainsi, 100 µL de chaque dilution ont été ajoutés à 900 µL d'eau. Pour les échantillons, ce sont 20 µL qui ont été complétés à 1mL avec de l'eau. Sont ensuite introduits 2 mL de réactif préparé avec 50 mL de BCA et 1 mL de $CuSO_4,5H_2O$ à 4%. Après incubation à 37 °C pendant 30 min, les DO sont mesurées à 562 nm.

C (µg/mL)	Absorbance
0	0
20	0,061
40	0,104
60	0,159
80	0,211
100	0,256

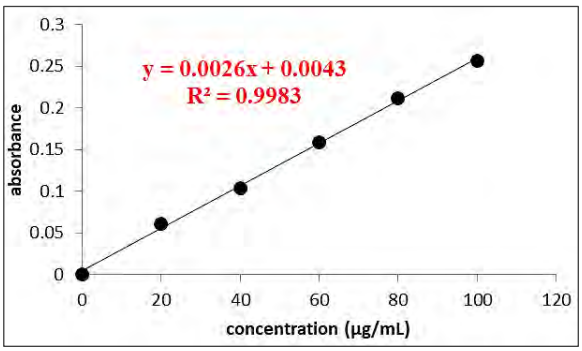


Figure 134. Droite de calibration de la protéine BSA

IV.6 METHODES DE DÉTERMINATION DES POTENTIELS BIOLOGIQUES DES EXTRAITS

IV.6.1 ÉVALUATION DE L'ACTIVITE ANTIOXYDANTE

L'un des objectifs de ce travail a été de déterminer le potentiel antioxydant des extraits de violette dans une optique de sélection d'extraits et de valorisation dermo-cosmétique. En effet, une forte activité antioxydante d'un extrait sera indicative d'une variété et/ou d'une richesse moléculaire. Deux principales méthodes ont été mises en œuvre pour le déterminer sur des extraits complexes et des molécules purifiées.

IV.6.1.1 Méthode DPPH

Le 2,2-diphényl-1-picryl-hydrazyl (DPPH$^{\bullet}$) est un radical organique stable présentant une absorption caractéristique à la longueur d'onde 517 nm en solution éthanolique. En milieu polaire cette longueur d'onde est ajustée à 523 nm et à 519 nm en milieu apolaire. En présence d'un antioxydant donneur de proton type polyphénol, ce radical est réduit (Figure 135), entrainant ainsi une perte de son absorbance à la longueur d'onde de travail (**Nguyen et al., 2013**). Cette technique n'est pas spécifique, mais permet de déterminer la réserve totale d'antioxydants hydrophiles d'un extrait.

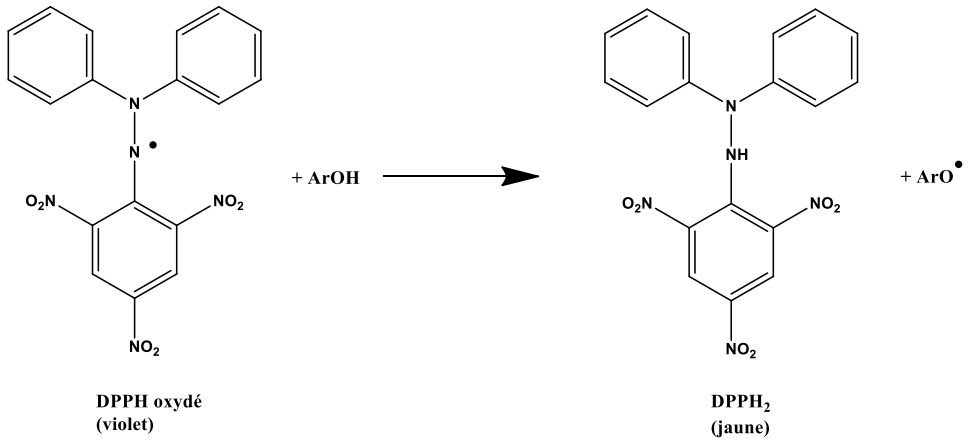


Figure 135. Réduction du radical DPPH en présence d'un antioxydant donneur de proton

Par conséquent, la capacité antioxydante d'un extrait peut être déterminée en calculant le pourcentage d'inhibition de l'absorbance du DPPH$^{\bullet}$. Celle-ci est mesurée après trente minutes de réaction à l'abri de la lumière. Elle est exprimée de la manière suivante (A = absorbance) :

$$\%\ inhibition = \frac{A_{contrôle} - A_{test}}{A_{contrôle}} * 100$$

A partir de cette information, la concentration inhibitrice à 50% (CI_{50}), qui représente la concentration massique des extraits diminuant de 50% l'absorbance du DPPH$^{\bullet}$, peut être obtenue à l'aide de la régression linéaire de la droite d'équation suivante : $\%\ inhibition = f(concentration\ du\ mélange\ testée)$.

Par comparaison, l'extrait présentant la valeur de CI_{50} la plus faible possèdera la réserve antioxydante la plus importante. Le Trolox est généralement la référence utilisée pour valider les expériences. Pour un même extrait, cinq dilutions (5, 50, 266,67, 500 et 1000 mg/mL) ont été préparées dans des tubes eppendorfs à un volume final de 1 mL à partir d'une solution mère de 2 mg/mL avec une solution éthanolique à 80% préalablement désoxygénée à l'azote. Chacune de ces dilutions a été analysée en triplicata comme présenté dans le plan de plaque 96 puits Tableau 44.

Tableau 44. Plan de plaque pour les analyses DPPH

	1	2	3	4	5	6	7	8	9	10	11	12
	Analyse Echantillon1				Analyse Echantillon2				Analyse Echantillon3			
A	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle
B	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
C	Ech1 C_1	Ech1 C_1	Ech1 C_1	Ech1 C_1	Ech2 C_1	Ech2 C_1	Ech2 C_1	Ech2 C_1	Ech3 C_1	Ech3 C_1	Ech3 C_1	Ech3 C_1
D	Ech1 C_2	Ech1 C_2	Ech1 C_2	Ech1 C_2	Ech2 C_2	Ech2 C_2	Ech2 C_2	Ech2 C_2	Ech3 C_2	Ech3 C_2	Ech3 C_2	Ech3 C_2
E	Ech1 C_3	Ech1 C_3	Ech1 C_3	Ech1 C_3	Ech2 C_3	Ech2 C_3	Ech2 C_3	Ech2 C_3	Ech3 C_3	Ech3 C_3	Ech3 C_3	Ech3 C_3
F	Ech1 C_4	Ech1 C_4	Ech1 C_4	Ech1 C_4	Ech2 C_4	Ech2 C_4	Ech2 C_4	Ech2 C_4	Ech3 C_4	Ech3 C_4	Ech3 C_4	Ech3 C_4
G	Ech1 C_5	Ech1 C_5	Ech1 C_5	Ech1 C_5	Ech2 C_5	Ech2 C_5	Ech2 C_5	Ech2 C_5	Ech3 C_5	Ech3 C_5	Ech3 C_5	Ech3 C_5
H	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle	Contrôle

Chaque puit a été rempli avec 200 µL de dilution de l'échantillon ou de référence Trolox (Fluka, pureté > 98%) ou de contrôle (solvant seul), puis 100 µL de solution de DPPH à 300 µM ont été ajoutés colonne par colonne à l'aide d'une pipette multicanaux. L'absorbance de l'extrait a été vérifiée en remplaçant dans une colonne le DPPH par l'éthanol à 80%. Ainsi, si l'échantillon possède une absorbance à la longueur d'onde de travail, elle pourra être prise en compte dans les calculs de CI_{50}. L'absorbance du milieu a ensuite été mesurée à l'aide d'un lecteur de plaque EON à 517 nm après 30 minutes de réaction dans le noir. La valeur de CI_{50} est déterminée graphiquement pour une régression linéaire comme suit: $CI_{50} = 0{,}5 - \frac{ordonnée\ à\ l'origine}{pente}$. Visuellement, un extrait antioxydant conduira à l'observation d'une couleur jaune dans le puit comme c'est le cas pour la référence Trolox, tandis que pour les extraits non antioxydants, une couleur violette sera remarquée, comme pour le contrôle solvant (Figure 136).

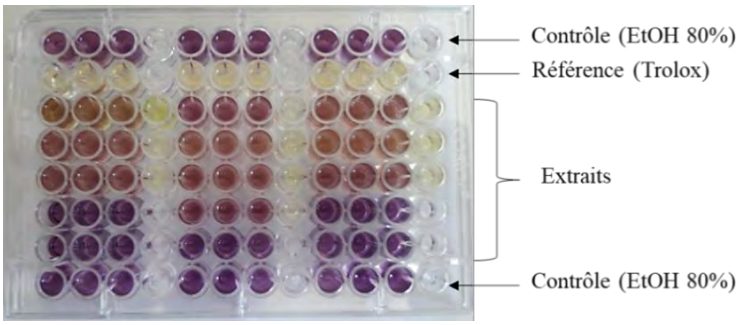


Figure 136. Exemple de plaque d'analyse DPPH d'extraits de violettes

Pour le criblage antioxydant de la collection, un protocole haut débit a été proposé. Une partie du surnageant obtenu après extraction et purification de la chlorophylle a été mis de côté et une dilution au 20ième a été réalisée. Pour cette étude, la CI_{50} n'a pas été déterminée en raison de la charge de travail, mais seul le potentiel antioxydant a été étudié à partir d'extraits strictement préparés sur le même protocole. Une simple comparaison des absorbances a donc été faite pour comparer les potentiels de chaque violette.

IV.6.1.2 Méthode de résonance paramagnétique électronique

La capacité des molécules à réduire le radical superoxyde ($O_2^{\bullet -}$) a été étudiée par résonance paramagnétique électronique (RPE), également nommée résonance de spin électronique (ESR) (**Hubert et al., 2008**). La formation de radicaux peut être directement détectée par RPE si la durée de vie est suffisante. Dans le cas contraire, une méthode de piégeage de ces radicaux appelée « spin-trapping » est utilisée. Pour cela, il faut disposer d'un composé capable de capter ces radicaux dans le milieu en formant un adduit de durée de vie plus longue et facile à détecter. Les nitrones sont souvent généralement les molécules d'intérêt de par la présence d'un carbone électrophile, lié à la polarisation de la liaison nitrone, qui favorise l'addition des radicaux et la formation d'adduits résistants et donc détectables par RPE (**Finkelstein et al., 1980**). Le DMPO (5,5-diméthyl-1-pyrroline-*N*-oxyde) est souvent utilisé (Figure 137 a) et forme avec le radical superoxyde un spectre très caractéristique de type triplet dédoublé avec des constantes de couplage A_N = 12 G, $A_{H\alpha}$ = 10,7 G et $A_{H\gamma}$ = 1,5 G. Une allure de faux quadruplet est distinguée en raison de la similitude des valeurs de constantes (Figure 137 b).

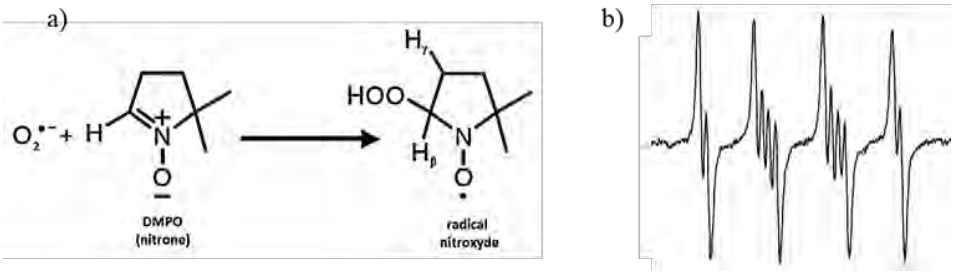


Figure 137. Piégeage du radical superoxyde $O_2^{\bullet -}$ par le DMPO (a) et spectre caractéristique de l'adduit DMPO-OOH (b)

Chaque extrait a été préparé de manière à obtenir une concentration finale à 50 mg/L. Dans un eppendorf, 320 µL de solution de diméthylsulfoxyde (DMSO) ont été ajoutés suivi de 100 µL d'eau distillée préalablement désoxygénée à l'azote, 25 µL d'extrait à 1 mg/mL, 50 µL de DMPO, puis pour finir 5 µL de solution d'éther de couronne saturée en dioxyde de potassium (KO_2). Après passage rapide au vortex, un capillaire de 50 µL a été rempli de ce mélange réactionnel puis introduit dans l'aimant. L'analyse a été lancée précisément deux minutes après la fin du mélange sur un spectromètre Brucker EMX 8/2,7. Le spectre correspondant au signal du DMPO superoxyde (DMPO-OOH) a été enregistré à l'aide de l'instrument Magnettech Miniscope MS100 dans la configuration suivante : centre du champ à 3376 mG, largeur du spectre de 60 mG, durée d'enregistrement de 30 s, modulation de l'amplitude à 1000 mG, atténuation à 5 db et le gain à 8.10^2. Les résultats ont ensuite été analysés à l'aide du logiciel WinEPR : après avoir effectué une correction de la ligne de base, une intégration du second ordre des différents pics observés a été effectuée. Pour cette étude, le potentiel antioxydant est déterminé à partir de l'intégration du signal obtenu et plus celle-ci est petite, plus l'extrait présente un fort potentiel antioxydant. Pour cette analyse, la référence choisie a été la rutine (Sigma, pureté > 94%), puisque le Trolox, utilisé en DPPH, n'était pas adapté comme en témoignent leurs signaux respectifs comparés au blanc DMSO (Figure 138).

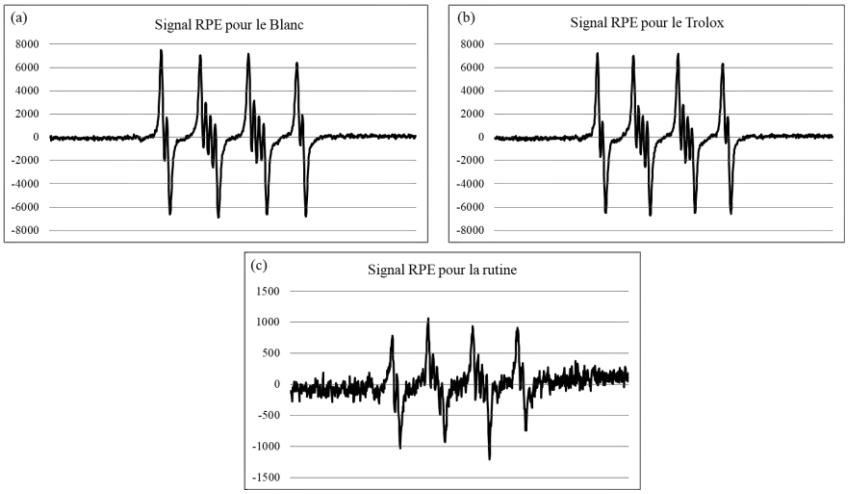


Figure 138. Signaux RPE du blanc solvant DMSO (a), du Trolox (b) et de la rutine (c)

IV.6.2 ÉVALUATION DE LA CYTOTOXICITE SUR CELLULES VERO

La recherche de principes actifs étant destinés à l'industrie cosmétique, des analyses de cytotoxicité ont été réalisées afin d'étudier la potentielle toxicité des extraits. Pour cela, des tests sur cellules VERO non cancéreuses (Sigma-Aldrich) ont été effectués. La lignée cellulaire Vero a été isolée à partir de rein de singe vert adulte d'Afrique (Certopithecus aethiops) en mars 1962 par Y. Yasumura et Y. Kawakita à l'université de Chiba au Japon. Ce nom a été choisi par le Dr. Yasumura et provient de la langue

Esperanto : « Vero » signifiant la vérité, mais est également l'acronyme de « Verda » et de « Reno » (respectivement vert et rein).

IV.6.2.1 Repiquage des cellules

Les cellules Vero prolifèrent relativement rapidement, c'est pourquoi trois repiquages par semaine sont réalisés à 48 h d'intervalle afin de ne pas saturer le milieu et conduire à la mort des cellules. Ces manipulations sont réalisées en conditions stériles sous poste de sécurité microbiologique (PSM).

Le milieu de culture est une solution de MEM (Milieu Essentiel Minimum) contenant 50 mL de sérum de veau (SVF) soit 10%, 5 mL d'acides aminés non essentiels (AANE) et 5 mL d'antibiotiques (100 U/mL de pénicilline et 100 µg/mL streptomycine). Le milieu de culture contenu dans une flasque de 25 mL est enlevé puis 5 mL de PBS (Phosphate Buffered Saline) sont ajoutés afin d'éliminer toute trace de sérum restant. La solution est jetée et 2 mL de trypsine sont ajoutés afin de décoller les cellules avant incubation 10 min à 37 °C. Par la suite 8 mL de milieu MEM sont ajoutés et bien répartis sur l'ensemble de la flasque par système d'aspiration/reflux. Un quart de cette solution, soit 2,5 mL, est prélevé puis introduit dans une nouvelle flasque contenant 17,5 mL de milieu MEM. La flasque est ensuite placée à incuber à 37 °C et 5% de CO_2 pendant 48 h.

IV.6.2.2 Dénombrement des cellules

Le dénombrement des cellules est une étape essentielle afin de préparer la bonne dilution permettant d'ensemencer le nombre de cellules souhaité par puit. Pour cela, après avoir trypsinisé les cellules et ajusté le volume à 10 mL dans un flacon, une centrifugation à environ 800 rpm pendant 4 min est réalisée afin de culotter les cellules. Le milieu est ensuite éliminé puis de nouveau 10 mL sont ajoutés et les cellules sont remises en suspension. Une dilution au demi est ensuite réalisée avec une solution de Trypan Blue Dye 0,40% (Bio Rad) puis 20 µL sont introduits dans un compartiment de la plaque de lecture, de manière à occuper l'ensemble du volume sans surcharger afin de ne pas créer un mouvement du fluide pouvant faire se mouvoir les cellules et compliquer le comptage. Pour le dénombrement, cinq carrés sont analysés et si les cellules sur le côté droit et un haut sont comptées, alors celles du bas et à gauche ne le sont pas (ou inversement) (Figure 139). Une moyenne est ensuite calculée permettant de remonter à la quantité de cellules dans les 10 mL de départ. La dilution adaptée pour l'ensemencement de 5 000 cellules par puit pour le nombre de plaques souhaité en est ensuite déduite.

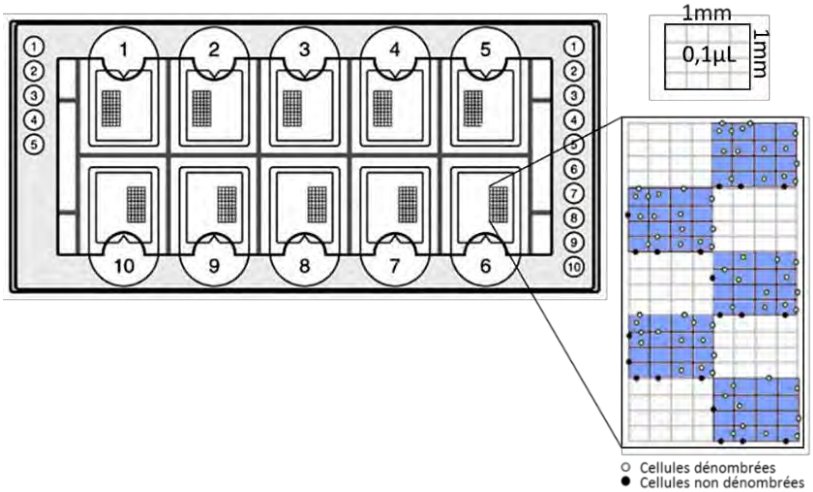


Figure 139. Plaque et grille de dénombrement des cellules. La grille comporte 10 carrés subdivisés en 10 secteurs.

IV.6.2.3 Test de cytotoxicité

Les cellules sont déposées en plaque 96 puits en ajoutant 100 µL de la dilution adaptée par puit tel que chacun d'entre eux contienne 5 000 cellules. La plaque est ensuite incubée à 37 °C et 5% de CO_2 pendant 24 h. Il est à noter que les extrémités de la plaque ne sont pas utilisées car des effets de bords sont observés et peuvent influencer les résultats. Ensuite, 100 µL d'extrait dilué dans le DMSO stérile à diverses concentrations ou du blanc ou du témoin positif (Doxorubicine, Sigma Aldrich) sont ajoutés dans la configuration du plan de plaque Tableau 45. La plaque est de nouveau incubée pendant 72 h. Chaque concentration est analysée en triplicata. Pour la majorité des plaques, une gamme de doxorubicine en trois points est ajoutée et pour la première et la dernière plaque analysées, une gamme en 6 points est faite afin de valider les résultats par le calcul de la CI_{50} de cette référence. Par ailleurs, une gamme DMSO a été réalisée afin de vérifier sa non-influence sur la croissance des cellules aux concentrations utilisées.

Tableau 45 : Plan de plaque pour les analyses de cytotoxicité

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B		Ech1 C1	Ech1 C1	Ech1 C1	Ech3 C1	Ech3 C1	Ech3 C1	Ech5 C1	Ech5 C1	Ech5 C1	Doxo C1	
C		Ech1 C2	Ech1 C2	Ech1 C2	Ech3 C2	Ech3 C2	Ech3 C2	Ech5 C2	Ech5 C2	Ech5 C2	Doxo C2	
D		Ech1 C3	Ech1 C3	Ech1 C3	Ech3 C3	Ech3 C3	Ech3 C3	Ech5 C3	Ech5 C3	Ech5 C3	Doxo C3	
E		Ech2 C1	Ech2 C1	Ech2 C1	Ech4 C1	Ech4 C1	Ech4 C1	Ech6 C1	Ech6 C1	Ech6 C1	blanc	
F		Ech2 C2	Ech2 C2	Ech2 C2	Ech4 C2	Ech4 C2	Ech4 C2	Ech6 C2	Ech6 C2	Ech6 C2	blanc	
G		Ech2 C3	Ech2 C3	Ech2 C3	Ech4 C3	Ech4 C3	Ech4 C3	Ech6 C3	Ech6 C3	Ech6 C3	blanc	
H												

La croissance cellulaire globale est généralement mesurée au moyen du test colorimétrique MTT (3[4,5-diméthylthiazol-2yl]-bromure de diphenyltétrazolium) (**Mosmann, 1983**). Ce test est une méthode rapide de numérisation des cellules vivantes. Le MTT est alors réduit en formazan par la succinate déshydrognéase mitochondriales des cellules vivantes, formant un précipité de couleur violette. La quantité de précipité formée est alors proportionnelle à la quantité de cellules vivantes. Les puits sont analysés au microscope pour détecter la présence éventuelle d'un précipité et vérifier la forme et l'abondance des cellules. Après avoir retiré le surnageant et laver au PBS, 50 µL de MTT à 0,5 mg/mL dans du PBS sont ajoutés dans chaque puit. Les plaques sont incubées pendant 3 h à 37 °C. Sont ajoutés finalement 100 µL de DMSO 100% afin de dissoudre les cristaux de formazan formés. Après 5 min, les plaques sont vigoureusement agitées à 300 rpm pendant 30 secs et l'absorbance mesurée à 570 nm à l'aide d'un spectrophotomètre EON.

IV.6.3 ÉVALUATION DES ACTIVITES PHYTOSANITAIRES

IV.6.3.1 Activité antifongique

Les maladies des plantes sont causées par différents facteurs dont le stress abiotique tels que la sécheresse, le manque de minéraux et la mise en jeu d'autres organismes comme les insectes, les virus, les bactéries ou les champignons. Ces derniers ont colonisés les plantes pour leur réserve constante de nourriture en tant que pathogènes ou de manière symbiotique. Les plantes ont alors développé plusieurs niveaux de défense pour lutter contre les infections des champignons phytopathogènes. Les principales molécules suspectées de conférer l'activité phytosanitaires sont les cyclotides. L'enrichissement en cyclotides est obtenu avec une solution méthanolique à 50%. Cette extraction a donc été réalisée et les extraits secs ont été solubilisés à 10 mg/ml.

IV.6.3.1.1 Préparation des souches

Les souches sont tout d'abord cultivées sur milieu spécifique avant sporulation. Les protocoles respectifs de chacune d'entre elles sont présentés ci-après :

- Pour *Colletotrichum higginsianum*

La culture est effectuée sur milieu Mathur constitué de 2,8 g/l de glucose, 1,22 g/l de $MgSO_4$, 2,72 g/l de KH_2PO_4, 2,18 g/l de peptone bactériologique (Peptone mycologique Oxoid™, ThermoFisher) et 30 g/l d'agar pour gélifier le milieu. On prélève ensuite les spores âgées maximum de 15 jours dans 10 ml d'eau stérile avant de les placer dans un pilulier stérile. Les spores sont par la suite dénombrées sur lame de Fuchs Rosenthal en déposant 12 à 13 µl de spores en haut de la lamelle. La moyenne des spores comptées sur deux grands carrés est calculée afin de remonter à la concentration dans les 10 mL (en multipliant par 5000). Une solution en spores équivalente à 10^6 spores/ml est ensuite préparée puis on ensemence 1 ml de cette suspension dans cinq erlenmeyers de 250 ml contenant 100 ml de milieu Mathur avant incubation à l'obscurité dans une étuve à 23 °C.

- Pour *Phytophthora capsici*

La culture est effectuée sur milieu V8 5% constitué de 5% de V8, 95 % d'H_2O et 20 g/l d'agar. Le pH du milieu doit être ajusté à 5 avec de la soude. Pour la sporulation, on coupe en deux la gélose de milieu V8 (5%) en utilisant une culture de 10 jours maximum (voire idéalement 7 jours) puis on la transfère dans une boite stérile. La gélose est recoupée en morceaux de 0,75 x 0,75 cm de manière à générer un stress physique. Sont ensuite ajoutés 10 ml d'eau stérile avant d'incuber 1h à température ambiante. L'eau est ensuite renouvelée avant incubation pendant deux jours dans l'étuve lumineuse à 22 °C pour améliorer le rendement. Un choc de température est ensuite effectué en plaçant la boite 1h précise à 4 °C pour libérer les spores

- Pour *Botrytis cinerea*, *Fusarium graminearum* et *Alternaria brassicicola*

Leur culture s'effectue sur milieu dextrosé à la pomme de terre PDA (Potato Dextrose Agar) contenant 4,0 g/l d'extrait de pomme de terre, 20 g/l de glucose et 15g/l d'agar. Les spores âgées maximum de 15 jours sont prélevées dans 10 mL d'eau stérile pour les placer dans un pilulier stérile. Dans le cas de *Botrytis*, la boîte de culture aura été placée au minimum 72h à la lumière avant récolte.

IV.6.3.1.2 Test antifongique en milieu liquide

L'activité antifongique *in vitro* des extraits de violettes a tout d'abord été évaluée par essais microspectrophotométriques avec des plaques ELISA stériles à fond plat transparent. Pour cela les spores de la souche à tester collectées dans 10 ml d'eau stérile sont dénombrées sur cellules de Fuchs-Rosenthal en microscopie afin de déterminer leur concentration. On réalise ensuite 10 ml de suspension de spores à environ 5000 spores/ml dans du PDB à 5,3 g/L pour toutes les souches sauf *Phytophthora capsici* pour laquelle on utilise du milieu V8. Chaque puit de la plaque est ensuite rempli avec 90 µl de cette suspension à laquelle sont ajoutés 10 µl d'extrait à tester, pour un volume final de 100 µl par puit. Ainsi, une dilution au 10ème de l'extrait à 10 mg/mL est analysée. Un témoin négatif est intégré à chaque plaque en ajoutant 10 µL d'eau stérile. Plusieurs mesures de DO à 595 nm sont ensuite relevées à l'aide d'un lecteur de plaque au temps initial (T0), puis 3 jours (T3) et 5 jours (T5) après. Entre temps les plaques sont placées à 22 °C dans l'obscurité. Chaque extrait est analysé en triplicata et une CI_{50} est calculée afin de mettre en évidence les violettes présentant des activités pertinentes. Un premier criblage de la collection a été réalisé sur les souches *Botrytis cinerea, Colletotrichum higginsianum* et *Fusarium*

graminearum. Les violettes ayant présenté une activité d'intérêt ont ensuite été analysées sur *Alternaria brassicicola* et *Phytophthora capsici*.

IV.6.3.2 Activité de défense des plantes

Le gène PR1 (pour Pathogenesis related 1) est un bon marqueur des réponses de défense chez les plantes car il est induit par de nombreux stress comme l'acide salicylique, l'éthylène ou encore toutes attaques de microorganismes phytopathogènes (**Ohashi et Ohshima, 1992**). Pour savoir si la plante répond à un stress, des plantes modifiées sont utilisées : elles expriment la région régulatrice de ce gène, appelé le promoteur, couplé à un gène marqueur, le gène GUS responsable de la synthèse de la β-glucuronidase qui va réagir avec son substrat le X-Gluc (5-Bromo-4-chloro-3-indolyl-β-D-Glucuronic acid) pour conduire à la formation d'un précipité bleu (Figure 140). L'activité Gus est donc détectée par accumulation du produit de réaction bleu lors de l'hydrolyse du substrat X-Gluc par la β-Glucuronidase lors de la transcription du gène PR1. *Arabidopsis thaliana* est généralement la plante utilisée car elle est facile à transformer de façon stable, sa croissance est rapide et elle est suffisamment petite pour entrer dans des plaques 48 puits, adaptée aux analyses haut débit.

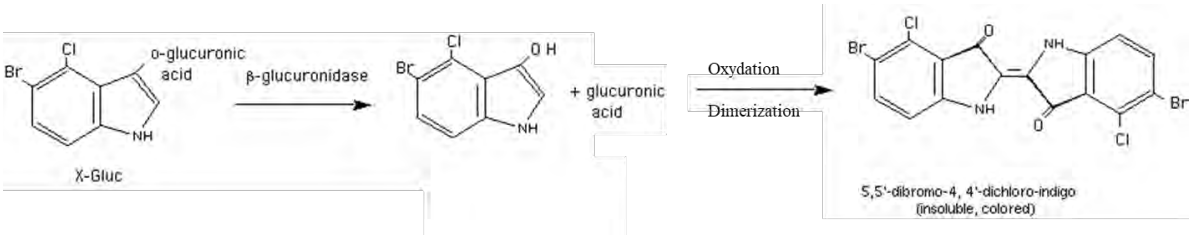


Figure 140. Clivage du X-Gluc par la β-glucuronidase conduisant à la production du chloro-bromoindigo se dimérisant en dichloro-dibromoindigo de couleur bleu

Les graines d'*Arabidopsis thaliana* sont trempées pendant 10 minutes dans une solution de stérilisation contenant 1 volume d'hypochlorite de sodium (NaClO) à 2,6 %, 5 volumes d'éthanol 100 % et quelques gouttes de triton. Deux lavages rapides à l'éthanol 95 % suivent avant séchage quelques heures sous PSM, jusqu'à ne plus observer d'agglomération des graines. Après cette stérilisation en surface, les graines sont transférées, à l'aide d'une pointe de cône préalablement stérilisée au bec bunsen, dans des puits de plaques 48 puits, contenant 500 µl de milieu MS (Murashige et Skoog, Sigma-Aldrich) supplémentés avec 1% de saccharose. Les plaques sont incubées à 23 °C avec une période d'illumination de 16 h le jour et de 8 h la nuit sous agitation rotative de 90 tr/min pendant environ 5 jours. Un tri est ensuite opéré afin de conserver une seule plantule par puit et les plaques sont de nouveau incubées, cette fois-ci sur un plateau statique pendant environ 4 jours. Le milieu MS additionné de 1% de saccharose est ensuite remplacé par 500 µl de milieu MS sans saccharose, qui pourrait perturber la coloration et induire des erreurs de lecture. Après 48 heures d'incubation, le milieu MS est renouvelé (450 µl) et 50 µl d'extrait de violette sont ajoutés. Chaque extrait est analysé en triplicata avec intégration d'un témoin positif, une solution de Bion à 40 µg/ml, d'un blanc H_2O et d'un blanc MeOH à 50% pour valider les résultats.

Après deux jours de mise en contact, le milieu est retiré et 500 µl de solution de coloration GUS, composée de 483,9 µl de Tampon phosphate pH 7, 4,03 µl de solution FerriC K et de solution de FerroC K à 0,1 M et 8,06 µl de substrat X-Gluc à 50 mg/ml dans du NN-diméthylformamide, sont ajoutés dans chaque puit. On applique deux fois 30 minutes de vide sur les plaques non couvertes et on incube à 37 °C pendant minimum 2h, jusqu'à ce que les témoins positifs soient colorés en bleu. Une fois la coloration réalisée, la solution est retirée et les feuilles décolorées en ajoutant dans les puits de l'éthanol à 70% afin de visualiser les spots bleus sur les plantules.

REFERENCES

Aceto, Serena, Paolo Caputo, Salvatore Cozzolino, Luciano Gaudio, and Aldo Moretti. **1999**. Phylogeny and Evolution of Orchis and Allied Genera Based on ITS DNA Variation: Morphological Gaps and Molecular Continuity. *Molecular Phylogenetics and Evolution* 13 (1): 67–76.

Ahmad, H., Khan, S. M., Ghafoor, S., et Ali, N. **2009**. Ethnobotanical Study of Upper Siran. *Journal of Herbs Spices & Medicinal Plants*, 15: 86–97.

Anizon, J.-Y., Lamaire, B., et Surbled, M. **2003**. Extraction assistée par micro-ondes. *Techniques de l'Ingénieur*, 1–10.

Ansari, M., et Emami, S. **2016**. β-Ionone et its analogs as promising anticancer agents. *European Journal of Medicinal Chemistry*, 123: 141–154.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., et al. **1995**. The ITS Region of Nuclear Ribosomal DNA: A Valuable Source of Evidence on Angiosperm Phylogeny. *Annals of the Missouri Botanical Garden*, 82(2): 247.

Ballard, H. E., Sytsma, K. J., et Kowal, R. R. **1999**. Shrinking the violets: phylogenetic relationships of infrageneric groups in *Viola* (Violaceae) based on internal transcribed spacer DNA sequences. *Systematic Botany*, 23(4): 439–458.

Barbeta, B. L., Marshall, A. T., Gillon, A. D., Craik, D. J., et Anderson, M. A. **2008**. Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. *Proceedings of the National Academy of Sciences*, 105(4): 1221–1225.

Barry, D. G., Daly, N. L., Clark, R. J., Sando, L., et Craik, D. J. **2003**. Linearization of a Naturally Occurring Circular Protein Maintains Structure but Eliminates Hemolytic Activity. *Biochemistry*, 42(22): 6688–6695.

Bartle, K. D., et Myers, P. **2002**. History of gas chromatography. *Trends in Anlytical Chrmistry*, 21(02): 547–557.

Bedoukian, P. Z. 1982. Violet Fragrance Compounds. *Fragrance Chemistry: The Science of the Sense of Smell* (Academic Press Inc.): 285–315. London: Ernst T. Theimer.

Beille, L. 1909. Famille des Violacées. *Botanique pharmaceutique*, vol. II: 411–415. A. Maloine.

Beirnaert, C., Meysman, P., Vu, T. N., Hermans, N., Apers, S., et al. **2018**. speaq 2.0: A complete workflow for high-throughput 1D NMR spectra processing et quantification. (D. Schneidman, Ed.)*PLOS Computational Biology*, 14(3): e1006018.

Benaissi, K. 2013. Le CO2 supercritique appliqué à l'extraction végétale. *Techniques de l'Ingénieur*, 1–16.

Benayad, Zakia, Carmen Gómez-Cordovés, and Nour Es-Safi. **2014**. Characterization of Flavonoid Glycosides from Fenugreek (Trigonella Foenum-Graecum) Crude Seeds by HPLC–DAD–ESI/MS Analysis. *International Journal of Molecular Sciences* 15 (11): 20668–85. https://doi.org/10.3390/ijms151120668.

Bhatt, V. P., et Negi, G. C. S. **2006**. Ethnomedicinal plant resources of Jaunsari tribe of Garhwal Himalaya, Uttaranchal. *Indian Journal of Traditional Knowledge*, 5(3): 331–335.

Blomberg, J., Schoenmakers, P. J., Beens, J., et Tijssen, R. **1997**. Compehensive two-dimensional gas chromatography (GCxGC) and its applicability to the characterization of complex (petrochemical) mixtures. *Journal of High Resolution Chromatography*, 20(10): 539–544.

Bonnier, G., et Douin, R. **1990**. *La Grande Flore en couleurs*, vol. 3. Belin.

Burgot, G., et Pellerin, F. **2003**. Microextraction en phase solide (SPME). *Techniques de l'ingénieur*, 33(0): 0–20.

Burketova, L., Trda, L., Ott, P. G., et Valentova, O. **2015**. Bio-based resistance inducers for sustainable plant protection against pathogens. *Biotechnology Advances*, 33(6): 994–1004.

Cai Ke, J. **2007**. Violaceae. *Flora of China* (Science Press (Beijing) and Missouri Botanical Garden (St. Louis)), vol. 13: 72–111. Wu Zhengyi, Peter H. Raven & Hong Deyuan.

Calderón-Santiago, M., Fernández-Peralbo, M. A., Priego-Capote, F., et Luque de Castro, M. D. **2016**. MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes. *Metabolomics*, 12(3): 1–12.

Cao, J., Yin, C., Qin, Y., et Cheng, Z. **2013**. Chemical Constituents of *Viola yedoensis* and Their Antioxidant Activity. *Chinese Journal of Experimental Traditional Medical Formulae*, 19: 77–81.

Caude, M., et Jardy, A. **2016**. *Chromatographie en phase liquide - Théorie et méthodes de séparation Chromatographie en phase liquide Théorie et méthodes de séparation*, 33(0).

Chauchard, B., Munzinger, J., Henry, M., et Cedex, N. **2013**. L'étude caryologique de deux cultivars de Violettes odorantes remet en cause leur origine taxonomique. *Acta Botanica Gallica*, 150: 85–93.

Chen, B., Colgrave, M. L., Daly, N. L., Rosengren, K. J., Gustafson, K. R., et al. **2005**. Isolation and Characterization of Novel Cyclotides from *Viola hederaceae*: Solution structure and anti-HIV activity of vhl-1, a leak-specific expressed cyclotide. *Journal of Biological Chemistry*, 280(23): 22395–22405.

Chervin, J., Perio, P., Martins-Froment, N., Pharkeovilay, C., Reybier, K., et al. **2017**. Dereplication of natural products from complex extracts by regression analysis and molecular networking: case study of redox-active compounds from *Viola alba subsp. dehnhardtii*. *Metabolomics*, 13(8). https://doi.org/10.1007/s11306-017-1227-6.

Costa, R., Fanali, C., Pennazza, G., Tedone, L., Dugo, L., et al. **2015**. Screening of volatile compounds composition of white truffle during storage by GCxGC-(FID/MS) et gas sensor array analyses. *LWT - Food Science et Technology*, 60(2): 905–913.

Cotton, J. 2015. *Application de l'analyse métabolomique à la détection ciblée et globale de contaminants organiques dans des matrices agroalimentaires et environnementales par spectrométrie de masse à ultra-haute résolution*. Université Pierre et Marie Curie Paris 6.

Craik, D. J. 2012. Host-Defense Activities of Cyclotides. *Toxins*, 4: 139–156.

C**u, J. Q.**, Perineau, F., et Gaset, A. **1992**. Volatile components of violet leaves. *Phytochemistry*, 31(2): 571–573.

de Hoffmann, E., & Stroobant, V. **2005.** *Spectrométrie de masse* (3e édition). Dunod.

De Lamarck, J.-B. P. A. de M., et De Candolle, A. P. **1805**. *Flore française, ou description succinctes de toutes les plantes qui croissent naturellement en France*, vol. 4. Paris: Desray.

Dudareva, N., Klempien, A., Muhlemann, J. K., et Kaplan, I. **2013**. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist*, 198: 16–32.

Dunn, W. B., et Ellis, D. I. **2005**. Metabolomics: Current analytical platforms and methodologies. *Trends in Analytical Chemistry*, 24(4): 285–294.

Eriksson, L, T Byrne, E Johansson, J Trygg, et C. Vikström. **2013.** 'Multivariate Characterization'. In *Multi- and Megavariate Data Analysis*, Third revised edition, 131–45. Umea, Sweden: Umetrics Academy.

Eriksson, L., Trygg, J., et Wold, S. **2008**. CV-ANOVA for significance testing of PLS and OPLS® models. *Journal of Chemometrics*, 22(11–12): 594–600.

Favier, A. 2003. *Le stress oxydant*. L'actualité chimique.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4): 783–791.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., et Whitehouse, C. M. **1990**. Electropsray ionization-principles and practice. *Mass Spectrometry Reviews*, 9: 32–70.

Fernandez, X., Breme, K., et Varlet, V. **2009**. Couplage chromatographie en phase gazeuse / olfactométrie. *Technique de l'Ingénieur*, 33(Base documentaire : Gestion des odeurs et des nuisances olfactives): 26.

Fernandez, X., Cabrol-bass, D., Fernandez, X., et Cabrol-bass, D. **2007**. Analyse des arômes. *Technique de l'Ingénieur*, 33(0): 1–22.

Ferrari, G., Donsı, F., et Pataro, G. **2010**. Applications of Pulsed Electric Field Treatments for the Enhancement of Mass Transfer from Vegetable Tissue. *Food Eng Rev (2010)*, 2: 109–130.

Ferreira, S. 2004. Doehlert matrix: a chemometric tool for analytical chemistry - review. *Talanta*, 63(4): 1061–1067.

Ferreres, Federico, Branca M. Silva, Paula B. Andrade, Rosa M. Seabra, and Margarida A. Ferreira. **2003**. Approach to the Study of C-Glycosyl Flavones by Ion Trap HPLC-PAD-ESI/MS/MS: Application to Seeds of Quince (Cydonia Oblonga). *Phytochemical Analysis* 14 (6): 352–59. https://doi.org/10.1002/pca.727.

Finkelstein, E., Rosen, G. M., et Rauckman, E. J. **1980**. Spin trapping of superoxide and hydroxyl radical: practical aspects. *Archives of Biochemistry and Biophysics*, 200(1): 1–16.

Flamini, G. 2007. Flavonoids and other compounds from the aerial parts of *Viola etrusca*. *Chemistry & Biodiversity*, 4(2): 139–44.

Gacche, R. N., et Jadhav, S. G. **2012**. Antioxidant Activities and Cytotoxicity of Selected Coumarin Derivatives: Preliminary Results of a Structure-Activity Relationship Study Using Computational Tools. *Journal of Experimental and Clinical Medicine*, 4(3): 165–169.

Gerlach, S. L., Rathinakumar, R., Chakravarty, G., Göransson, U., Wimley, W. C., et al. **2010**. Anticancer and chemosensitizing abilities of cycloviolacin O2 from *Viola odorata* and psyle cyclotides from *Psychotria leptothyrsa*. *Biopolymers*, 94(5): 617–625.

Gil, Hee-Young, et **Seung-Chul Kim**. **2016**. *Viola Woosanensis*, a Recurrent Spontaneous Hybrid between *V. Ulleungdoensis* et *V. Chaerophylloides* (Violaceae) Endemic to Ulleung Island, Korea. *Journal of Plant Research* 129 (5): 807–22. https://doi.org/10.1007/s10265-016-0830-3.

Goff, S. A. 2006. Plant Volatile Compounds: Sensory Cues for Health and Nutritional Value? *Science*, 311(5762): 815–819.

Górecki, Tadeusz, Frederic Lynen, Roman Szucs, et Pat Sandra. **2006**. 'Universal Response in Liquid Chromatography Using Charged Aerosol Detection'. Analytical Chemistry 78 (9): 3186–92. https://doi.org/10.1021/ac060078j.

Goudable, J., et Favier, A. **1997**. Radicaux libres oxyénés et antioxydants. *Nutur Clin Métabol*, 11: 115–120.

Gran, L., Sandberg, F., et Sletten, K. **2000**. *Oldenlandia affinis* (R&S) DC A plant containing uteroactive peptides used in African traditional medicine. *Journal of Ethnopharmacology*, 70: 197–203.

Grapov, D., Wanichthanarak, K., et Fiehn, O. **2015**. MetaMapR: pathway independent metabolomic network analysis incorporating unknowns. *Bioinformatics*, 31(16): 2757–2760.

Grasse, M.-C., Ellena, J.-C., et Peyron, L. **2009a**. Survol historique de la culture des Violettes et de leur valorisation à Tourettes-sur-Loup. *Un jour, une plante : Les violettes en parfumerie*.

Grasse, M.-C., Ellena, J.-C., et Peyron, L. **2009b**. Utilisation des feuilles de violette à Grasse. *Un jour, une plante : Les violettes en parfumerie*.

Gruber, C. W., Čemažar, M., Anderson, M. A., et Craik, D. J. **2007**. Insecticidal plant cyclotides and related cystine knot toxins. *Toxicon*, 49(4): 561–575.

Gustafson, K. R., McKee, T. C., et Bokesch, H. R. **2004**. Anti-HIV cyclotides. *Current Protein and Peptide Science*, 5: 331–340.

Haleng, J., Pincemail, J., Defraigne, J.-O., Charlier, C., et Chapelle, J.-P. **2007**. Le stress oxydant. *Revue Médicale de Liège*, 62(10): 628–38.

Hammami, I., Kamoun, N., et Rebai, A. **2011**. Biocontrol of Botrytis cinerea with essential oil and methanol extract of *Viola odorata L.* flowers. *Archive of Applied Science Research*, 3(5): 44–51.

Hanem Fathy, K. 2012. Prospect of Botanical Biopesticides in Insect Pest Management. *Pharmacologia*, 3(12): 641–656.

Harbourne, N., Marete, E., Jacquier, J. C., et O'Riordan, D. **2009**. Effect of drying methods on the phenolic constituents of meadowsweet (*Filipendula ulmaria*) and willow (*Salix alba*). *LWT - Food Science et Technology*, 42(9): 1468–1473.

Harman, D. 1992. Free radical theory of aging. *Mutation Research/DNAging*, 275(3–6): 257–266.

Hasegawa, M., Kishino, H., et Yano, T. **1985**. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22(2): 160–174.

Hashempour, H., Koehbach, J., Daly, N. L., Ghassempour, A., et Gruber, C. W. **2013**. Characterizing circular peptides in mixtures: sequence fragment assembly of cyclotides from a violet plant by MALDI-TOF/TOF mass spectrometry. *Amino Acids*, 44(2): 581–595.

Hausbeck, Mary K., and **Kurt H. Lamour**. **2004**. *Phytophthora Capsici* on Vegetable Crops: Research Progress and Management Challenges. *Plant Disease* 88 (12): 1292–1303. https://doi.org/10.1094/PDIS.2004.88.12.1292.

Hekking, W. H. A. (Ed.). **1988**. Violaceae ; P. 1. *Rinorea* and *Rinoreocarpus*. *Flora Neotropica*: 1–207. New York: New York Botanical Garden.

Herrmann, A., Burman, R., Mylne, J. S., Karlsson, G., Gullbo, J., et al. **2008**. The alpine violet, *Viola biflora*, is a rich source of cyclotides with potent cytotoxicity. *Phytochemistry*, 69: 939–952.

I**brahim, Ruqaya Mohammed. 2017**. Immunological, Cytogenetic and Hepatoprotective Effect of Viola Odorata Methanolic Extract on Methotrexate Albino Male Mice. *Journal of Biotechnology Research Center* 11 (2): 47–55.

Ireland, David C., Colgrave, M. L., et Craik, D. J. **2006**. A novel suite of cyclotides from *Viola odorata* : sequence variation and the implications for structure, function and stability. *Biochemical Journal*, 400(1): 1–12.

Ireland, David C., Colgrave, M. L., Nguyencong, P., Daly, N. L., et Craik, D. J. **2006**. Discovery and Characterization of a Linear Cyclotide from *Viola odorata*: Implications for the Processing of Circular Proteins. *Journal of Molecular Biology*, 357: 1522–1535.

James W., B., Mark W., C., Maarten J. M., C., Michael F., F., Walter S., J., et al. **2016**. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, 181(1): 1–20.

Jennings, C., West, J., Waine, C., Craik, D., et Anderson, M. **2001**. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. *Proceedings of the National Academy of Sciences*, 98(19): 10614–10619.

Jeong, Y. H., Oh, Y.-C., Cho, W.-K., Shin, H., Lee, K. Y., et al. **2016**. Anti-inflammatory effects of *Viola yedoensis* and the application of cell extraction methods for investigating bioactive constituents in macrophages. *BMC Complementary and Alternative Medicine*, 16: 1–16.

Jukes, T. H., et Cantor, C. R. **1969**. Evolution of protein molecules. *Mammalian Protein Metabolism* (Academic Press), vol. III: 21–132. Munro H. N.

Jung, M., Geiger, H., Zinsmeistert, H. D., et Botanik, F. **1995**. Tri- and tetrahydroxycoumarin derivatives from *tetraphis pellucida*. *Phytochemistry*, 39(2): 379–381.

Kanchanapoom, T. 2007. Aromatic diglycosides from *Cladogynos orientalis*. *Phytochemistry*, 68(5): 692–696.

Kanti Das, T., Wati, M. R., et Fatima-Shad, K. **2014**. Oxidative Stress Gated by Fenton and Haber Weiss Reactions and Its Association with Alzheimer's disease. *Archives of Neuroscience*, 2(3): 1–8.

Katajamaa, M., Miettinen, J., et Oresic, M. **2006**. MZmine : toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics*, 22(5): 634–636.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16: 111–120.

Le Floch, F., Tena, M. T., Rios, A., et Valcarcel, M. **1998**. Supercritical fluid extraction of phenol compounds from olive leaves. *Talanta*, 46(5): 1123–1130.

Lee, S.-W., Chen, Z.-T., et Chen, C.-M. **1993**. The 4α-Methylsterol Violasterol A from *Viola Formosana Hayata*. *Journal of the Chinese Chemical Society*, 40(3): 305–307.

Leopoldini, M., Marino, T., Russo, N., et Toscano, M. **2004**. Antioxidant Properties of Phenolic Compounds: H-Atom versus Electron Transfer Mechanism. *The Journal of Physical Chemistry A*, 108(22): 4916–4922.

Li, Y., Fabiano-Tixier, A. S., Vian, M. A., et Chemat, F. **2013**. Solvent-free microwave extraction of bioactive compounds provides a tool for green analytical chemistry. *Trends in Analytical Chemistry*, 47: 1–11.

Ligor, Magdalena, Sylwia Studzińska, Aleš Horna, et Bogusław Buszewski. **2013**. Corona-Charged Aerosol Detection: An Analytical Approach. Critical Reviews in *Analytical Chemistry,* 43 (2): 64–78. https://doi.org/10.1080/10408347.2012.746134.

Liochev, S. I., et Fridovich, I. **2002**. The Haber-Weiss cycle—70 years later: an alternative view. *Redox Report*, 7(1): 55–57.

Little, J. R., et McKinney, L. E. **2016**. Violaceae Batsch. *Flora of North America*, vol. 6. Oxford University Press.

Liza, M. S., Abdul Rahman, R., Mandana, B., Jinap, S., Rahmat, A., et al. **2012**. Supercritical fluid extraction of bioactive flavonoid from Strobilanthes crispus (Pecah kaca) and its comparison with solvent extraction. *International Food Research Journal*, 19(2): 503–508.

Luzzatto-Knaan, T., Melnik, A. V., et Dorrestein, P. C. **2015**. Mass spectrometry tools and workflows for revealing microbial chemistry. *Analyst*, 140(15): 4949–4966.

Mahieu, N. G., et Patti, G. J. **2017**. Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. *Analytical Chemistry*, 89(19): 10397–10406.

Malécot, V., Marcussen, T., Munziger, J., Yockteng, R., et Henry, M. **2007**. On the origin of the sweet-smelling Parma violet cultivars (Violaceae): wide intraspecific hybridization, sterility, and sexual reproduction. *American Journal of Botany*, 94(1): 29–41.

Malien-Aubert, C., Dangles, O., et Amiot, M. J. **2001**. Color Stability of Commercial Anthocyanin-Based Extracts in Relation to the Phenolic Composition. Protective Effects by Intra- and Intermolecular Copigmentation. *Journal of Agricultural and Food Chemistry*, 49(1): 170–176.

Mandal, V., Mohan, Y., et Hamalatha, S. **2007**. Microwave Assisted Extraction - An Innovative and Promising Extraction Tool for Medicinal Plant Research. *Pharmacognosy Reviews*, 1(1): 1–18.

Marcussen, T., Karlsson, T., Jonsell, B., et Wind, P. **2010**. Violaceae. *Flora Nordica*, vol. 6: 12–52. Stockholm: The Bergius Foundation.

Marzouk, M. M., Elkhateeb, A., Ibrahim, L. F., Sameh, R., et Kawashty, S. A. **2012**. Two Cytotoxic Coumarin Glycosides from the Aerial Parts of *Diceratella elliptica* (DC.) Jonsell Growing in Egypt. *Rec. Nat. Prod.*, 3: 237–241.

Michener, C. D., et Sokal, R. R. **1957**. A quantitative approach to a problem in classification. *Evolution*, 11(2): 130–162.

Milman, B. L., et Zhurkovich, I. K. **2016**. Mass spectral libraries: A statistical review of the visible use. *TrAC Trends in Analytical Chemistry*, 80: 636–640.

Molnár, P., et Szabolcs, J. **1980**. Occurrence of 15-cis-violaxanthin in *Viola tricolor*. *Phytochemistry*, 19(4): 623–627.

Moon, H.-I., Jung, J.-C., et Lee, J. **2007**. Antiplasmodial activity of triterpenoid isolated from whole plants of *Viola* genus from South Korea. *Parasitology Research*, 100(3): 641–644.

Muhammad, N., Saeed, M., Awen, A., et Khan, H. **2012**. Ethnomedicinal, phytochemical and pharmacological profile of genus *Viola*. *Phytopharmacology*, 3(1): 214–226.

Murray, A. P., Nacional, U., Rodriguez, S., et Nacional, U. **2015**. Chemical Constituents and Biological Activities of Plants from the Genus *Viola*. *Chemistry & Biodiversity*, 12: 1777–1808.

Nakamura, Koh, Tetsuo Denda, Goro Kokubugata, Chiun-Jr Huang, Ching-I Peng, et Masatsugu Yokota. **2015**. 'Phylogeny et Biogeography of the *Viola Iwagawae-Tashiroi* Species Complex (Violaceae, Section Plagiostigma) Endemic to the Ryukyu Archipelago, Japan'. *Plant Systematics and Evolution* 301 (1): 337–51. https://doi.org/10.1007/s00606-014-1076-y.

Naruya, S., et Masatoshi, N. **1987**. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology et Evolution*, 4(4): 406–425.

Nejia, H. 2013. *Extraction et purification de substances naturelles : comparaison de l'extraction au CO2-supercritique et des techniques conventionnelles*. Université de Toulouse.

Nepveu, F. 2008. APPRT- Région Midi-Pyrénées VALorisation d'Agro-Produits : Identification de Marqueurs spécIfiques par leurs Propriétés redox Porteur *: UPS*: 1–39. Document interne, Toulouse: Université Paul Sabatier.

Nguyen, H. T., Vo, B. T., Nguyen, L. H., Bernad, J., Alaeddine, M., et al. **2013**. Extracts of *Crinum latifolium* inhibit the cell viability of mouse lymph oma cell line EL4 et induce activation of anti-tumour activity of macrophages in vitro. *Journal of Ethnopharmacology*, 149(1): 75–83.

Niederbacher, B., Winkler, J. B., et Schnitzler, J. P. **2015**. Volatile organic compounds as non-invasive markers for plant phenotyping. *Journal of Experimental Botany*, 66(18): 5403–5416.

Nkhili, E., Tomao, V., El Hajji, H., El Boustani, E.-S., Chemat, F., et al. **2009**. Microwave-assisted water extraction of green tea polyphenols. *Phytochemical Analysis*, 20(5): 408–415.

Ohashi, Y., et Ohshima, M. **1992**. Stress-induced expression of genes for pathogenesis-related proteins in plants. *Plant and Cell Physiology*, 33(7): 819–826.

Özel, M. Z., Göğüş, F., et Lewis, A. C. **2006**. Comparison of direct thermal desorption with water distillation and superheated water extraction for the analysis of volatile components of *Rosa damascena Mill*. using GCxGC-TOF/MS. *Analytica Chimica Acta*, 566(2): 172–177.

Papet, Y., Brunet, B., et Mura, P. **2010**. Headspace (HS) et micro-extraction en phase solide (SPME). Théorie et applications. *Annales de Toxicologie Analytique*, 22(2): 75–79.

Parsley, N. C., Kirkpatrick, C. L., Crittenden, C. M., Rad, J. G., Hoskin, D. W., et al. **2018**. PepSAVI-MS reveals anticancer et antifungal cycloviolacins in *Viola odorata*. *Phytochemistry*, 152: 61–70.

Pichersky, E. 2006. Biosynthesis of Plant Volatiles: Nature's Diversity and Ingenuity. *Science*, 311(5762): 808–811.

Pranting, M., Loov, C., Burman, R., Goransson, U., et Andersson, D. I. **2010**. The cyclotide cycloviolacin O2 from *Viola odorata* has potent bactericidal activity against Gram-negative bacteria. *Journal of Antimicrobial Chemotherapy*, 65(9): 1964–1971.

Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., et al. **1987**. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*, 238(4825): 336–341.

Procházkoá, D., Boušová, I., et Wilhelmová, N. **2011**. Antioxidant and prooxidant properties of flavonoids. *Fitoterapia*, 82(4): 513–523.

Prost, Isabelle, Sandrine Dhondt, Grit Rothe, Jorge Vicente, Maria José Rodriguez, Neil Kift, Francis Carbonne, et al. **2005**. Evaluation of the Antimicrobial Activities of Plant Oxylipins Supports Their

Involvement in Defense against Pathogens. *Plant Physiology* 139 (4): 1902–13. https://doi.org/10.1104/pp.105.066274.

Qin, Y., Yin, C., et Cheng, Z. **2013**. A New Tetrahydrofuran Lignan Diglycoside from *Viola tianshanica* Maxim. *Molecules*, 18(12): 13636–13644.

Quideau, S., Deffieux, D., Douat-Casassus, C., et Pouységu, L. **2011**. Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. *Angewandte Chemie International Edition*, 50(3): 586–621.

Restek. 2000. *A Technical Guide for Static Headspace Analysis Using GC*.

Ribeiro, S. M., Porto, W. F., Silva, O. N., Santos, M. de O., Dias, S. C., et al. **2013**. Plant Antifungal Peptides. *Handbook of Biologically Active Peptides*: 169–179. Elsevier.

Rodrigues, V. H., Cardoso, C., Paixão, J. A., et Costa, M. M. R. R. **2007**. 2,2,6,6-Tetramethyl-4-oxopiperidinium nitrate. *Acta Crystallographica Section C Crystal Structure Communications*, 63(4): 243–245.

Rouadjia, I. 2009. *Infos Matériaux : Le "champ des odeurs."* Matériauthèque Ensad.

Ruzicka, L., Schinz, H., et Susz, B.-P. **1944**. Veilchenriechstoffe.(14. Mitteilung). Zur Stereoisomerie von Hexen-(3)-ol-(1), Nonadien-(2, 6)-al-(1) und Nonadien-(2, 6)-ol-(1). *Helvetica Chimica Acta*, 27(1): 1561–1569.

Saint-lary, L., Debrauwer, M. L., Gougeon, M. R., Paris, M. J., Thomas, M. O., et al. **2015**. *Evaluation de l'approche métabolomique pour l'authentification des extraits naturels utilisés dans le secteur arômes et parfums*. Université Nice Sophia-Antipolis.

Saldaña, G., Álvarez, I., Condón, S., et Raso, J. **2014**. Microbiological Aspects Related to the Feasibility of PEF Technology for Food Pasteurization. *Critical Reviews in Food Science and Nutrition*, 54(11): 1415–1426.

Salem, M. A., Jüppner, J., Bajdzienko, K., et Giavalisco, P. **2016**. Protocol: a fast, comprehensive and reproducible one-step extraction method for the rapid preparation of polar and semi-polar metabolites, lipids, proteins, starch et cell wall polymers from a single sample. *Plant Methods*, 12: 1–15.

Sanger, F., et Coulson, A. R. **1975**. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94: 441–448.

Sangster, T., Major, H., Plumb, R., Wilson, A. J., et Wilson, I. D. **2006**. A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis. *The Analyst*, 131(10): 1075.

Segovia, F. J., Luengo, E., Corral-Pérez, J. J., Raso, J., et Almajano, M. P. **2015**. Improvements in the aqueous extraction of polyphenols from borage (*Borago officinalis L*.) leaves by pulsed electric fields: Pulsed electric fields (PEF) applications. *Industrial Crops and Products*, 65: 390–396.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., et al. **2003**. Cytoscape : A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, 13: 2498–2504.

Shapiro, A. D., et C. Zhang. **2001**. The Role of NDR1 in Avirulence Gene-Directed Signaling and Control of Programmed Cell Death in *Arabidopsis*. *Plant physiology* 127 (3): 1089–1101. https://doi.org/10.1104/pp.010096.

Shen, X.-Y., et Xie, C. **2009**. Study on Antimicrobial Activity of *Viola tianshanica* Maxim.Extracts. *Biotechnology*, 19: 71–73.

Sies, H. 1985. Oxidative Stress : Introductory Remarks. *Oxidative stress*.

Singleton, Vernon, Rudolf Orthofer, et Rosa M Lamuela-raventós. **1999**. 'Analysis of Total Phenols et Other Oxidation Substrates et Antioxidants by Means of Folin-Ciocalteu Reagent'. In *Methods in enzymology*, Academic Press, 299:152–78.

Skowyra, M., Calvo, M. I., Gallego, M. G., Azman, N. A. M., et Almajano, M. P. **2014**. Characterization of Phytochemicals in Petals of Different Colours From *Viola × wittrockiana Gams*. and Their Correlation With Antioxidant Activity. *Journal of Agricultural Science*, 6(9): 93–105.

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., et Siuzdak, G. **2006**. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Analytical Chemistry*, 78(3): 779–787.

Sumner, Lloyd W., Alexander Amberg, Dave Barrett, Michael H. Beale, Richard Beger, Clare A. Daykin, Teresa W.-M. Fan, et al. **2007**. Proposed Minimum Reporting Standards for Chemical Analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 3 (3): 211–21. https://doi.org/10.1007/s11306-007-0082-2.

Sun, Y., Du, L., Zhou, L., Miao, F., Yang, X., et al. **2011**. Study on antibacterial active components from *Viola yedoensis*. *Zhongguo Zhong Yao Za Zhi.*, 36(19): 2666–2671.

Tholl, D., Boland, W., Hansel, A., Loreto, F., Röse, U. S. R., et al. **2006**. Practical approaches to plant volatile analysis. *The Plant Journal*, 45(4): 540–560.

Trabi, M. 2004. Tissue-Specific Expression of Head-to-Tail Cyclized Miniproteins in Violaceae and Structure Determination of the Root Cyclotide *Viola hederacea* root cyclotide1. *The Plant Cell*, 16(8): 2204–2216.

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., et al. **2015**. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature Methods*, 12: 523–526.

Tsugawa, H., Kind, T., Nakabayashi, R., Yukihira, D., Tanaka, W., et al. **2016**. Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. *Analytical Chemistry*. https://doi.org/10.1021/acs.analchem.6b00770.

Tutin, T. G., Moore, D. M., Valentine, D. H., Walters, S. M., & Webb, D. A. n.d. Violaceae. *Flora Europaea*, vol. 2: 270–275. Cambridge University Press.

Uhde, G., et Ohloff, G. **1972**. Parmon, eine phantomverbindung im veilchenblütenöl. *Helvetica Chimica Acta*, 55(7): 2621–2625.

van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., et van der Werf, M. J. **2006**. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(1): 142.

Vercammen, J., Sandra*, P., Baltussen, E., Sandra, T., et David, F. **2000**. Considerations on Static and Dynamic Sorptive et Adsorptive Sampling to Monitor Volatiles Emitted by Living Plants. *Journal of High Resolution Chromatography*, 23(9): 547–553.

Vergnes, S., Ladouce, N., Fournier, S., Ferhout, H., Attia, F., et al. **2014**. Foliar treatments with Gaultheria procumbens essential oil induce defense responses and resistance against a fungal pathogen in *Arabidopsis*. *Frontiers in Plant Science*, 5(477): 1–8.

Villaverde, J. J., Sandín-España, P., Sevilla-Morán, B., López-Goti, C., et Alonso-Prados, J. L. **2016**. Biopesticides from Natural Products: Current Development, Legislative Framework, and Future Trends. *BioResources*, 11(2). https://doi.org/10.15376/biores.11.2.Villaverde.

Vukics, V., Ringer, T., Kery, A., Bonn, G. K., et Guttman, A. **2008**. *Analysis of heartsease (Viola tricolor L .) flavonoid glycosides by micro-liquid chromatography coupled to multistage mass spectrometry*, 1206: 11–20.

Wang, C. K., Colgrave, M. L., Gustafson, K. R., Ireland, D. C., Goransson, U., et al. **2007**. Anti-HIV cyclotides from the Chinese medicinal herb *Viola yedoensis*. *Journal of Natural Products*, 71: 47–52.

Wardencki, W., Michulec, M., et Curylo, J. **2004**. A review of theoretical and practical aspects of solid-phase microextraction in food analysis. *International Journal of Food Science and Technology*, 39(7): 703–717.

Wätjen, W., Michels, G., Steffan, B., Niering, P., Chovolou, Y., et al. **2005**. Low Concentrations of Flavonoids Are Protective in Rat H4IIE Cells Whereas High Concentrations Cause DNA Damage and Apoptosis. *The Journal of Nutrition*, 135(3): 525–531.

Watson, J. D., Gilman, M., Witkowski, J., et Zoller, M. **1994a**. La réaction de polymérisation en chaîne. *ADN Recombinant*, 33(0): 79–98.

Watson, J. D., Gilman, M., Witkowski, J., et Zoller, M. **1994b**. La réaction de polymérisation en chaîne. *ADN Recombinant*, 33(0): 79–98.

White, T. J., Bruns, T., Lee, S., et Taylor, J. **1990**. Amplification and Direct Sequencing of Fungal Ribosomal Rna Genes for Phylogenetics. *PCR Protocols*, (January 2016): 315–322.

Willmer, P. 2011. The Timing et Patterning of Flowering. *Pollination and Floral Ecology*: 483–502. Princeton: Princeton University Press.

Wodehouse, R. P. 1935. *Pollen grains*. Mcgraw-Hill Book Company, Inc; New York; London.

Wolfender, J.-L., Marti, G., Thomas, A., et Bertrand, S. **2015**. Current approaches and challenges for the metabolite profiling of complex natural extracts. *Journal of Chromatography A*, 1382: 136–164.

Wolfender, J.-L., Queiroz, E. F., et Hostettmann, K. **2006**. The importance of hyphenated techniques in the discovery of new lead compounds from nature. *Expert Opinion on Drug Discovery*, 1(3): 237–260.

Xie, C., Veitch, N. C., Houghton, P. J., et Simmonds, M. S. J. **2003**. Flavone C-glycosides from *Viola yedoensis* MAKINO. *Chemical & Pharmaceutical Bulletin*, 51(10): 1204–7.

Yang, J., Guo, J., et Yuan, J. **2008**. In vitro antioxidant properties of rutin. *LWT - Food Science and Technology*, 41(6): 1060–1066.

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., et al. **2013**. Molecular Networking as a Dereplication Strategy. *Journal of Natural Products*, 76: 1686–1699.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution*, 39(3): 306–314.

Yeshak, M. Y., Burman, R., Eriksson, C., et Göransson, U. **2012**. Optimization of cyclotide extraction parameters. *Phytochemistry Letters*, 5: 776–781.

Zhang, Bo, Yan Huang, Qiaoli Zhang, Xiaofen Liu, Fang Li, et Kunsong Chen. **2014**. Fragrance Discrimination of Chinese *Cymbidium* Species and Cultivars Using an Electronic Nose. *Scientia Horticulturae* 172: 271–77. https://doi.org/10.1016/j.scienta.2014.04.019.

Zhou, H. Y., Hong, J. L., Shu, P., Ni, Y. J., et Qin, M. J. **2009**. A new dicoumarin and anticoagulant activity from *Viola yedoensis* Makino. *Fitoterapia*, 80(5): 283–285.

Zhu, Hong, Shuang-Shuang Qin, Na Zhang, Da-Wei Yang, Hua-Rui Han, Kun-Hua Wei, et Min-Hui Li. **2015**. 'Chemical Constituents et Biological Activities of Plants from the Genus *Viola*'. *Chemistry & Biodiversity* 12 (12): 1777–1808.

ANNEXES

ANNEXE 1. BASES THEORIQUES DES METHODES UTILISEES

A. BASES THEORIQUES DE LA PHYLOGENIE MOLECULAIRE

La phylogénie moléculaire est une science qui permet d'étudier le lien entre espèces apparentées par l'intermédiaire des séquences ADN. L'analyse phylogénétique est la plus couramment utilisée de nos jours et permet de compléter les méthodes traditionnelles reposant sur les caractéristiques morphologiques et anatomiques.

A.1 La réaction de polymérisation en chaine (PCR)

C'est une technique inventée en 1983 par K. Mullis très utilisée en phylogénie puisqu'elle permet de multiplier une région spécifique de l'ADN, assurant ainsi une meilleure lecture du code génétique. Cette réplication sélective *in vitro* est appelée amplification et repose sur l'utilisation d'une ADN polymérase. Cette réaction se déroule dans un milieu réactionnel constitué de l'extrait d'ADN, d'une Taq polymérase, des amorces et des quatre désoxyribonucléosides triphosphate constitutifs de l'ADN apportés en excès. Elle comprend trois étapes effectuées à des températures optimales grâce à l'utilisation d'un thermocycleur (Figure S1).

La première étape est la dénaturation du brin d'ADN qui survient à une température de 94°C. A cette température, les liaisons hydrogènes sont rompues et l'ADN double-brin est alors transformé en ADN simple-brin. Afin de pouvoir amplifier sélectivement une région de l'ADN, il est indispensable de disposer d'amorces (encore appelées Primers) qui sont de petits fragments d'ADN complémentaires aux extrémités de la séquence à répliquer et généralement constitués de 10 à 30 nucléotides. Elles vont permettre de localiser et d'isoler le fragment d'ADN d'intérêt. L'hybridation des amorces, qui repose donc sur le principe d'appariement des bases, est obtenue pour des températures comprises entre 40 et 70 °C. L'élongation est réalisée à une température de 72 °C grâce à l'intervention de la Taq polymérase qui catalyse la réplication en se liant aux amorces et en utilisant les divers désoxyribonucléosides triphosphates du milieu réactionnel. A chaque nouveau cycle, les brins nouvellement synthétisés deviennent des matrices à répliquer ; la réplication devient alors exponentielle. En moyenne de 20 à 40 cycles sont nécessaires pour obtenir une quantité détectable et analysable par électrophorèse sur gel d'agarose par exemple. La révélation est obtenue par coloration au bromure d'éthidium et est visualisée aux ultraviolets (280-320 nm).

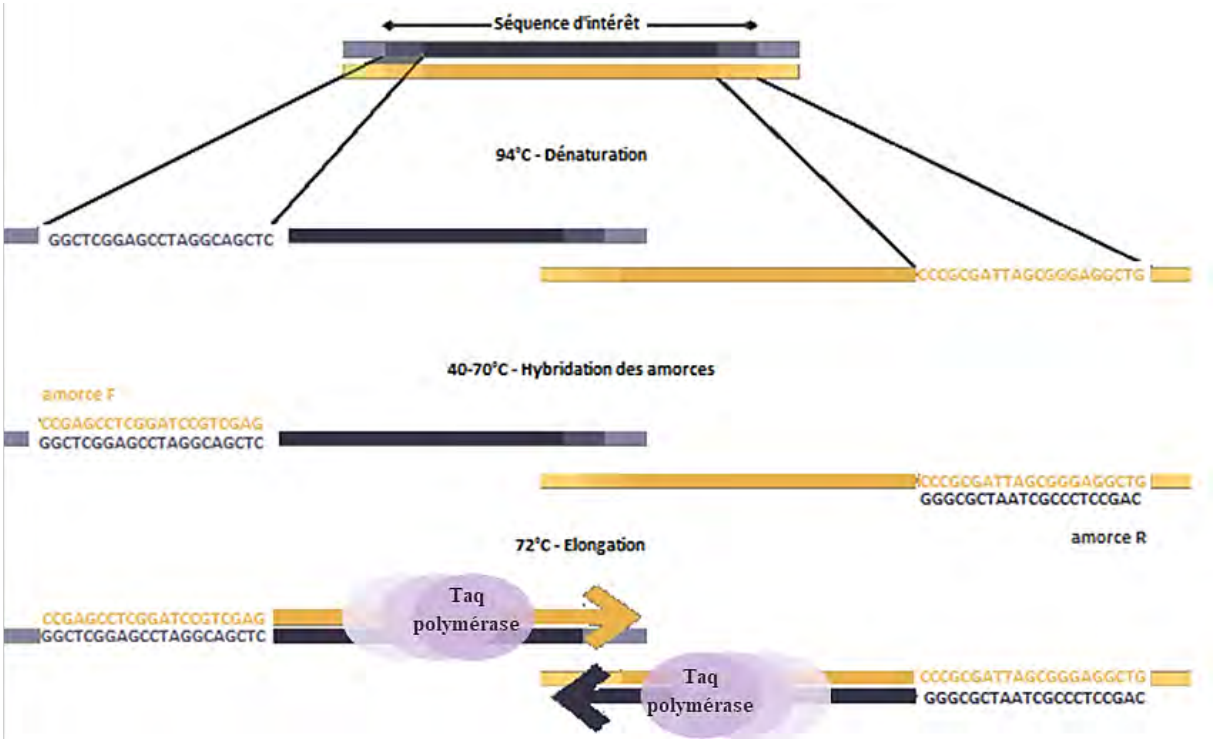


Figure S1. Schéma des étapes de l'amplification d'une séquence par PCR (Watson et al., 1994)

A.2 Arbres phylogénétiques

L'analyse phylogénétique permet d'aborder différentes questions telles que l'histoire évolutive des espèces, leur identification et leur classification ainsi que l'évolution des caractères par exemple. Avant ces derniers provenaient des comparaisons morphologiques, comportementales ou encore des répartitions géographiques. Aujourd'hui les phylogénies sont principalement obtenues à partir des séquences moléculaires (ADN, ARN, protéines). Elle conduit à la construction d'arbres phylogénétiques reposant sur différentes méthodes et son point de départ est l'alignement des séquences ADN (Figure S2).

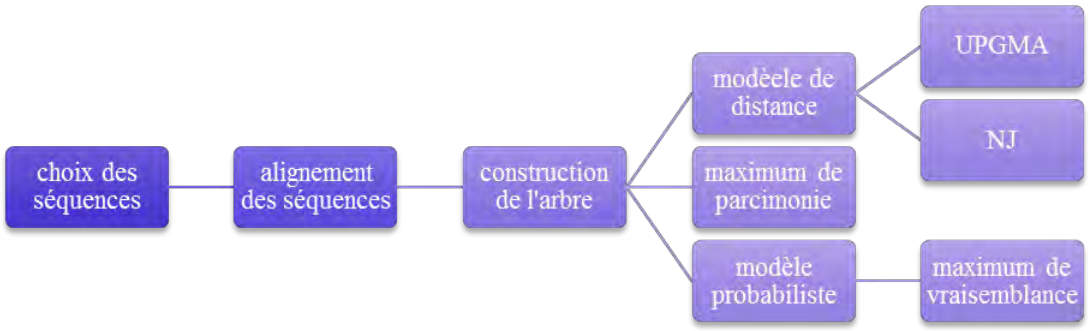


Figure S2. Cheminement de la construction d'un arbre phylogénétique

A.2.1 Alignement des séquences

Les approches phylogénétiques débutent toutes par un séquençage et un alignement des séquences d'ADN ou de protéines homologues. L'alignement est une étape essentielle et doit être de bonne qualité. C'est pourquoi les zones non fiables doivent être éliminées du reste de l'analyse. Les sites sont superposés de manière à conserver l'alignement le plus parcimonieux, c'est-à-dire à maximiser la similarité des séquences et donc à minimiser le nombre d'évènements mutationnels. Plusieurs

mutations peuvent être observées : des substitutions, des insertions et des délétions rendant cette étape relativement complexe. La majorité des méthodes ne tient compte que des substitutions. L'alignement définit les correspondances entre nucléotides et constitue donc une donnée fondamentale dans la reconstruction phylogénétique.

A.2.2 Méthodes de construction des arbres phylogénétiques

Il existe différentes méthodes de construction d'arbre qui sont sélectionnées selon les données et les besoins de l'étude. On distingue des méthodes de distance pour étudier principalement la similarité des séquences, et des méthodes basées sur les caractères et assurant l'obtention d'arbres à caractère évolutif.

A.2.2.1 Méthode de distance

Les méthodes de distance telles que UPGMA (unweighted pair group method with arithmetic averages) (**Michener et Sokal, 1957**) et NJ (Neighbor-joining) conçue par Saitou et Nei (**Naruya et Masatoshi, 1987**) permettent la reconstruction d'arbre phylogénétique par la recherche de séquences les plus proches. Elles reposent sur la construction de matrices de distance entre les séquences, sont rapides et donnent de bons résultats pour des séquences présentant une forte similarité. Une estimation de distance peut être calculée en alignant les séquences et en comptant ensuite le nombre de positions où les séquences diffèrent : rapport du nombre de nucléotides différents et du nombre de sites examinés. Le nombre de différences, appelé nombre observé de substitutions, est généralement une sous-estimation de la distance réelle car de multiples mutations pourraient avoir eu lieu à n'importe quelle position. Pour corriger ces substitutions cachées, un modèle de substitution, tel que Jukes-Cantor ou Kimura 80, peut être utilisé pour obtenir une estimation de distance plus précise. Initialement, chaque organisme d'entrée est placé dans son propre groupe qui correspond à un nœud dans l'arbre résultant (arbre en étoile), puis des groupes sont itérativement formés.

A.2.2.2 Méthodes fondées sur les caractères

Ces méthodes sont plus longues mais plus précises et reposent sur un ou plusieurs caractères à étudier.

a) Maximum de vraisemblance

Cette méthode de reconstruction phylogénétique évalue, en termes de probabilités, l'ordre des branchements et la longueur des branches d'un arbre sous un modèle évolutif donné. C'est la méthode la mieux justifiée sur un plan théorique. Elle cherche à identifier l'arbre le plus probable compte tenu des données. Autrement dit, cette méthode cherche l'arbre conduisant à la plus forte probabilité de produire les séquences observées. Ainsi, tous les arbres possibles sont modélisés en calculant une estimation de maximum de vraisemblance. Cependant ce calcul étant long et le nombre de possibilités exponentiellement proportionnel au nombre de taxons, des méthodes heuristiques sont employées afin de converger plus rapidement vers l'arbre le plus proche de l'arbre réel. Ces méthodes nécessitent la mise au point d'un arbre phylogénétique initial comme point de départ. Il est généralement obtenu à partir d'une méthode de distance.

b) Parcimonie

Cette méthode s'appuie sur plusieurs hypothèses. Les sites évoluent indépendamment les uns des autres (la séquence peut être considérée comme une suite de caractères non ordonnés) et la vitesse d'évolution est lente et constante au cours du temps. Elle conduit à des arbres non enracinés.

Le temps de calcul est plus réduit mais conduit à des arbres moins précis. Elle consiste à choisir l'arbre qui minimise le nombre d'évènements évolutifs le long de ses branches, c'est-à-dire qui minimise le nombre de changements d'états de caractères, ici les nucléotides.

A.2.2.3 Modèles de substitutions et calcul de distance

Pour estimer une distance évolutive entre les taxons, il faut avoir recours à un modèle de la fréquence à laquelle les mutations se produisent au niveau de l'ADN. Ces modèles sont appelés modèles de substitutions au nombre de cinq : Jukes-Cantor (**Jukes et Cantor, 1969**), Felsenstein 81 **(Felsenstein, 1985**), Kimura 80 (**Kimura, 1980**), HKY (**Hasegawa et al., 1985**) et GTR (**Z. Yang, 1994**). Leurs points communs sont qu'ils partent du postulat que les mutations sur différents sites du génome sont indépendantes et que les mutations sur chaque site suivent la même distribution probabiliste. Les modèles Jukes-Cantor et Kimura 80 partent du principe que les quatre bases présentent des fréquences égales tandis que les modèles HKY et GTR assurent une différenciation estimées par les fréquences des bases azotées observés dans l'alignement des séquences. Par ailleurs, pour le modèle Jukes-Cantor, toutes les substitutions sont supposées à poids égal tandis que pour Kimura 80 et HKY, les taux de transition (substitution entre A et G ou entre C et T) et transversion (substitutions entre A/G et C/T) peuvent être différents. Quant au modèle GTR il préconise que toutes les substitutions peuvent se produire à différentes fréquences.

A.2.3 Enracinement

La majorité des méthodes conduit à des arbres non enracinés puisqu'elle s'appuie sur des différences de séquences sans orientation temporelle. L'enracinement d'un arbre, qui permet d'intégrer une dimension temporelle et ainsi de déterminer les relations de parentés, peut s'effectuer de deux manières : l'enracinement au poids moyen pour lequel le postulat que les lignées évoluent toutes à la même vitesse dès leur divergence est appliqué, impliquant que la racine sera située à équidistance des nœuds ; et l'enracinement par un groupe extérieur de séquence homologue aux séquences analysées. Par exemple, dans le cas des violettes, on pourra intégrer des séquences de pensées dont on sait qu'elles seront proches car de la même famille, mais différentes car d'une autre espèce. La racine est alors sur la branche reliant le groupe externe aux autres séquences.

A.2.4 Analyse bootstrap

L'analyse bootstrap est une approche probabiliste et l'un des moyens les plus couramment usités pour évaluer la fiabilité de la topologie d'un arbre phylogénétique. Elle permet d'étudier l'incidence de la pondération par site (équivalent d'un nucléotide), c'est-à-dire que les branches sont évaluées en utilisant la technique de rééchantillonnage. Plusieurs arbres sont créés en piochant au hasard des sites puis en le remettant modifiant ainsi les alignements originels. Pour chaque branche de l'arbre d'origine, son occurrence dans les autres arbres est observée et un score de 1 au nœud à la racine de la branche est indexé si elle est retrouvée. La procédure est alors répétée plusieurs fois, généralement 100 à 1000 fois. La valeur bootstrap est alors calculée en tant que pourcentage d'arbres réévalués contenant la même branche. Ainsi, plus la valeur bootstrap est élevée et plus la branche de l'arbre associée est confirmée et robuste, puisque sa variabilité est faible. Cependant, un biais est inhérent à cette méthode. En effet, elle n'est pas appropriée aux séquences proches car elles possèdent peu de sites pour les différencier et donc les valeurs de bootstrap seront faibles.

B. PRINCIPE DE LA METABOLOMIQUE

La métabolomique (**Cotton, 2015**) est une technique de plus en plus utilisée et touche de nombreux domaines : la santé, la nutrition et l'analyse d'extraits végétaux. Elle fait partie des méthodes chimiométriques qui mettent en relation des mesures faites sur un système, par exemple un échantillon, avec son état, défini par des grandeurs qualitatives ou quantitatives, par des méthodes mathématiques et/ou statistiques. C'est une technique qui permet l'étude des métabolites, molécules de faibles masses moléculaires, synthétisés à partir de protéines et qui participent au bon fonctionnement de la cellule, notamment au niveau de la croissance et de la défense. Le terme métabolomique définit l'analyse complète et quantitative de tous les métabolites d'un système. Si l'on s'intéresse à une partie des métabolites, par exemple une classe chimique particulière, on parlera de profilage des métabolites et enfin on utilisera le terme empreinte métabolique lors de l'analyse d'un grand nombre d'échantillons dans un but de classification (**Dunn et Ellis, 2005**). Deux approches sont distinguées en métabolomique :

a) approche ciblée

Les méthodes ciblées se focalisent sur un nombre limité de métabolites. Ainsi, l'objectif est d'établir une relation entre l'état et les mesures qualitative et quantitative de certains métabolites d'intérêt ou de familles particulières de métabolites. C'est donc une approche a priori dont la principale limite réside dans le fait de devoir connaître au préalable les composés cibles. C'est pourquoi une approche non ciblée peut également être envisagée.

b) approche non ciblée

Les méthodes non ciblées se basent sur l'ensemble des signaux détectables par une méthode physico-chimique. Elles assurent une détection d'un grand nombre de composés, sans recherche de caractérisation ou de quantification. Dans un second temps, les signaux d'intérêtssont identifiés à l'aide de méthodes d'analyses adaptées. L'analyse complète des métabolites d'un système, comme un extrait brut de plante par exemple, est une entreprise très ambitieuse compte tenu du nombre de composés contenu dans cet extrait. Néanmoins, la combinaison de méthodes d'extraction, d'analyse et de traitement des données permet de donner un bon aperçu du métabolome d'un organisme. L'analyse métabolomique présente un intérêt dans l'étude d'extraits végétaux puisqu'elle permet d'identifier et de quantifier des métabolites qui peuvent être mis en relation avec des activités biologiques ciblées, mais également répondre à des questions pertinentes telles que l'origine géographique, l'espèce botanique ou encore l'influence de l'environnement.

B.1 Traitement de l'échantillon

La préparation de l'échantillon doit rester simple de manière à limiter les biais entre les échantillons et elle doit permettre de conserver au maximum l'hétérogénéité de l'échantillon de manière à obtenir une bonne représentativité du métabolome étudié. Lors de la récolte, il est primordial de veiller à éviter la formation d'artefacts due à des réactions enzymatiques ou d'oxydation par exemple. C'est pourquoi, les échantillons prélevés sont directement plongés dans un bain d'azote liquide puis conservés à très basse température (- 80°C). De même, lors du séchage, il est possible de relancer ces réactions non désirées (**Harbourne et al., 2009**), c'est pourquoi la lyophilisation est préférée puisqu'elle s'opère à partir du matériel gelé. Cette étape permet *in fine* une meilleure conservation de la matière en raison de l'élimination de l'eau, et donc de la limitation des réactions enzymatiques.

La majorité des techniques analytiques utilisées dans une étude métabolomique est de type chromatographique (gazeuse ou liquide) et spectroscopique (RMN, MS, …). La spectrométrie de masse est essentiellement couplée avec une colonne de chromatographique gazeuse (GC-MS) dans le cas de l'analyse de composés (semi-)volatils, ou liquide (LC-MS) dans le cas de l'analyse de composés peu ou non volatils. La technologie la plus utilisée en analyse métabolomique reste la LC-MS car elle permet de séparer et de détecter une large gamme de composés. La spectrométrie de masse haute résolution (HRMS) est de plus en plus utilisée dans les analyses métabolomiques en raison de sa sensibilité élevée et de l'obtention d'une masse exacte, indispensable au processus d'identification. La source d'ionisation prépondérante reste l'ionisation par électronébulisation (ESI) puisqu'elle assure l'ionisation à la fois des composés polaires et peu polaires. Le principal inconvénient de cette technique réside dans la production importante d'adduits. Les principaux analyseurs associés sont la FT-ICR (Fourier Transform Ion Cyclotron Resonance), l'Orbitrap et le ToF (Time of Flight).

B.2 Traitement des données

Différents logiciels sont ensuite utilisés de manière à effectuer une détection automatique des informations spectrales et chromatographiques des échantillons analysés en UHPLC-MS et GC-MS, à l'aide de divers algorithmes afin de sélectionner les informations pertinentes. Ces derniers peuvent être libres ou commerciaux et les plus utilisés sont MZmine **(Katajamaa et al., 2006**), XCMS (**Smith et al., 2006**) et la combinaison MS-DIAL (**Tsugawa et al., 2015**) et MS-FINDER (Tsugawa et al., 2016b). Des conversions des fichiers .raw sont souvent nécessaires avant implémentation dans les logiciels comme par exemple la conversion en .CDF avec le convertisseur DataBridge pour MZmine ou en .abf avec le convertisseur ABFconverter pour MS-DIAL. Le but final est d'obtenir par la suite une liste de pics qui servira aux analyses statistiques et à l'identification des composés à l'aide de bases de données. Deux approches de traitement statistique peuvent être mises en œuvre :

a) une analyse non supervisée

Les méthodes non supervisées s'appuient uniquement sur les données spectrales et permettent de comprendre leur structuration. Elle assure une observation de la variabilité générale pour une étude globale des données brutes. La méthode d'analyse statistique non supervisée la plus connue est l'analyse en composantes principales (ACP). C'est la première étape de la démarche d'analyses multivariées avec aucun à priori de l'appartenance d'un échantillon à un groupe donné. Elle permet d'extraire et de synthétiser les informations essentielles des données en nouvelles variables appelées composantes principales, qui correspondent à une combinaison linéaire des variables originelles. L'ACP va alors permettre d'identifier les directions, c'est-à-dire les axes principaux pour maximiser la variation des données. Ainsi, elle conduit à la réduction des dimensions, généralement en 2 ou 3 composantes, permettant de visualiser graphiquement l'information en en perdant le moins possible. Ces composantes sont orthogonales les unes par rapport aux autres et permettent la formation d'un nouveau système de coordonnées pour la projection des échantillons. Le plan ainsi formé est appelé carte factorielle des individus (score plot en anglais). Chaque composante explique un certain pourcentage du jeu de données total. La présence de groupes représentés par la proximité des échantillons révèle leur similarité de comportement par rapport aux variables étudiées.

Des techniques de classification peuvent également être utilisées pour mettre en évidence des groupes supposant l'existence de ressemblance entre les individus. Elles sont complémentaires de l'ACP. La principale technique est la classification ascendante hiérarchique (CAH). C'est un processus itératif qui permet la formation de groupes matérialisés par la suite sous forme de dendrogramme. Au départ, il y a autant de groupes que de variables, puis des fusions entre les groupes les plus similaires sont ensuite

effectuées jusqu'à obtenir des groupes clairement distincts. En amont, un prétraitement des variables est généralement effectué, dont une normalisation afin de réduire la variance entre les variables et ne pas favoriser l'expression de certaines variables par rapport à d'autres. En effet, d'un point de vue biologique, les métabolites abondants ne sont pas nécessairement plus intéressants que les métabolites peu concentrés. Ces prétraitements permettent ainsi de réduire l'influence du bruit de fond et de se focaliser sur les informations biologiques pertinentes (**van den Berg et al., 2006**).

b) une analyse supervisée

Les méthodes supervisées utilisent les données spectrales et une ou des informations quantitatives ou qualitatives basées sur l'état du système. Elles assurent une mise en relation des variables après retraitement et normalisation dans une optique de recherche de modèles prédictifs. Parmi les méthodes d'analyse supervisée est retrouvée la régression des moindres carrés partiels (Partial Least Squares [PLS] en anglais) basée sur les régressions linéaires. Une régression orthogonale (OPLS) peut également être utilisée. Dans ces modèles, une variable Y, nommée variable à expliquer, est exprimée en fonction d'une autre variable X, nommée variable explicative. La variable Y peut aussi bien être qualitative conduisant à l'application d'une PLS-DA (analyse discriminante) ou quantitative. Deux paramètres R^2Y et Q^2 sont indicateurs de la validité du modèle. Le premier coefficient est à rapprocher de l'appartenance à différentes classes, tandis que le second reflète l'erreur de prédiction. Plus ces derniers sont proches de 1 et plus le modèle est robuste. Un test de permutation est aussi usuellement utilisé pour renforcer la robustesse du modèle ainsi que la table ANOVA, en particulier le paramètre *p* qui indique la significativité du modèle. Il est d'usage d'interpréter une valeur de *p* inférieure à 0,05 comme le résultat d'un modèle significatif et exploitable (**Eriksson et al., 2008**). Ces modèles permettent de mettre en évidence les variables responsables de la discrimination des échantillons et couplés à des analyses biologiques, ils permettent d'identifier les variables potentiellement responsables de ces activités.

C. RESEAU MOLECULAIRE

Les réseaux moléculaires (« molecular networking ») sont une approche bio-informatique qui permet d'interpréter les données de spectres de masse qui peuvent être très étendues selon la complexité de l'extrait analysé. Cet outil a été développé par les groupes de Pieter C. Dorrestein et de Nuno Bandeira de l'université de Californie, de manière à aider l'interprétation des données provenant d'un mélange complexe comme le sont les extraits d'origine naturelle (**J. Y. Yang et al., 2013**). Il permet d'organiser les données de masse sous forme d'une carte de similarité spectrale formant un réseau mettant ainsi en évidence des groupes spectraux caractérisés par des similitudes de fragmentation. En effet, cet outil repose sur le postulat que des molécules structurellement proches vont partager des voies de fragmentations similaires et donc des fragments ou des pertes de neutre. Ainsi, les spectres d'un jeu de données sont comparés deux à deux par un alignement spectral et un score de similitude (cosine score), leur est attribué et va déterminer la proximité des deux molécules dans le réseau moléculaire qui sera formé. Cet alignement est basé sur l'application d'algorithmes prenant en compte les valeurs des rapports m/z ainsi que les intensités relatives des fragments. Le cosine score déterminé est compris entre 0 pour des spectres totalement différents et 1 pour des spectres identiques. Il est généralement acquis qu'un cosine score de 0,7 reflète une forte similarité spectrale. L'interprétation sous la forme de réseau moléculaire (Figure S3) permet de mettre en évidence des groupes d'analogues structuraux et de procéder à une *de novo* déreplication dans le cas de molécules inconnues fortement associées à des composés identifiés, par l'interprétation des pertes de masse.

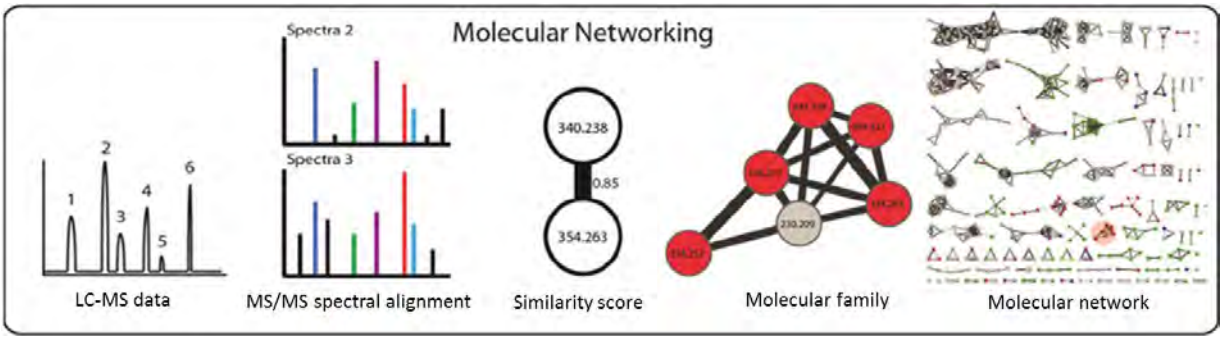


Figure S3. Cheminement de la construction d'un réseau (image issue de (Luzzatto-Knaan et al., 2015)

ANNEXE 2. VALIDATION DE LA METHODE D'ANALYSE SOUS CLOCHE

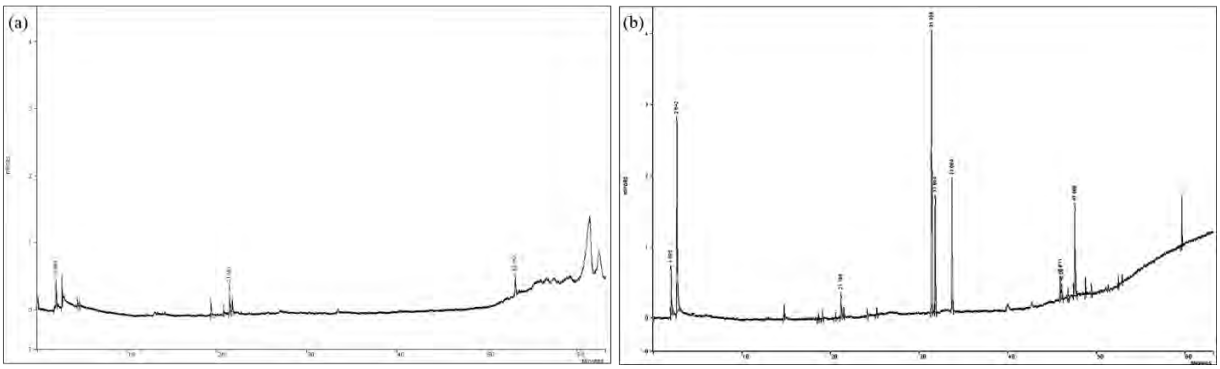


Figure S4. Profils volatils d'un plant de violette de Toulouse sans fleurs (a) et avec fleurs (b)

ANNEXE 3. STANDARDS GC-MS

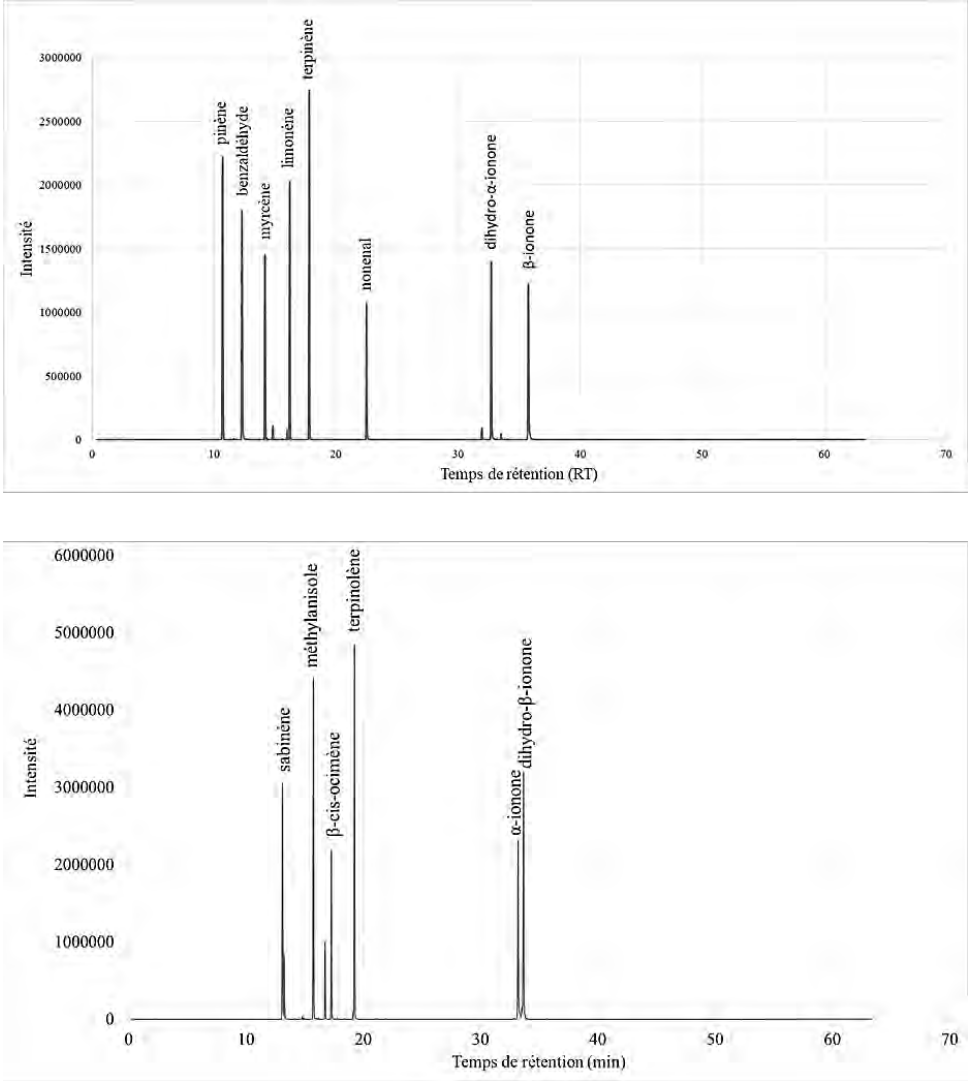
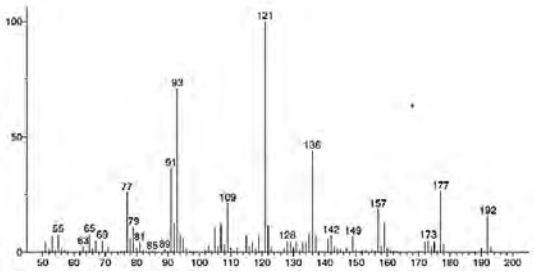
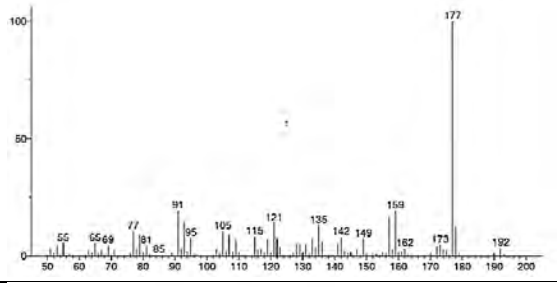
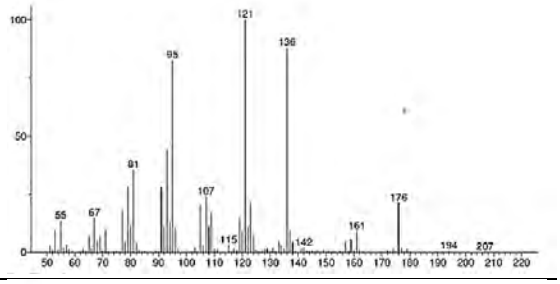
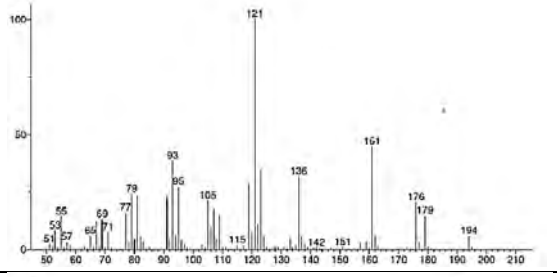
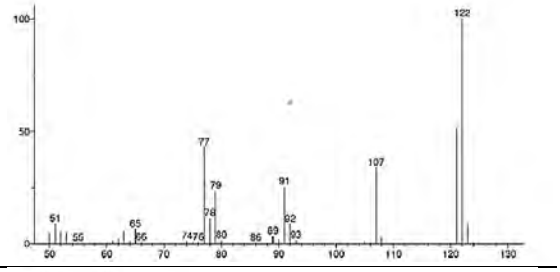
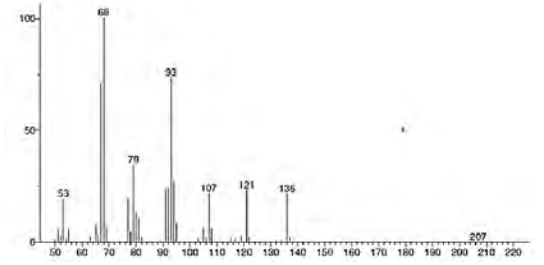
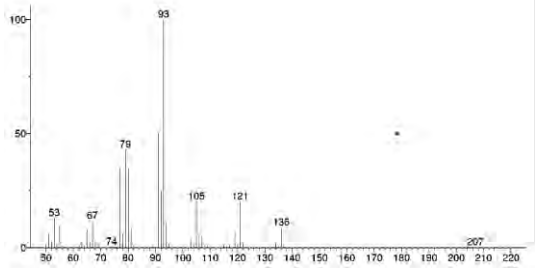


Figure S5. Profils chromatographiques des deux mixtes de standards anaylsés

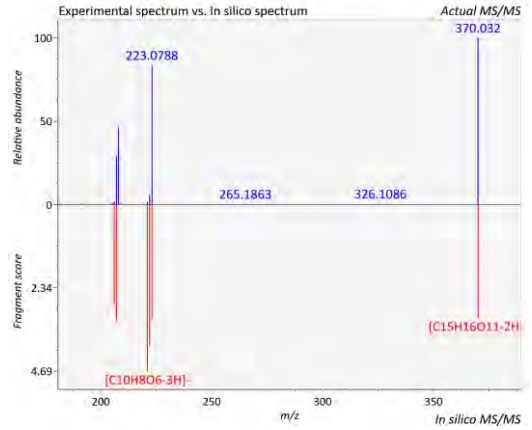
Tableau S1. Caractéristiques chromatographiques des composés identifiés chez les violettes et spectres MS/MS associés

Nom	Référence Sigma Alrich	RT standard (min)	RT extrait (min)	Ecart-type (%)	Spectre de masse expérimental du standard
α -ionone	I12409	33,2	33,3	7,1	

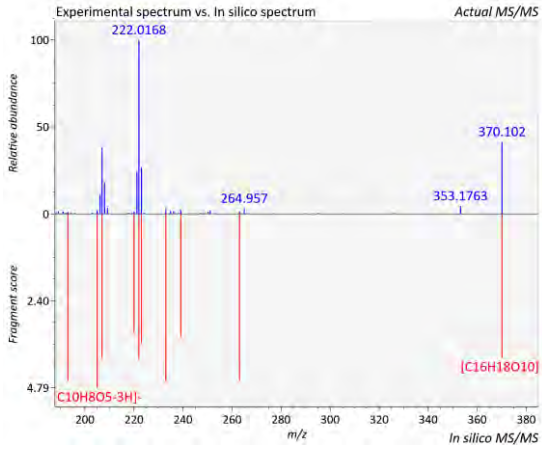
β -ionone	I12603	35,7	35,8	7,1	
Dihydro-α-ionone	00139	32,7	32,8	7,1	
Dihydro-β-ionone	W362603	33,7	33,8	7,1	
4-méthylanisole	148091	15,7	15,8	7,1	
Limonène	62118	16,2	16,3	7,1	
β-cis-ocimène		17,3	17,4	7,1	

ANNEXE 4. IDENTIFICATION *IN SILICO* DES BIOMARQUEURS NON VOLATILS DES FEUILLES

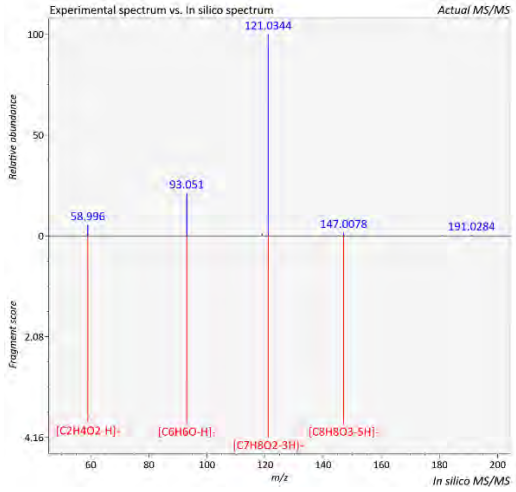
a) composé 1



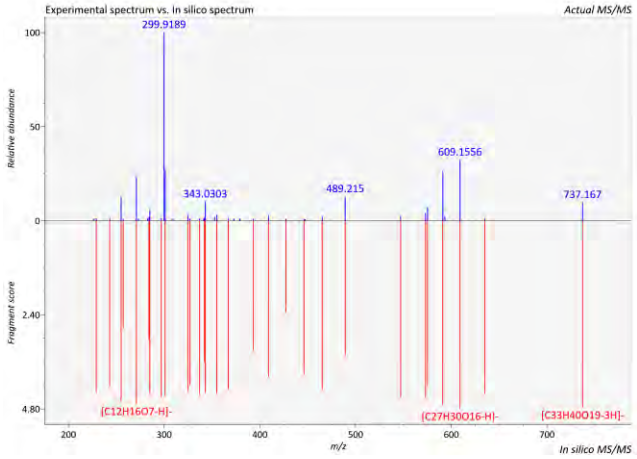
b) composé 2



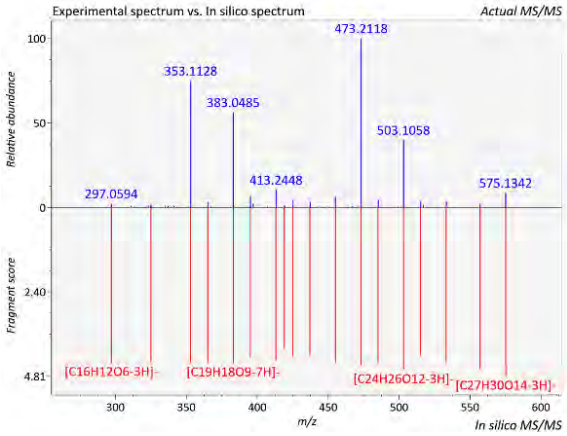
c) composé 3



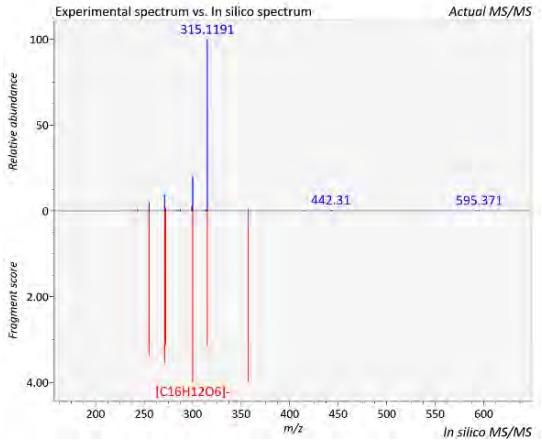
d) composé 4



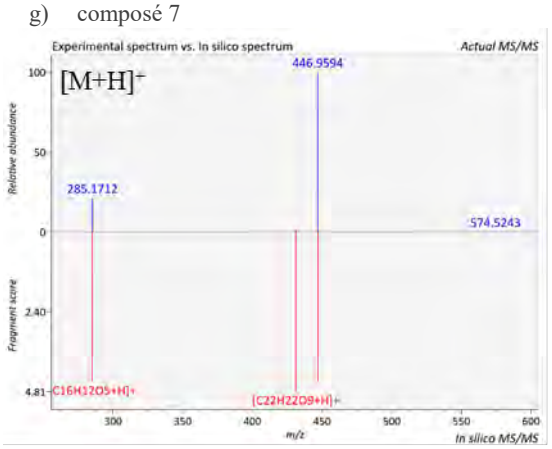
e) composé 5 (**Benayad et al., 2014**)

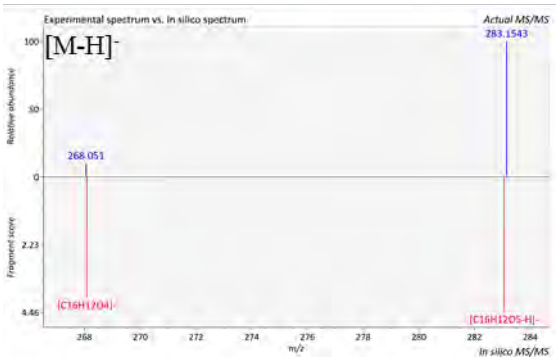


f) composé 6

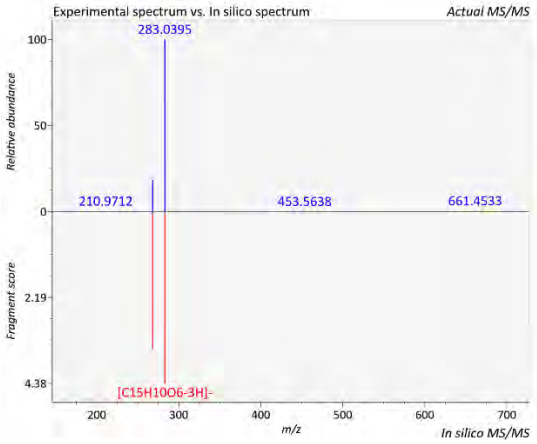


g) composé 7

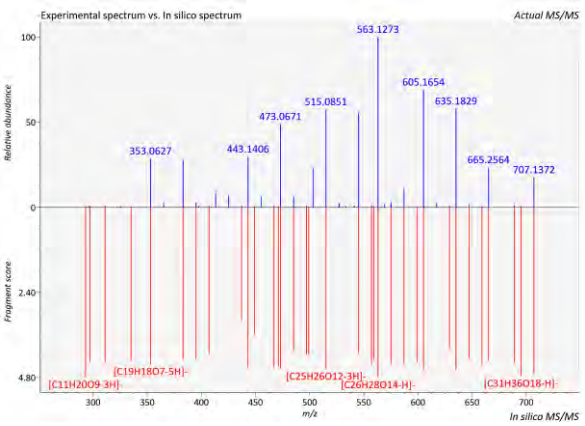




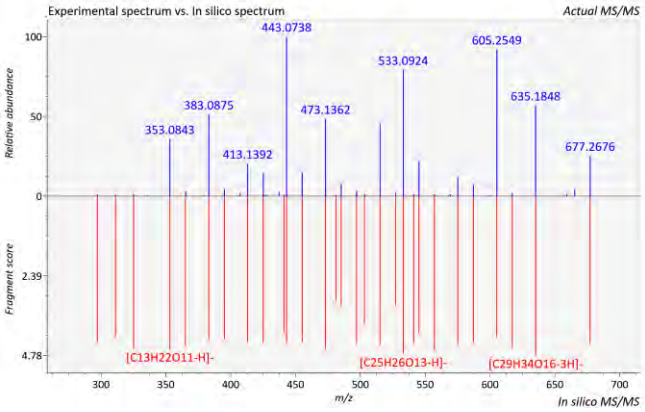
h) composé 8



i) composé 10



j) composé 11 (**Benayad et al., 2014**)



k) composé 9
[M+3H]3+
[M+4H]4+
803.3678
[M+4H]4+
[M+3H]3+
1070.8168
l) composé 11
[M+3H]3+
[M+4H]4+
817.1260
[M+4H]4+
[M+3H]3+
1089.1653
m) composé 12
[M+3H]3+
[M+3H]3+
1052.4648

ANNEXE 5. CARACTERISTIQUES ANATOMIQUES DES ESPECES DE VIOLETTE DE LA COLLECTION

Tableau S2. Caractéristiques anatomiques au niveau de la bractéole

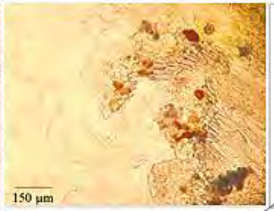
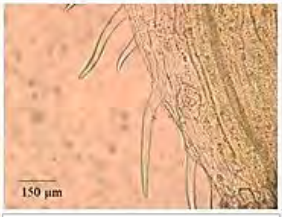
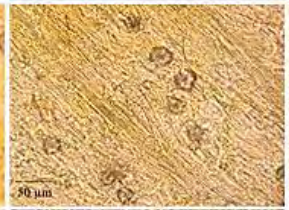
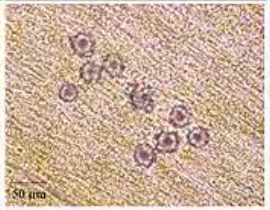
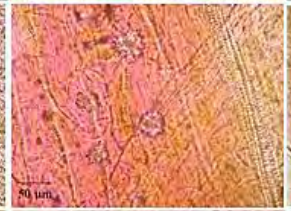
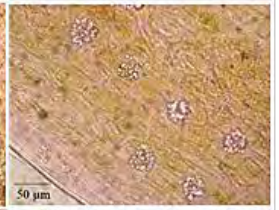
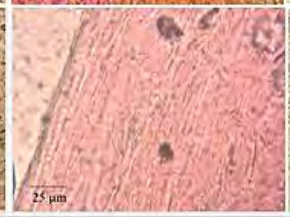
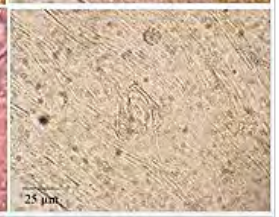
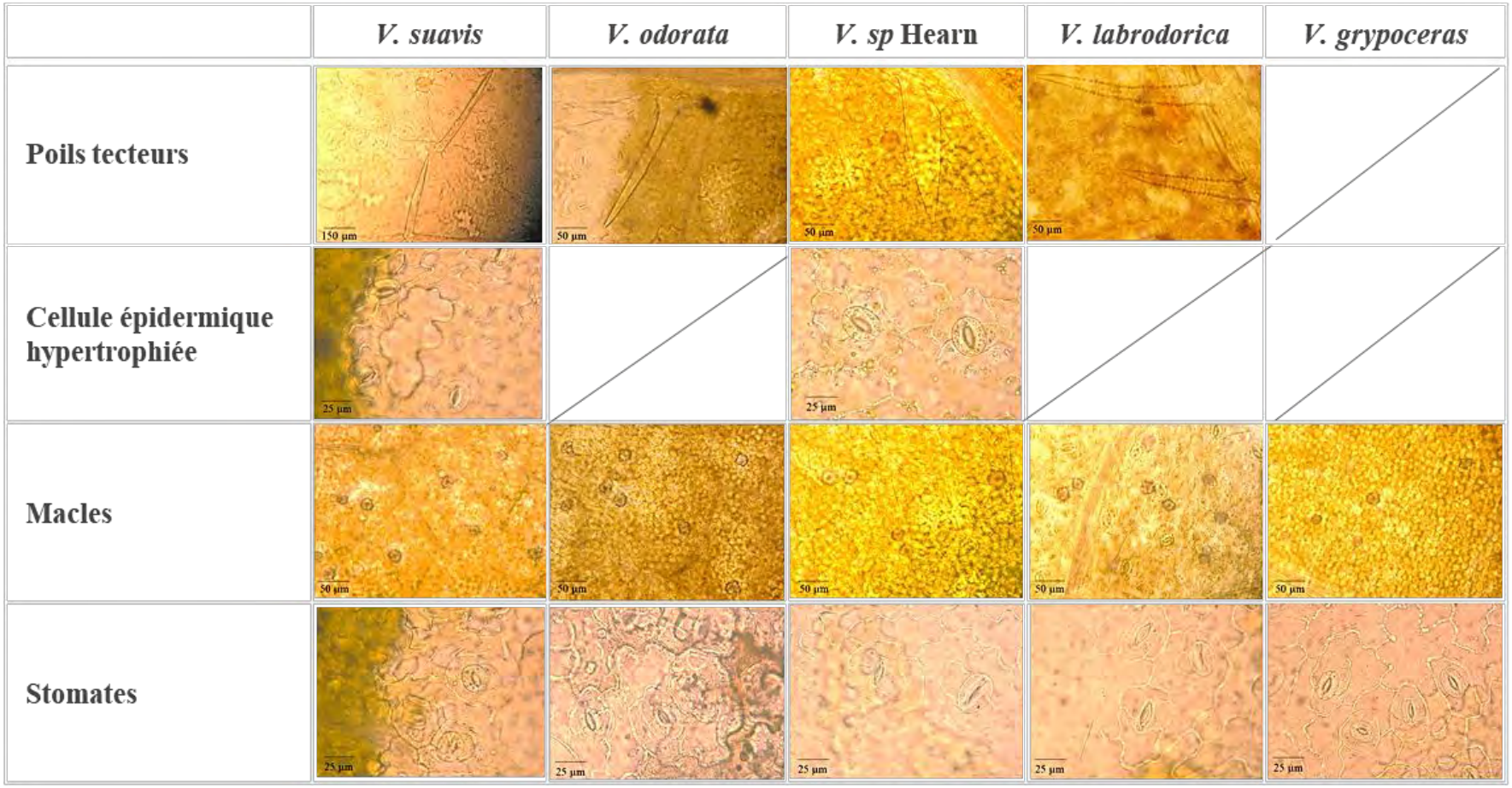
	V. suavis	*V. odorata*	*V. sp* Hearn	*V. labrodorica*	*V. grypoceras*
Poils sécréteurs					
Poils tecteurs					
Cellule à mucilage					
Macles					
Stomates					

Tableau S3. Caractéristiques anatomiques au niveau de l'épiderme



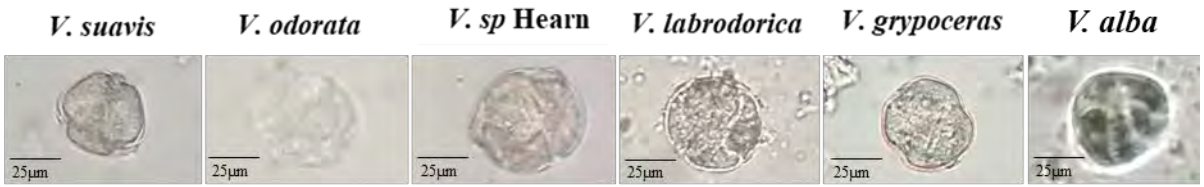
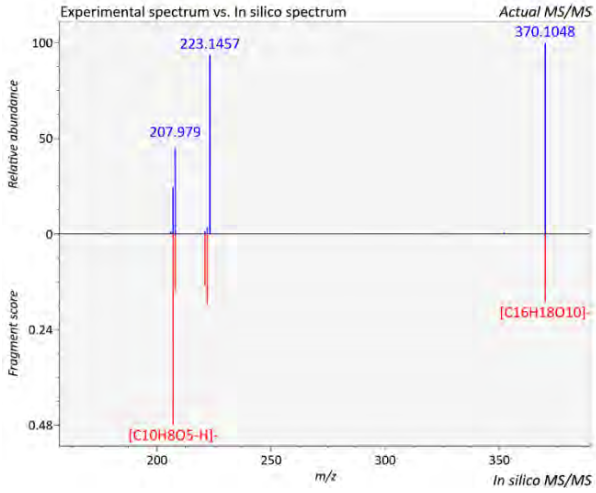


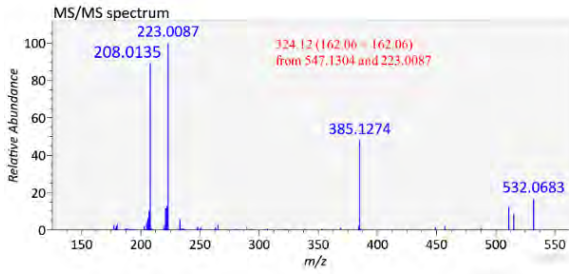
Figure S6. Grains de pollen observés chez les espèces de violette de la collection identifiées (x400)

ANNEXE 6. COMPARAISON DES SPECTRES DE MASSE EXPERIMENTAUX ET *IN SILICO* DES COMPOSES

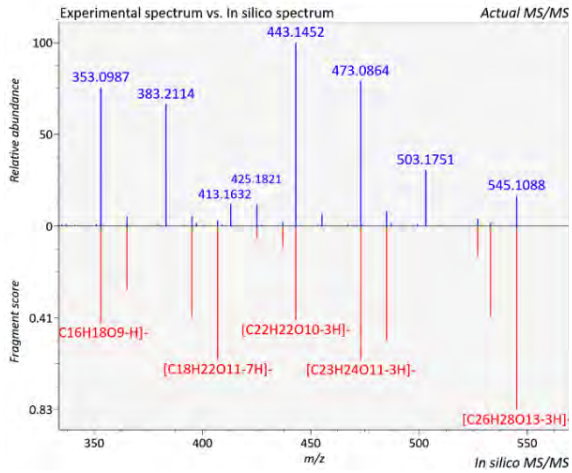
a) 7-Methoxy-5,6,8-trihydroxycoumarin-5-β-glucopyranoside (composés **1** et **2**)



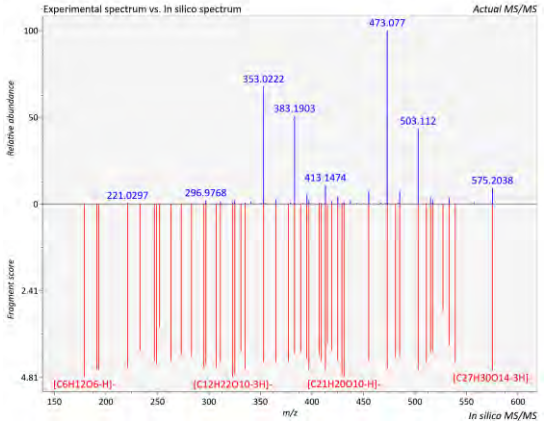
b) 7,8-dimethoxy-5,6,8-trihydroxycoumarin-5-β glucopyranosyl-glucopyranoside (compound **20**)



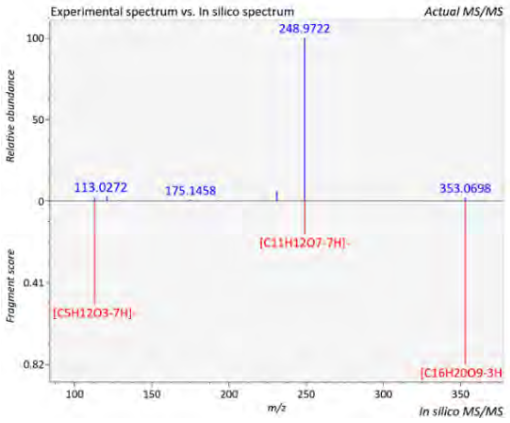
c) Schaftoside (composé **21**) (**Ferrere et al., 2003**)



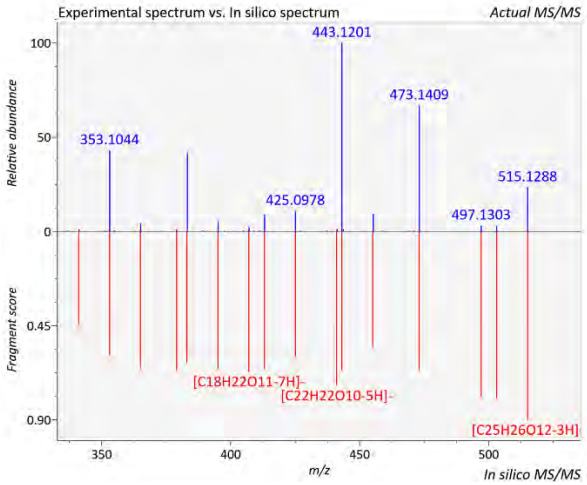
d) Isovitexin 6"-O-beta-D-glucopyranoside (composé **5**) (**Benayad et al., 2014**)



e) 8-*C*-Glucosyl-3,4-dihydro-5,7-dihydroxy-6-methoxycoumarin (composé **22**)



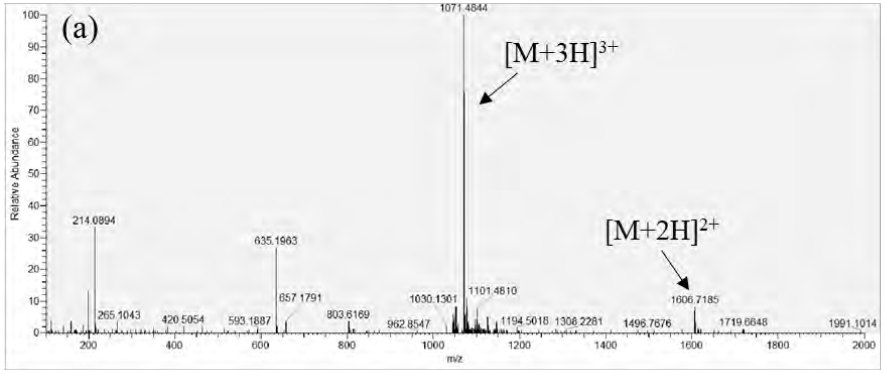
f) 6,8-Di-*C*-α-L-arabinopyranosylapigenin (composé **19**)

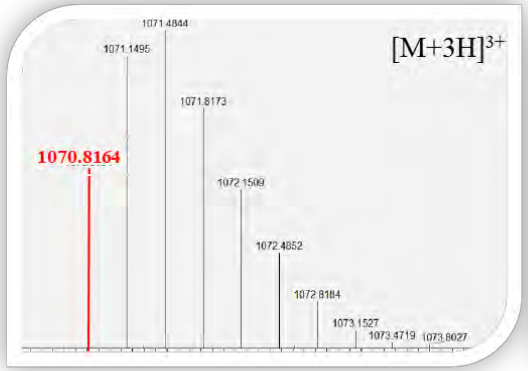


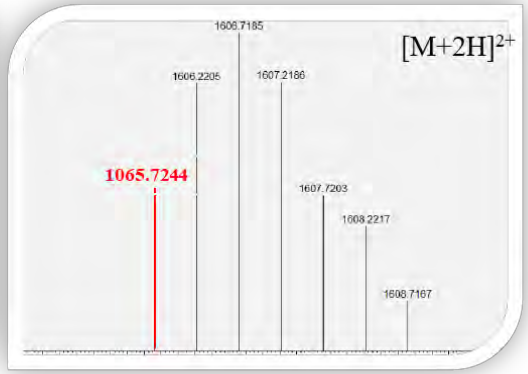
ANNEXE 7. SPECTRE DE MASSE DES CYCLOTIDES IDENTIFIES

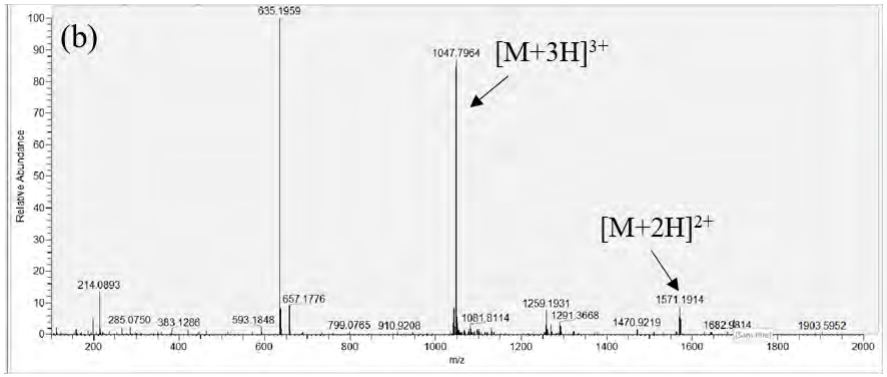
Table S4. Caractéristiques des cycloviolacines et du cyclotide Cter D identifiés chez certaines violettes de la collection

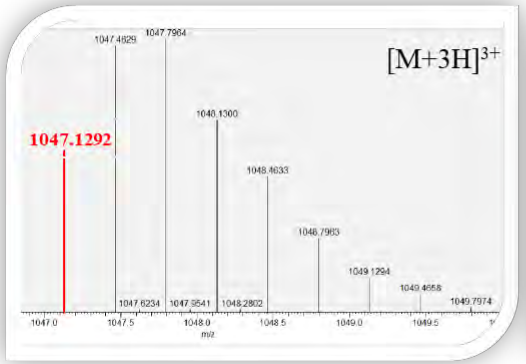
Nom	N°	Formule brute	Masse exacte	$[M+2H]^{2+}$	$[M+3H]^{3+}$	$[M+4H]^{4+}$	$[M+5H]^{5+}$
Cter D	12	$C_{142}H_{221}N_{35}O_{41}S_6$	3264,4609	1633,2384	1089,1616	817,1232	653,9001
Cycloviolacine O_{11}	26	$C_{138}H_{216}N_{36}O_{40}S_6$	3209,4299	1605,7229	1070,8179	803,3654	642,8939
Cycloviolacine O_{13}	27	$C_{133}H_{207}N_{37}O_{38}S_6$	3122,3727	1562,1943	1041,7988	781,6011	625,4825
Cycloviolacine O_4	28	$C_{134}H_{209}N_{37}O_{39}S_6$	3152,3833	1577,1996	1051,8024	789,1038	631,4846
Cycloviolacine O_9	14	$C_{133}H_{207}N_{37}O_{39}S_6$	3138,3676	1570,1917	1047,1305	785,5998	628,6815

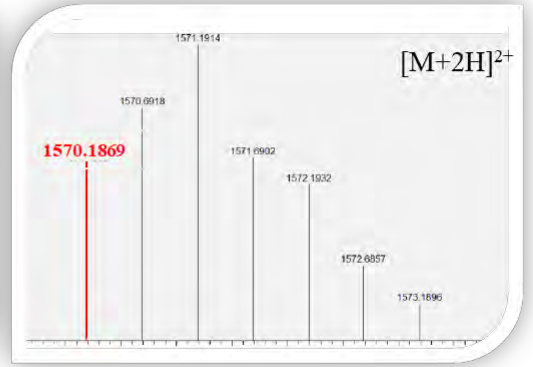


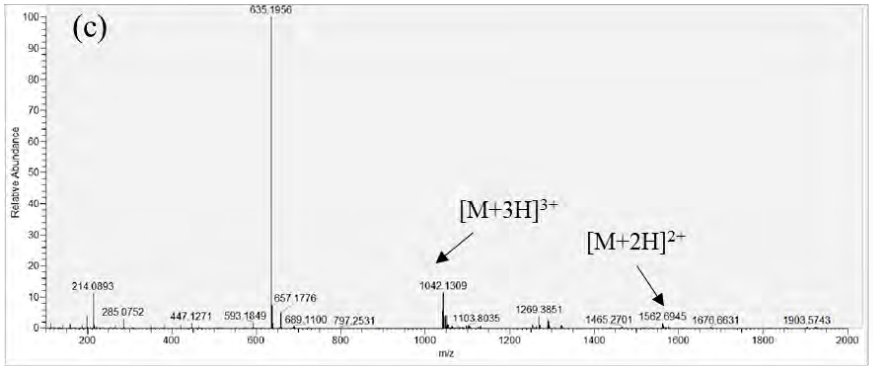


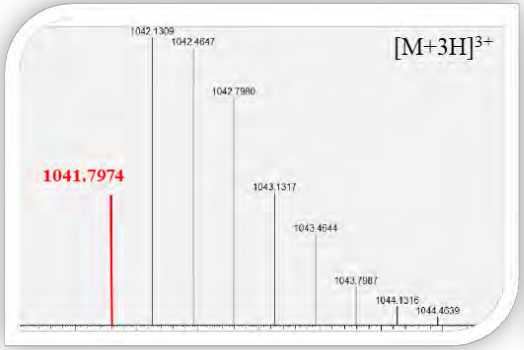


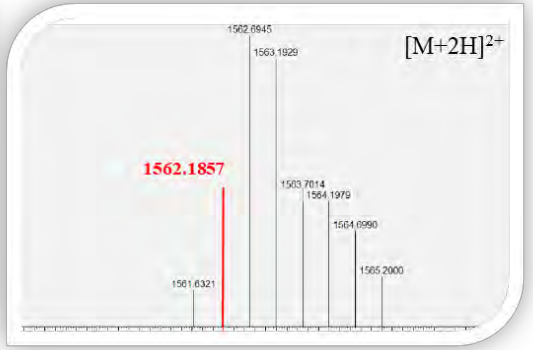

(b)
[M+3H]3+
[M+2H]2+
Relative Abundance
m/z
635.1959
1047.7964
214.0893
285.0750
383.1288
593.1848
657.1776
799.0765
910.9208
1081.8114
1259.1931
1291.3668
1470.9219
1571.1914
1903.5952


[M+3H]3+
1047.1292
1047.4629
1047.7964
1048.1300
1048.4633
1048.7963
1049.1294
1049.4658
1049.7974


[M+2H]2+
1570.1869
1570.6918
1571.1914
1571.6902
1572.1932
1572.6857
1573.1896


(c)
[M+3H]3+
[M+2H]2+
Relative Abundance
m/z
635.1956
214.0893
285.0752
447.1271
593.1849
657.1776
689.1100
797.2531
1042.1309
1103.8035
1269.3851
1465.2701
1562.6945
1676.6631
1903.5743


[M+3H]3+
1041.7974
1042.1309
1042.4647
1042.7980
1043.1317
1043.4644
1043.7987
1044.1316
1044.4639


[M+2H]2+
1562.1857
1562.6945
1563.1929
1563.7014
1564.1979
1564.6990
1565.2000
1561.6321

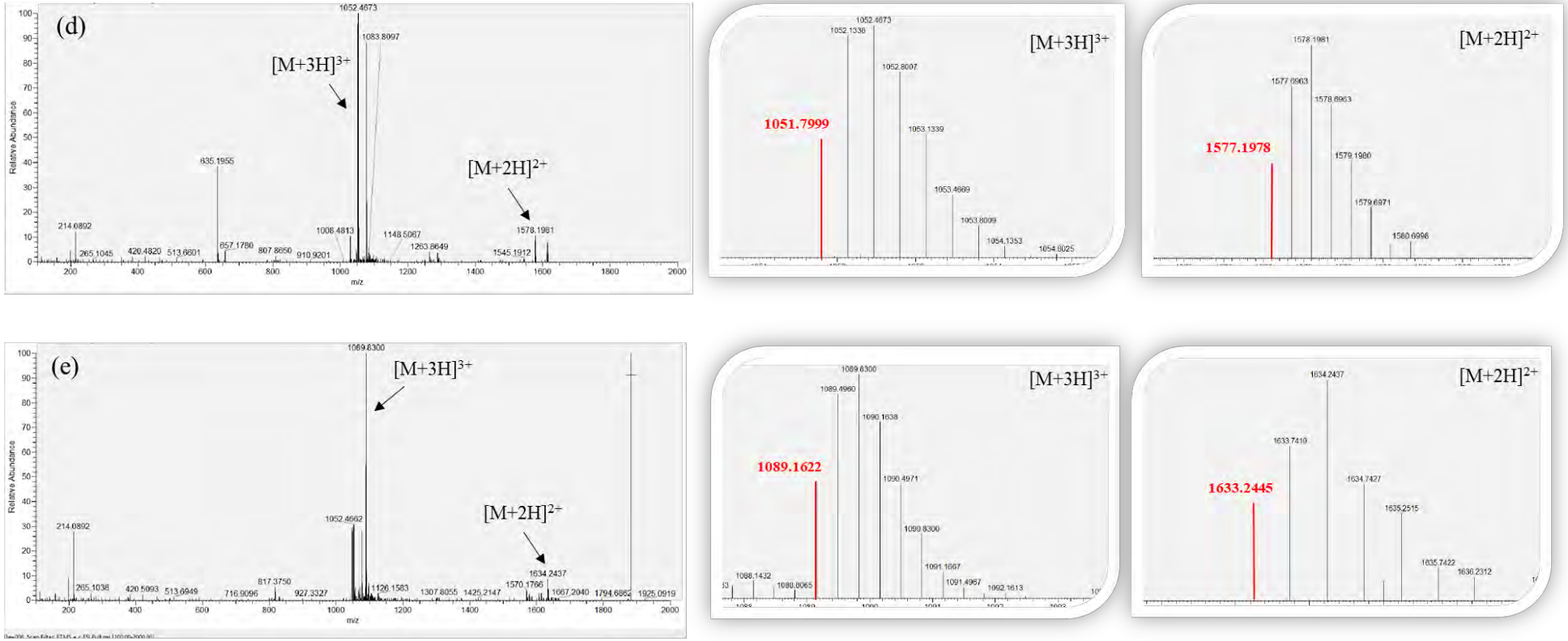
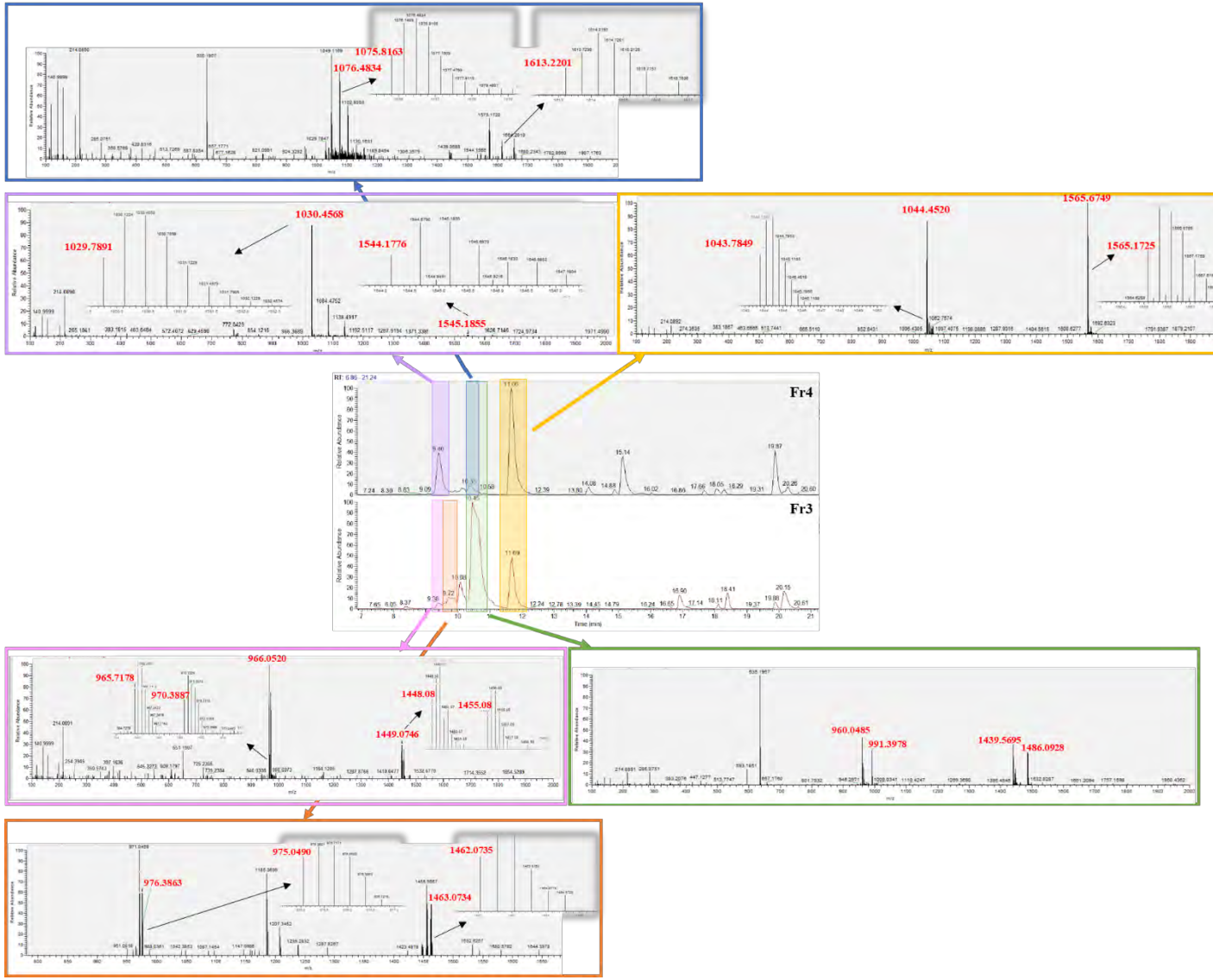


Figure S7. Spectres de masse et profils isotopiques des fragments $[M+2H]^{2+}$ et $[M+3H]^{3+}$ des cycloviolacines O11 (a), O9 (b), O3 (c) et O4 (d) ainsi que du cyclotide Cter D (e)

ANNEXE 8. SPECTRE DE MASSE DES CYCLOTIDES OBTENUS PAR SPE DE *V. sp* HEARN



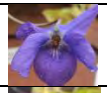
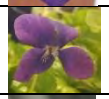
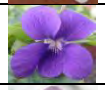
ANNEXE 9. LISTE DES STANDARDS

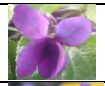
Nom	Classe chimique	Formule brute	Masse exacte	Détection	RT (min)
Acide chlorogenique	Aromatique simple	$C_{16}H_{18}O_9$	354,095	neg/pos	6,43
Acide ferulique	Aromatique simple	$C_{10}H_{10}O_4$	194,184	pos	10,52
Acide paracoumarique	Aromatique simple	$C_9H_8O_3$	164,047	pos	9,08
Ajmaline	Aromatique simple	$C_{20}H_{26}N_2O_2$	326,199	pos	10,11
Aloin	Drogue	$C_{21}H_{22}O_9$	418,126	pos	15,59
Arborinine	Alcaloïde	$C_{16}H_{15}NO_4$	285,1	pos	23,32
Arbutin	Aromatique simple	$C_{12}H_{16}O_7$	272,09	neg	1,36
Arecoline	Alcaloïde	$C_8H_{13}NO_2$	155,095	pos	0,98
Berberine	Alcaloïde	$C_{20}H_{18}NO_4^+$	336,124	pos	16,28
Bergapten	Benzopyranoïde	$C_{12}H_8O_4$	216,042	pos	20,33
Cafeine	Alcaloïde	$C_8H_{10}N_4O_2$	194,08	pos	6,78
Cinchonin	Alcaloïde	$C_{19}H_{22}N_2O$	294,173	pos	7,1
Cucurbitacin E	Terpenoïde	$C_{32}H_{44}O_8$	556,304	pos	27,9
Ellagic acid	Tanin	$C_{14}H_6O_8$	302,006	neg	11
Emodine	Aromatique polycyclique	$C_{15}H_{10}O_5$	270,053	neg	28,89
Escin	Terpenoïde	$C_{55}H_{86}O_{24}$	1130,551	neg/pos	26,33
Esculin	Benzopyranoïde	$C_{15}H_{16}O_9$	340,079	neg/pos	4,9
Eupatorin	Aromatique polycyclique	$C_{18}H_{16}O_7$	344,09	neg/pos	23,1
Flavone	Flavonoïde	$C_{15}H_{10}O_2$	222,068	pos	25,21
Fraxetin	Benzopyranoïde	$C_{10}H_8O_5$	208,037	neg/pos	8,78
Glutathion	Alcaloïde	$C_{10}H_{17}N_3O_6S$	307,084	neg/pos	0,98
Harpagoside	Terpenoïde	$C_{24}H_{30}O_{11}$	494,179	neg	18,22
Hypericin	Aromatique polycyclique	$C_{30}H_{16}O_8$	504,085	neg	14,73
Isoquercitrin	Flavonoïde	$C_{21}H_{20}O_{12}$	464,095	neg/pos	11,83
L-Asparagine	Acide aminé	$C_4H_8N_2O_3$	132,053	neg/pos	0,9
Lobeline	Drogue	$C_{22}H_{27}NO_2$	337,204	pos	17,43
L-picein	Aromatique simple	$C_{14}H_{18}O_7$	298,105	neg/pos	4,57
L-Proline	Acide aminé	$C_5H_9NO_2$	115,063	pos	0,93
Ouabaine	Stéroïde	$C_{29}H_{44}O_{12}$	584,283	neg/pos	9,55
Phloridzin	Chalcone	$C_{21}H_{24}O_{10}$	436,137	neg/pos	14,73
Phloroglucinol	Aromatique simple	$C_6H_6O_3$	126,032	neg/pos	1,44
Protocatechic	Aromatique simple	$C_7H_6O_3$	138,032	neg/pos	4,67
Protopine	Alcaloïde	$C_{20}H_{19}NO_5$	353,126	pos	13,07
Quassine	Flavonoïde	$C_{22}H_{28}O_6$	388,189	pos	18,66
Quercetine	Flavonoïde	$C_{15}H_{10}O_7$	302,043	neg/pos	16,53
Quercitrin	Flavonoïde	$C_{21}H_{20}O_{11}$	448,101	neg/pos	13,26
Quinidine	Alcaloïde	$C_{20}H_{24}N_2O_2$	324,184	pos	8,46
Quinine	Flavonoïde	$C_{27}H_{30}N_2O_5$	324,184	pos	8,97
rosmarinic acid	Aromatique simple	$C_{18}H_{16}O_8$	360,085	neg/pos	14,19
Rutoside	Flavonoïde	$C_{27}H_{30}O_{16}$	610,153	neg/pos	11,53
Theobromine	Alcaloïde	$C_7H_8N_4O_2$	180,065	pos	2,58
Theophyline	Alcaloïde	$C_7H_8N_4O_2$	180,065	neg	4,01
Tryptophane	Acide aminé	$C_{11}H_{12}N_2O_2$	204,09	neg/pos	4,19
Umbelliferone	Benzopyranoïde	$C_9H_6O_3$	162,032	neg/pos	14,19

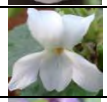
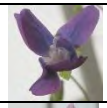
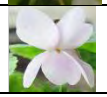
ANNEXE 10. RESUME DES DONNEES OBTENUES SUR CHAQUE VIOLETTE DE LA COLLECTION

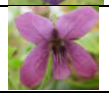
			Identification génétique par ITS			Pourcentage d'inhibition ± SD (%)					
N°	Nom	Fleur	Espèce	Score d'id.	DPPH	*Colleto.*	*Fusa.*	*Botrytis*	*Alternaria*	*Phytophthora*	Coloration SDP
1	Doreen		*Viola odorata isolate ODO178 (V11)*	99,37	30,5 ± 3,9	- 31 ± 15	- 42 ± 8	98 ± 0	12 ± 7	72 ± 9	
2	Alice		*Viola odorata isolate ODO214 (V8)*	100	27,3 ± 3,5	- 17 ± 24	- 9 ± 5	99 ± 0	125 ± 7	98 ± 2	
3	Viola Baronne de Rothschild		*Viola suavis isolate SUB211 (V7)*	98,69 - 99	29,3 ± 3,1	- 31 ± 41	- 37 ± 2	97 ± 0	122 ± 7	92 ± 8	
4	Red Alice				37,9 ± 4,2	- 62 ± 17	17 ± 3	29 ± 8			
5	Viola la France				30,6 ± 2	- 70 ± 5	- 29 ± 6	13 ± 14			
6	Charles Winston Groves		*Viola odorata isolate ODO214 (V8)*	100	22,5 ± 2,6	- 36 ± 11	- 36 ± 7	100 ± 0	117 ± 11	99 ± 3	
7	Viola Jean Arnaud		*Viola odorata isolate ODO182 (V3)*	100	63 ± 3,4	- 75 ± 18	18 ± 5	28 ± 17			1 ± 0
8	Grey Lady		*Viola odorata isolate ODO178 (V11)*	100	41,6 ± 1,9	- 76 ± 24	24 ± 1	14 ± 31			1,3 ± 0,6
9	Viola Perle rose		*Viola odorata isolate ODO214 (V8)*	100	60,4 ± 0,2	- 89 ± 3	19 ± 6	- 2 ± 9			1 ± 0
10	Viola Elisabeth Lee		*Viola odorata isolate ODO182 (V3)*	99,84	28,1 ± 3,6	- 38 ± 8	27 ± 5	8 ± 20			
11	Viola Jack Sampson		*Viola odorata isolate ODO178 (V11)*	100	35,5 ± 4,1	- 86 ± 37	21 ± 9	- 6 ± 11			1 ± 0

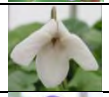
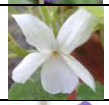
12	Annie		*Viola tokubuchiana f. albiflora*	79,55	42,3 ± 1,6	- 67 ± 8	16 ± 2	99 ± 0	4 ± 3	73 ± 4	
13	Viola Rouge Prouteau		*Viola odorata isolate ODO178 (V11)*	100	34,7 ± 2	- 60 ± 6	24 ± 11	- 21 ± 11			0,5 ± 0,7
14	Viola Lydia Groves				22,9 ± 1,3	- 69 ± 3	15 ± 5	1 ± 5			2 ± 1
15	Malmsbury pink		*Viola odorata isolate ODO214 (V8)*	100	26,4 ± 3,6	- 77 ± 6	19 ± 11	- 5 ± 19			0,7 ± 0,6
16	Blanche de chevreuses		*Viola odorata isolate ODO178 (V11)*	100	49,3 ± 1	- 69 ± 11	17 ± 5	14 ± 19			1,3 ± 0,6
17	Rubra odorata		*Viola odorata isolate ODO214 (V8)*	100	33,4 ± 0,1	- 48 ± 6	11 ± 4	2 ± 3			0,7 ± 0,6
18	Viola Wren's Pink		*Viola odorata isolate ODO182 (V3)*	99,65	18,4 ± 0,7	- 60 ± 30	21 ± 9	- 3 ± 15			0,5 ± 0,7
19	Viola Comte de Brazza		*Viola alba isolate ALB184 (V6)*	85,06	48,1 ± 1,1	- 104 ± 34	- 13 ± 7	98 ± 0	8 ± 7	16 ± 3	
20	Joan Yardley		*Viola odorata isolate ODO214 (V8)*	100	31,2 ± 1	- 64 ± 6	27 ± 12	12 ± 11			0,7 ± 0,6
21	Viola rose d'Amérique du Sud		*Viola odorata isolate ODO182 (V3)*	100	41,2 ± 1,8	- 75 ± 11	9 ± 10	- 12 ± 20			1,3 ± 0,6
22	Viola cœur d'Alsace		*Viola odorata isolate ODO178 (V11)*	100	18,9 ± 0,7	- 66 ± 20	19 ± 6	4 ± 17			1 ± 0
23	Viola Naepolitean		*Viola odorata isolate ODO182 (V3)*	100	25,9 ± 0,6	- 92 ± 26	6 ± 11	- 11 ± 15			0,7 ± 0,6
24	Viola labradorica		*Viola labradorica*	100	78,9 ± 0,1	- 85 ± 3	- 3 ± 10	- 6 ± 25			0,5 ± 0,7
25	Viola Princesse de galles				28,1 ± 1,6	- 59 ± 13	24 ± 5	7 ± 15			1 ± 0

26	Viola souvenir de ma fille		*Viola odorata isolate ODO178 (V11)*	100	20,1 ± 0,2	- 88 ± 10	12 ± 2	10 ± 9			1,7 ± 0,6
27	Crepuscule					- 45 ± 11	40 ± 4	80 ± 8	118 ± 9	96 ± 3	
28	Charente		*Viola odorata isolate ODO182 (V3)*	99,82	24,7 ± 1,1	- 57 ± 9	8 ± 4	11 ± 14			1,3 ± 0,6
29	Viola Pinehurst		*Viola sp. Hearn cult33*	99,84	32,1 ± 1,1	- 65 ± 8	21 ± 10	- 10 ± 15			1 ± 0
30	Viola sororia dark freckles		*Viola sp. Hearn cult33*	99,51	17,4 ± 1,6	2 ± 20	100 ± 0	20 ± 8	- 38 ± 89	36 ± 10	2 ± 0
31	Sororia blanche		*Viola sp. Hearn cult33*	99,51	9,2 ± 0,5	26 ± 7	97 ± 9	115 ± 31	124 ± 3	96 ± 5	
32	Princess Alexander		*Viola sp. Hearn cult33*	99,84		- 88 ± 9	19 ± 12	- 15 ± 9			0,7 ± 0,6
33	Viola Nora Church				29,5 ± 0,7	- 40 ± 8	- 51 ± 2	99 ± 0	- 83 ± 13	11 ± 9	
34	Viola Green Jade		*Viola mandshurica*	99,85	6,2 ± 1,4	- 1 ± 2	101 ± 1	20 ± 10	118 ± 12	95 ± 5	1 ± 0
35	Violette chinoise		*Viola sp. Hearn cult33*	100		- 65 ± 6	26 ± 5	7 ± 10			2 ± 0
36	Ash Vale Blue		*Viola alba isolate ALB184 (V6)*	98,69	40,6 ± 1,5	- 64 ± 41	45 ± 6	18 ± 15	98 ± 26	- 13 ± 16	0,7 ± 0,6
37	Hopley's white		*Viola alba isolate ALB184 (V6)*	92,12	78,2 ± 0,7	- 50 ± 4	21 ± 7	- 35 ± 48			3 ± 1
38	Dick of the Hill		*Viola acuminata*	97,06	15,9 ± 0,6	44 ± 11	6 ± 4	- 20 ± 30			2 ± 0
39	Vin André Thorp		*Viola odorata isolate ODO182 (V3)*	96,41	31,3 ± 1,1	- 67 ± 29	16 ± 3	15 ± 10			1,3 ± 0,6
40	Alice Watter		*Viola odorata isolate ODO214 (V8)*	100		- 4 ± 18	17 ± 6	20 ± 9			2 ± 0
41	Le Gresley					- 22 ± 25	34 ± 2	- 10 ± 21	- 14 ± 143	18 ± 2	1,7 ± 0,6

42	Viola California		*Viola odorata isolate ODO182 (V3)*	99,82	29,8 ± 2,3	- 64 ± 22	19 ± 6	- 32 ± 39			1,7 ± 0,6
43	Viola Eizanensis					- 37 ± 29	52 ± 41	4 ± 6	35 ± 11	11 ± 14	2,7 ± 0,6
44	Viola substinuata					- 67 ± 6	7 ± 2	4 ± 13			2,3 ± 0,6
46	Odorata mauve		*viola alba subsp dehnhardtii*	100	21,6 ± 0,9	- 53 ± 6	21 ± 2	2 ± 5			1 ± 0
47	Viola Reid's Crimpson carpet		*Viola odorata isolate ODO178 (V11)*	100	19,7 ± 1,1	- 23 ± 3	31 ± 5	18 ± 15			0,7 ± 0,6
48	Viola Marquis de Brazais		*Viola sp. Hearn cult33*	100	28,8 ± 1	- 96 ± 2	1 ± 3	- 13 ± 25			2 ± 0
49	Viola Mirabilis				77,6 ± 0,2	- 88 ± 4	10 ± 7	- 1 ± 28			1,5 ± 0,7
50	Carol Lockston				11 ± 1,5	- 76 ± 40	19 ± 2	20 ± 2			1 ± 0
52	Viola Sulfurea		*Viola odorata isolate ODO214 (V8)*	100	13,8 ± 1,7	- 123 ± 29	6 ± 3	- 2 ± 11			2 ± 0
53	Viola Cendrillon				38,8 ± 4,1	- 28 ± 34	- 10 ± 6	- 5 ± 23			0,7 ± 0,6
54	Viola Lady Jane		*Viola odorata isolate ODO214 (V8)*	100	17,4 ± 1,9	- 87 ± 19	18 ± 4	8 ± 10			0,7 ± 0,6
55	Viola odorata ancienne		*Viola odorata isolate ODO214 (V8)*	100	41,5 ± 2,4	- 43 ± 24	19 ± 5	34 ± 18			
56	Irish Elegance		*Viola alba subsp. dehnhardtii*	100	21,6 ± 2,1	- 60 ± 58	21 ± 10	10 ± 3			0,5 ± 0,7
57	Viola the Czar		*Viola suavis isolate SUB211 (V7)*	99,35	19,5 ± 4	- 50 ± 40	12 ± 3	4 ± 11			1,3 ± 0,6
59	Guilleto Famin		*Viola acuminata*	96,9	44 ± 5,5	- 52 ± 8	17 ± 2	- 10 ± 10			0,5 ± 0,7

60	Viola Valentré				29,3 ± 3,7	- 64 ± 42	14 ± 15	- 21 ± 30			2,7 ± 0,6
61	Mother Pearl		*Viola odorata isolate ODO178 (V11)*	100	39,2 ± 2,2	- 47 ± 6	16 ± 7	11 ± 12			2 ± 1
62	Christmas		*Viola odorata isolate ODO214 (V8)*	99,84	36,2 ± 0,3	- 19 ± 6	53 ± 3	- 1 ± 13			1,7 ± 0,6
63	Parchment		*Viola kusanoana voucher Yoo863*	97,56	39,3 ± 1,1	- 58 ± 12	25 ± 4	- 11 ± 8			2 ± 0
64	Opera		*Viola odorata isolate ODO214 (V8)*	100	19,8 ± 1,7						
65	Viola Mme DUMAS					- 21 ± 22	48 ± 8	27 ± 17	6 ± 5	8 ± 16	1,7 ± 0,6
66	Mulberry				24,6 ± 2,7	- 39 ± 18	23 ± 4	- 4 ± 13			1,3 ± 0,6
67	Toscane		*Viola odorata isolate ODO178 (V11)*	99,67	24,6 ± 2,4	- 86 ± 8	18 ± 12	3 ± 14			3 ± 0
68	Viola San Remo		*Viola suavis isolate SUB211 (V7)*	98,85 - 99,8	25,4 ± 2,3	- 27 ± 18	32 ± 6	15 ± 1	14 ± 6	97 ± 2	
69	Viola Mrs Barton		*Viola odorata isolate ODO178 (V11)*	99,38	56,4 ± 1,6	- 71 ± 1	29 ± 8	- 5 ± 4			2,7 ± 1,2
70	Viola rose Barandou		*Viola odorata isolate ODO182 (V3)*	99,84	24,1 ± 0,8	- 30 ± 9	24 ± 10	19 ± 23			2 ± 0
71	Phill Doves		*Viola odorata isolate ODO182 (V3)*	97,72	29 ± 0,7	- 52 ± 13	17 ± 7	- 10 ± 8			1,7 ± 0,6
72	Viola De Bruneau					- 50 ± 6	29 ± 4	3 ± 13			3 ± 1
73	Viola Hirta					- 39 ± 8	54 ± 4	21 ± 11			2 ± 1

75	Lee's peachy pink					- 59 ± 12	30 ± 5	32 ± 16			1,7 ± 0,6
76	Viola Marie rose				35,6 ± 2,2	- 20 ± 19	28 ± 6	0 ± 4			1 ± 0
77	Cyclope		*Viola suavis isolate SUB211 (V7)*	99,84	18 ± 1,1	- 44 ± 15	26 ± 13	- 19 ± 25			1 ± 0
78	Helena		*Viola odorata isolate ODO182 (V3)*	99,84	27,2 ± 1,7	- 56 ± 11	24 ± 2	14 ± 16			1,7 ± 0,6
79	Elisabeth Bailes		*Viola odorata voucher NMW6028*	85,17	21,8 ± 0,6	- 51 ± 22	51 ± 31	58 ± 19	121 ± 5	78 ± 4	
80	Grypoceras variegata		*Viola grypoceras var. exilis*	99,16	78,5 ± 0,5	- 68 ± 3	20 ± 2	18 ± 14			2 ± 0
81	Viola peppered palm					- 85 ± 2	22 ± 5	15 ± 10			1 ± 0
82	Viola sororia blue diamond					- 60 ± 10	19 ± 5	6 ± 21			3,3 ± 1,2
83	Viola palmata					13 ± 33	101 ± 1	22 ± 10	127 ± 14	98 ± 4	3 ± 0
84	Viola sororia Binsted speckles		*Viola sp. Hearn cult33*	100		- 98 ± 18	12 ± 6	9 ± 22			1 ± 0
85	Viola striatta		*Viola labradorica*	98,53	37,1 ± 1,5	- 44 ± 28	32 ± 7	- 18 ± 15	123 ± 1	96 ± 3	2,7 ± 1,2
86	Royal elck		*Viola sp. Hearn cult33*	99,84	34,6 ± 1,4	- 67 ± 12	11 ± 5	0 ± 5			
87	Susan Chilcott				21,3 ± 1,4	- 56 ± 10	25 ± 7	30 ± 17			2 ± 0
88	Emile Prevot		*Viola sp. Hearn cult33*	99,84	31,9 ± 1	- 39 ± 11	5 ± 19	- 2 ± 6			1,3 ± 0,6
89	Dawnie		*Viola odorata isolate ODO214 (V8)*	100	38,7 ± 2,4	- 77 ± 7	29 ± 12	- 2 ± 22			3 ± 0

90	Viola Lavender Lady		*Viola odorata isolate ODO214 (V8)*	100	31,1 ± 4,4	- 81 ± 16	30 ± 6	20 ± 15			1,3 ± 0,6
91	Luxonne		*Ramularia glennii*	90,7	23,6 ± 5,4	- 53 ± 9	20 ± 5	5 ± 2			2 ± 0
92	Maroc		*Viola suavis isolate SUB211 (V7)*	99,84	17,4 ± 2,1	- 67 ± 9	55 ± 23	13 ± 9	85 ± 32	- 7 ± 17	2 ± 0
93	Ferndale		*Viola odorata isolate ODO214 (V8)*	100	41,8 ± 2,4	- 76 ± 11	13 ± 6	34 ± 13			0,7 ± 0,6
94	Zampieri		*Viola odorata voucher NMW6028*	100	22,5 ± 2,2	- 28 ± 5	24 ± 9	19 ± 16			2,7 ± 0,6
95	Viola Colombine				14,6 ± 2,2	- 80 ± 11	20 ± 4	24 ± 30			2 ± 1
96	Cinema bizarre				25,2 ± 1,6	±	18 ± 1	- 16 ± 44			0,7 ± 0,6
97	La Valade		*Viola ovato-oblonga voucher Yoo931*	97,05	38,3 ± 0,3	- 73 ± 27	18 ± 6	16 ± 22			1,7 ± 0,6
98	Governor Herrick		*Viola sp. Hearn cult33*	99,84	21,6 ± 1,3	- 63 ± 20	- 2 ± 3	1 ± 11			0,7 ± 0,6
99	Peace		*Viola suavis isolate SUW170 (V17)*	99,18	28,3 ± 2,3	20 ± 28	16 ± 9	6 ± 20			0,3 ± 0,6
100	Viola Amiral Avellan		*Viola suavis isolate SUB211 (V7)*	99,84	31,5 ± 0,5	45 ± 1	43 ± 11	4 ± 3	7 ± 3	89 ± 3	1,3 ± 0,6
101	Viola Hederacea epertion		*Viola schulzeana*	89,9	33,7 ± 1,2	- 79 ± 8	85 ± 4	98 ± 11	131 ± 1	73 ± 7	3 ± 0
102	Coven garden				36,9 ± 0,7	- 27 ± 11	36 ± 26	58 ± 44	91 ± 50	101 ± 0	1,7 ± 0,6

104	Kerry Girl		*Viola verecunda voucher Yoo975*	100		- 76 ± 6	22 ± 7	25 ± 12			2,7 ± 1,2
105	Odorata alba		*Viola sp. Hearn cult33*	100							
106	Odorata blanche		*Viola sp. Hearn cult33*	99,69	18,8 ± 0,8	- 8 ± 13	100 ± 0	28 ± 2	135 ± 0	95 ± 2	2,3 ± 0,6
107	Viola reine Charlotte		*Viola odorata isolate ODO178 (V11)*	100	44 ± 0,5	- 21 ± 3	24 ± 6	1 ± 18			1 ± 0
108	Orange		*Viola odorata isolate ODO214 (V8)*	100	19,9 ± 1,3	- 40 ± 29	21 ± 4	13 ± 16			0,7 ± 0,6
109	Fée Jalucine		*Viola alba isolate ALB184 (V6)*	98,85	68,8 ± 1	- 37 ± 32	26 ± 9	8 ± 9			0,7 ± 0,6
110	Violette de Toulouse		*Viola alba isolate ALB184 (V6)*	98,85	72,2 ± 1	- 70 ± 4	- 90 ± 14	97 ± 0	- 81 ± 14	- 21 ± 2	
111	Lady Hume Cambell					- 65 ± 9	45 ± 9	79 ± 14	34 ± 33	- 15 ± 12	1,7 ± 0,6

PUBLICATIONS

Metabolomics (2017) 13:96
DOI 10.1007/s11306-017-1227-6

ORIGINAL ARTICLE

Dereplication of natural products from complex extracts by regression analysis nd molecular networking: case study of redox-active compounds from *Viola alba subsp. dehnhardtii*

Justine Chervin[1,2,4] · Pierre Perio[1] · Nathalie Martins-Froment[3] · Chiobouaphong Pharkeovilay1 · Karine Reybier1 · Françoise Nepveu1 · Nicolas Fabre[1] · Thierry Talou[2] · Valérie Bonzon-Ponnet[4] · Guillaume Marti[1]

Received: 5 May 2017 / Accepted: 19 June 2017
© Springer Science+Business Media, LLC 2017

Abstract:

Introduction In natural product research, bioassay-guided fractionation was previously widely employed but is now judged to be inadequate in terms of time and cost, particularly if only known compounds are ultimately isolated. The development of metabolomics, along with improvements in analytical tools, allows comprehensive metabolite profiling. This enables dereplication to target unknown active compounds early in the purification workflow.

Objectives Starting from an ethanolic extract of violet leaves, this study aims to predict redox active compounds within a complex matrix through an untargeted metabolomics approach and correlation analysis.

Method Rapid fractionation of crude extracts was carried out followed by multivariate data analysis (MVA) of liquid chromatography-high resolution mass spectrometry (LC-HRMS) profiles. In parallel, redox active properties were evaluated by the capacity of the molecules to reduce 2,2-diphenyl-1-picrylhydrazyl (DPPH•) and superoxide ($O_2^{\bullet-}$) radicals using UV-visible and electron spin resonance (ESR) spectroscopies, respectively. A spectral similarity network (molecular networking) was used to highlight clusters involved in the observed redox activities.

Results Dereplication on *Viola alba subsp. dehnhardtii* highlighted a reproducible pool of redox active molecules. Polyphenols, particularly *O*-glycosylated coumarins and *C*-glycosylated flavonoids, were identified and *de novo* dereplicated through molecular networking. Confirmatory analyses were undertaken by thin layer chromatography (TLC)-

Electronic supplementary material The online version of this article (doi:10.1007/s11306-017-1227-6) contains supplementary material, which is available to authorized users.

* Guillaume Marti
guillaume.marti@univ-tlse3.fr

Justine Chervin
justine.chervin@ensiacet.fr

Pierre Perio
pierre.perio@univ-tlse3.fr

Nathalie Martins-Froment
martins@chimie.ups-tlse.fr

Chiobouaphong Pharkeovilay
chiobouaphong.phakeovilay@univ-tlse3.fr

Karine Reybier karine.reybier@univ-tlse3.fr
Françoise Nepveu francoise.nepveu@univ-tlse3.fr
francoise.nepveu@univ-tlse3.fr

Nicolas Fabre
nicolas.fabre@univ-tlse3.fr

Thierry Talou
thierry.talou@ensiacet.fr

Valérie Bonzon-Ponnet
valerie.bonzonponnet@parfumsberdoues.com

1 UMR 152 Pharmadev, Université de Toulouse, IRD, UPS, Toulouse, France

2 Laboratoire de Chimie Agro-industrielle, Université de Toulouse, INRA, INPT, Toulouse, France

3 Service commun de spectrométrie de masse, Université de Toulouse, ICT, UPS, Toulouse, France

4 Berdoues SAS, Cugnaux, France

DPPH-MS assays and nuclear magnetic resonance (NMR) spectra of the most active compounds. *Conclusion* Our dereplication strategy allowed the screening of leaf extracts to highlight new biologically active metabolites in few steps with a limited amount of crude material and reduced time-consuming manipulations. This approach could be applied to any kind of natural extract for the study of various biological activities.

Keywords: Dereplication; metabolomics; molecular networking; natural products; redox active properties; UHPLC-HRMS

1. Introduction

In natural product research, identification and characterisation of bioactive compounds from crude extracts remains challenging. To identify active molecules, conventional treatments include extraction, bio-guided fractionation, isolation, and characterisation steps. However, these procedures are time consuming as the composition of natural extracts can be complex. There is also a risk of redundant results, such as isolating known compounds or loss of bioactivity during manipulations (Ayouni et al., 2016). Holistic methods such as metabolomics can enhance the classical reductionist workflow based on iterative bioassays and paralleled fractionation to reach active compounds. This approach, referred to as untargeted metabolite profiling, and subsequent statistical analysis was revealed as a valuable tool in natural product research (Cox, Oh, Keasling, Colson, & Hamann, 2014a).

Ultra-high performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS) has been used to study complex mixtures, taking advantage of sensitive detection and high chromatographic and mass resolutions. This approach consists of the pre-identification of known compounds in a single analytical run, which is of particular interest in the dereplication process (Wolfender, Marti, Thomas, & Bertrand, 2015b). As a preliminary step, chemical compound databases are interrogated based on accurate mass of all detected features (*m/z*-retention time *(*RT) pairs), which produces molecular formula hits. However, several possibilities may be obtained, mainly due to isobaric matches. Therefore, additional filters are needed to rank database hits, such as cross-research with chemotaxonomic data or tandem mass spectrometry (MS/MS) fragmentation patterns. Indeed, development of *in silico* tools to mirror predicted fragmentation behaviour with experimental data limits the number of matches for a given feature (Xiao, Zhou, & Ressom, 2012) (Kind & Fiehn, 2016). Additionally, a mass spectral similarity network (e.g., molecular networking) was developed to help interpret LC-MS/MS spectra, complementing this dereplication strategy (Yang, Sanchez, Rath, Liu, Boudreau, Bruns, Glukhov, Wodtke, R. De Felicio, et al., 2013) (Grapov, Wanichthanarak, & Fiehn, 2015b). This tool organises the data in the form of a network, composed of nodes and edges with formation of clusters based on MS/MS spectral similarity score. Structurally identified nodes can aid structural elucidation of close neighbours, resulting in the propagation of assigned peaks (Allard et al., 2016).

MS/MS-based dereplication or *de novo* identification strategies are necessary to target active compounds early in the purification process as untargeted metabolite profiling methods generate tens to thousands of variables, depending of the analytical workflow (Wolfender, Marti, & Queiroz, 2010). Multivariate data analysis (MVA) is commonly used to exploit the variable space in order to rank the features most involved in sample separation. Orthogonal partial least square (OPLS) regression models enable the prediction of features involved in a given activity for each sample (Worley & Powers, 2013) (Rantalainen, Cloarec, Nicholson, Holmes, & Trygg, 2006).

Violets contain various natural products with diverse biological activities (Muhammad et al., 2012). These include anti-inflammatory, antimicrobial (Witkowska-Banaszczak, Bylka, Matławska, Goślińska, & Muszyński, 2005) and, predominantly, antioxidant activity (Vukics, 2009). Previous work describes over 200 compounds in this genus, mainly flavonoids, terpenoids and phenylpropanoids (Murray et al., 2015). However, only thirty of the 500 species

were studied. Thus, a potential supply of new biologically active compounds remains to be discovered, particularly non-volatile metabolites. Many interesting activities are related to the redox active functions of the molecules. Monitoring the ability to reduce free radicals by a single electron reaction is straightforward and automatable *in vitro* (López-alarcón & Denicola, 2013b). This makes it possible to identify reducing agents (also known as antioxidants), which are useful in liquid or soft medium as preservatives. *In vivo*, the antioxidant character of a compound is not predominantly linked to its ability to scavenge free radicals due to kinetic constraints. The redox property of the molecule, in its reversibility, is involved. It is the oxidised form that paradoxically activates the Nrf2 (NF-E2-related factor 2) signalling pathway and allows the expression of antioxidant enzymes and proteins (Forman & Davies, 2014). Thus evaluating *in vitro* reductants is the first step in identifying new structures capable of having active oxidised forms *in vivo*. The aim of this work was to identify metabolite clusters most probably involved in reduction reactions starting from a complex extract of leaves of violet of Toulouse (*Viola alba* subsp. *Dehnhardtii,* Violaceae*)*.

2. Materials and methods

2.1. Plant material

In spring 2016, seven distinct flowerpots of violet of Toulouse (*Viola alba* subsp. *dehnhardtii*) were collected from the Toulouse Municipal Greenhouses, France. Samples were washed before separation of flowers, aerial parts and roots. To stabilise the vegetable matter, each plant part was lyophilised and ground into powder.

2.2. Leaf extraction

Metabolites were extracted by adding ten volumes of 80% ethanol (EtOH) to the powdered leaf (10.0 ± 0.2 g). The solutions were then sonicated in a bath (Fisher Scientific) at room temperature for 30 min and filtered. This procedure was repeated once with fresh solvent. Each extract was evaporated under reduced pressure (Buchi rotavapor R-114) and 100 mg of the crude extracts were finally suspended in 1 mL of methanol prior to solid phase extraction (SPE) (1 g Sep-Pak® C_{18} cartridge, Waters, Milford, MA, USA). Rapid fractionation was achieved using five aqueous methanolic solutions of increasing organic solvent concentrations (F1: 95/5 H_2O/MeOH; F2: 75/25; F3: 50/50; F4: 25/75; F5: 0/100). The five collected fractions of each extract were dried under reduced pressure. One part was put aside for redox properties analysis and the other part was dissolved to 1 mg/mL for UHPLC-HRMS analysis.

2.3. Redox properties assay

The reductive properties of the crude extracts and their respective fractions were firstly evaluated by the DPPH radical scavenging assay (Nguyen et al., 2013). Each sample was analysed in triplicate and five concentrations were tested in order to graphically determine the half maximal inhibitory concentration (IC_{50}). Trolox (Fluka, purity > 98%) and rutoside (Sigma, purity > 94%) were used as positive controls. Secondly, the capacity of the molecules to reduce the superoxide radical ($O_2^{\bullet-}$) was analysed by electron spin resonance (ESR) (Hubert, Berger, & Dayde, 2008). Rutoside was employed as a positive control. The results were analysed with WinEPR software (v. 2.11b, Bruker): a baseline correction was firstly done before 2nd order integration of the signal.

2.4. UHPLC-HRMS profiling

All extracts were profiled using a UHPLC-DAD-LTQ Orbitrap XL instrument (Ultimate 3000, Thermo Fisher Scientific, Hemel Hempstead, UK). The UV detection was performed by a diode array detector (DAD) from 210 to 400 nm. Mass detection was performed using an electrospray source in positive ionisation (PI) and negative ionisation (NI) modes at 15000 resolving power (full width at half maximum (FWHM) at 400 *m/z*). The mass scanning range was *m/z* 100–1500 Da. The capillary temperature was 300°C and ISpray voltage was fixed at 4.2 kV (positive mode) and 3.0 kV (negative mode). Mass measurement was externally calibrated before starting the experiment. Each full MS scan was followed by data dependant MS/MS on the three most intense peaks using stepped collision-induced dissociation (CID) (35% normalised collision energy, isolation

width 2 Da, activation Q 0.250). The LC–MS system was run in binary gradient mode using a BEH C_{18} Acquity column (100 × 2.1 mm i.d., 1.7 µm, Waters, MA, USA) equipped with a guard column. Mobile phase A (MPA) was 0.1% formic acid (FA) in water and mobile phase B (MPB) was 0.1% FA in acetonitrile. Gradient conditions were: 0 min, 95% MPA; 0.5 min 95% MPA; 12 min, 5% MPA; 15 min, 5% MPA, 15.5 min, 95% MPA; 19 min, 95% MPA. The flow rate was 0.3 mL/min, column temperature 40°C and injection volume was 2 µL.

2.5. Data processing

The UHPLC-HRMS raw data were converted to abf files (REIFYCS Abf Converter) and processed with MS-DIAL version 2.54 (Tsuagawa, 2015) for mass signal extraction between 100 to 1500 Da from 0 to 12.5 min. Respective MS1 and MS2 tolerance were set to 0.01 and 0.4 Da in centroid mode. The optimised detection threshold was set to 2×10^4 for MS1 and 10 for MS2. Adducts and complexes were identified to exclude them from the final peak list. Finally, the peaks were aligned on a quality control (QC) reference file with a retention time tolerance of 0.1 min and a mass tolerance of 0.025 Da. The resulting peak list was then exported to comma-separated value (CSV) format prior to MVA using SIMCA-P+ (version 14.0, Umerics, Umea, Sweden).

2.6. Statistical analysis

CSV files were directly imported into SIMCA-P+ (version 14.0, Umerics, Umea, Sweden). For MVA, all data were log transformed and pareto scaled. The OPLS regression analysis was done with DPPH IC_{50} values as Y input. Coefficient scores were used to rank variables according to their reductive potential. For each model, a leave-one-subject-out cross-validation was performed to assess the model fit. The validity of the discriminant model was verified using permutation tests (Y-scrambling).

2.7. Identification of significant feaures

Molecular formulae of significant features were calculated with MS-FINDER 2.10 (Tsugawa et al., 2016b). Various parameters were used in order to reduce the number of potential candidates, such as the element selection exclusively including C, H, O; mass tolerance fixed to 10 ppm and the isotopic ratio tolerance set to 20%. Only natural product databases focused on plants were selected from Universal Natural Products Database (UNPD), KNApSAc, PlantCyc, Dictionary of Natural Products (DNP, CRC press, v25:2) and CheBI. The results were presented as a list of compounds sorted according to the score value of the match. This value encompassed uncertainty on accurate mass, the isotopic pattern score and the experimental MS/MS fragmentation mirrored to *in silico* matches. Chemical classes were retained for identified features with a score above 5 and only structures with a score above 7 were retained for thorough analysis.

2.8. Mass spectral similarity network

The text file format exported from MS-DIAL was cleaned-up by eliminating the identified adducts before importing into MetamapR (version 1.4.0) (Grapov et al., 2015b). A mass spectral similarity network was created with a maximum of 10 connections between nodes, a cut off fixed at 0.3 and a retention time filter at 1.5 min. The calculated edge list was then downloaded and processed with Cytoscape 2.8.3 (Shannon et al., 2003b). An attribute file containing all processed information, in particular *m/z* values, OPLS coefficients and chemical classes of identified features was imported to improve network visualisation.

2.9. Validation of the model by TLC-DPPH-MS

A TLC method was developed to confirm the statistical results. TLC separation was undertaken for fractions F2 and F3 with respective migration solvent composition: ethyl acetate, formic acid, acetic acid, water (50:5.5:5.5:13.5) and ethyl acetate, formic acid, water (60:6.5:6.5). 675 µL of a 10 mg/mL solution was deposed in band of 180 mm in a 20 cm TLC plate. The plate was then placed in an oven at 70°C for 15 min and then observed under UV at 254 nm for F2 and 366 nm for F3. A small part of the plate was revealed with a purple DPPH solution at 600 mol/L. Then, yellow active spots were collected on the non-treated surface with the use of a TLC-MS interface (Camag, Muttenz, Switzerland) and directly injected into the LTQ-Orbitrap instrument.

2.10. Purification of compounds of interest

Purification of active compounds was undertaken starting from 200 mg of F2 (supplementary material **S1**) on an HPLC-DAD-QTOF-MS instrument (HPLC Alliance 2695- QTof Premier, Waters, MA, USA).

3. Results

3.1. Radical scavenging properties in vitro

Redox properties of extracts were evaluated by studying their ability to scavenge DPPH• (2,2-diphenyl-1-picrylhydrazyl) radical. The analyses were first carried out on ethanolic extracts of flowers, leaves, and roots. The results showed that roots presented weak radical scavenging activity, with an IC_{50} of 1525 ± 122 mg/L. Flowers and leaves had greater activity, with IC_{50} values of 475 ± 15 mg/L and 467 ± 14 mg/L, respectively (see supplementary data **S2**). Leaves were chosen for further investigations as flowers are more fragile and less abundant.

Radical scavenging activity was measured for the crude ethanolic leaf extracts of seven individual plants of violet of Toulouse and respective fractions (N = 42) of decreasing polarity (F1 to F5) (**Fig. 1 A**; plain bar plots). While the seven crude extracts displayed an IC_{50} value of 467 ± 14 mg/L, high IC_{50} values were obtained for SPE fractions F1 and F5 (2605 ± 380 mg/L and 1337 ± 26 mg/L, respectively) demonstrating weak radical scavenging activity *in vitro* in these cases. Significant activities were measured for fractions F2, F3 and F4, with IC_{50} values of 53 ± 5 mg/L, 153 ± 9 mg/L and 429 ± 10 mg/L, respectively. For comparison, the positive controls Trolox and rutoside possess IC_{50} values of approximately 7 mg/L and 14 mg/L, respectively. Therefore, fractions F2 and F3 seemed promising for the discovery of antioxidant compounds *in vitro* as preservatives.

To confirm these results, ESR experiments were conducted to measure the capacity of extracts (final concentration of 50 mg/L) to trap superoxide radical (O_2•⁻). The corresponding values of the ESR signal double integrations are presented in **Fig. 1 A** (striped bar plots). As for DPPH assay, F2 and F3 appeared to have the most redox potential as demonstrated by the low intensity of the remaining signal.

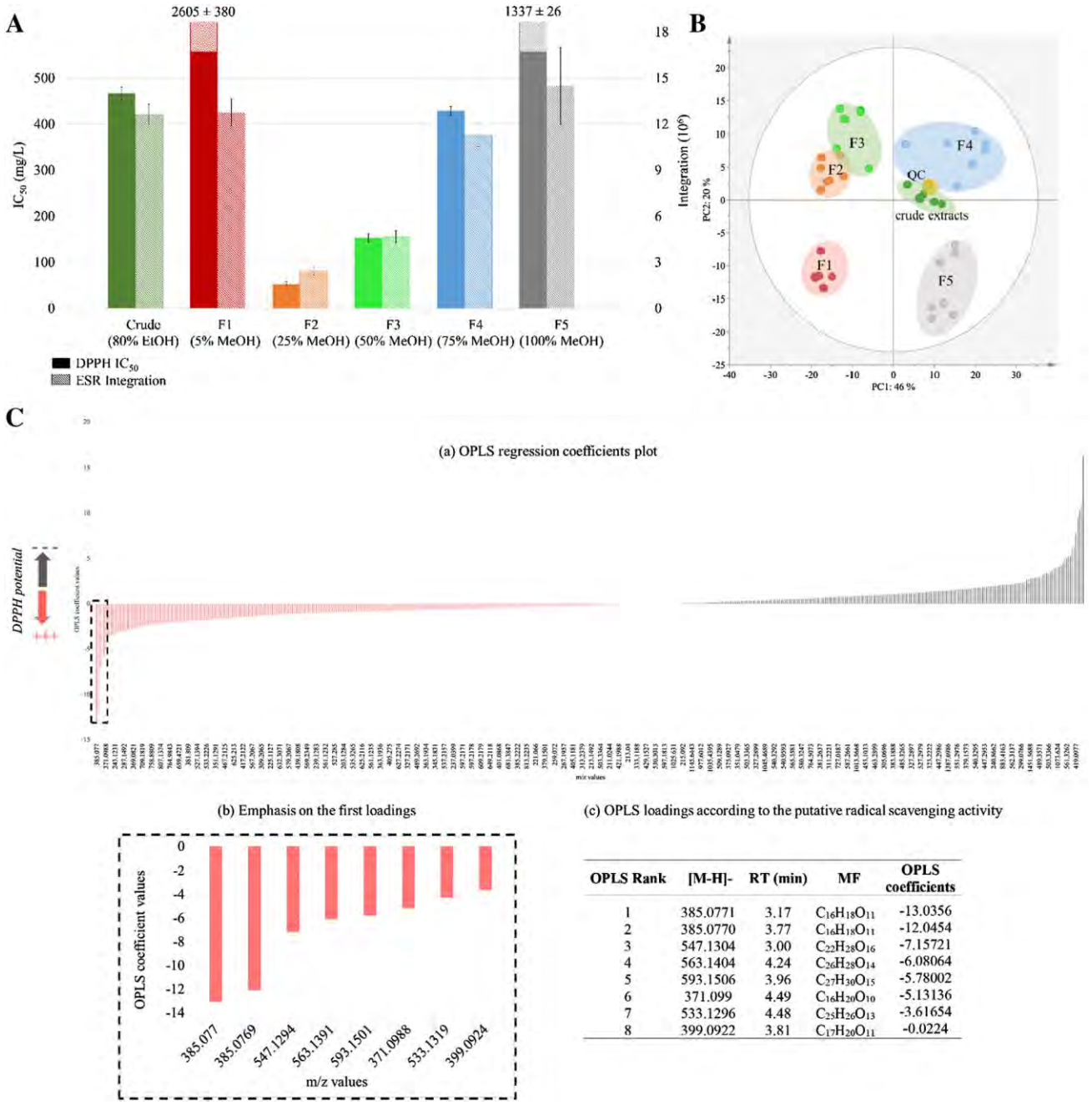
3.2. UHPLC-HRMS-based metabolomics approach

UHPLC-HRMS profiles of all the 46 extracts (7 crude extracts, 35 SPE fractions, and 6 QC samples prepared by pooling an aliquot of all fractions) afforded 434 and 527 features (*m/z*-RT pairs) in NI and PI modes, respectively.

The C18-SPE procedure is clearly effective at resolving the extracts according to polarity, with polar fractions 1 to 3 eluting early (< 6 min in our UHPLC conditions) and fractions 4 and 5, containing apolar compounds, eluting at later retention times (see supplementary data **S3**).

As a preliminary step, principal component analysis (PCA) was applied as an exploratory data analysis to provide an unsupervised overview of the LC-MS fingerprints. PCA clustered all independent biological replicates from the same fraction. Thus, one fraction was related to one position on the plot with slight variability highlighting a reproducible response. As expected, PCA grouped crude extracts and QC near the plot centre. By contrast, single fractions were well distributed (**Fig. 1 B**).These results indicated that SPE fractionation and LC-MS workflows highlighted variability in the data set, and were highly reproducible. Moreover, it demonstrated a stable chemical composition within aerial parts of the seven independent plants under study. After PCA, we applied OPLS regression analysis to obtain a classification of the loadings (i.e. m/z-RT pairs) regarding the DPPH IC50 value (input Y). The quality of model prediction was good (R2Y = 0.977, Q2Y = 0.952) and a permutation test assessed its validity (see supplementary data **S4**). This supervised method allowed the ranking of potential redox-active compounds according to their regression coefficient values (**Fig. 1 C**-a,b): negative coefficients were correlated to potential redox-active compounds and positive coefficients to less-active redox compounds. From the list containing the first hits (**Fig. 1 C** -c), we

carried out the identification procedure based on their MS/MS spectra.



OPLS Rank	[M-H]-	RT (min)	MF	OPLS coefficients
1	385.0771	3.17	$C_{16}H_{18}O_{11}$	-13.0356
2	385.0770	3.77	$C_{16}H_{18}O_{11}$	-12.0454
3	547.1304	3.00	$C_{22}H_{28}O_{16}$	-7.15721
4	563.1404	4.24	$C_{26}H_{28}O_{14}$	-6.08064
5	593.1506	3.96	$C_{27}H_{30}O_{15}$	-5.78002
6	371.099	4.49	$C_{16}H_{20}O_{10}$	-5.13136
7	533.1296	4.48	$C_{25}H_{26}O_{13}$	-3.61654
8	399.0922	3.81	$C_{17}H_{26}O_{11}$	-0.0224

Fig. 1 MVA workflow: (A) Comparison of radical scavenging properties of leaf crude extracts and respective fractions of violet of Toulouse obtained by C_{18}-SPE using DPPH assays (plain bar plots) and ESR (striped bar plots). (B) PCA score plot of the ESI-NI dataset (QC: Quality controls, F1 to F5 denote respective C_{18}-SPE fractions). (C-a) Coefficient plot obtained by OPLS regression; (C-b) Emphasis of the first loadings; (C-c) Ranking of the first seven loadings estimated as radical scavenger compounds according to the OPLS regression.

3.3. Identification of potential radical scavenger compounds based on in silico fragmentation

Using the OPLS regression analysis results, the top eight compounds were tentatively identified by interrogating local natural product databases integrated in MS-FINDER: PlantCyc, UNPD, KNApSAcK and ChEBI (**Table 1**, Hit NP databases).

For each compound, the results afforded several candidates and ranked them according to their

similarity score, which was based on comparison between experimental MS/MS fragments and *in silico* spectra of candidates. Interestingly, for compounds **1**, **2**, **3, 6** and **8**, the top three candidates ranked by MS-FINDER each contained a coumarin core, and for compounds **4**, **5** and **7**, each contained a flavonoid derivative. UV spectra of each peak (**Table 1**; UV λ_{max}) supported our assumptions. The coumarin nucleus displays two absorption bands near 270 nm and 310 nm. Substitutions on this nucleus tend to produce a bathochromic shift (Masrani, Rama, & Bafna, 1974). For the flavonoids, two absorption bands of interest are around 250 to 295 nm and 310 to 370 nm, depending on the flavonoid class and the substitution pattern

(Olsen et al. 2009). Therefore, the higher wavelength band gives differentiation between coumarins and flavonoids. Still, the candidates were numerous and not always relevant, e.g., compound **5** had hundreds of candidates and feruloyl derivatives were proposed for compound **1**. To refine the results, we carried out a second identification step by creating a user-defined database restricting results of UNPD to compounds found in *Viola* genus (**Table 1**, Hit UNPD-Viola). Thus, fewer and more pertinent candidates could be proposed. In our case, this led to the identification of compounds **4**, **5** and **7** as flavonoids, indicating that all the other significant features have never been described in *Viola* genus. Moreover, compounds **3** and **8** remained totally unknown as no hit matched with any database.

3.4. Extending the dereplication process via a mass spectral similarity network

The mass spectral similarity network displayed two clusters of radical scavenger compounds, i.e., antioxidants (**Fig. 2**). Each cluster was closely related to one SPE fraction: according to our previous identification results, flavonoids were mainly found in F3 while coumarins were mainly found in F2.

The flavonoid cluster displayed three compounds with related structures active in reducing DPPH• (**Fig. 3**). Compounds **4** and **5** presented a statistically significant radical scavenging potential according to their coefficient score (**Table 1**; OPLS coefficient). Interestingly, the molecular formulae of the compounds from this cluster matched structures already identified in Violaceae. Experimental MS/MS patterns mirrored to *in silico* fragmentation allowed the preliminary identification of schaftoside (**4**), isovitexin 6"-O-β-D-glucopyranoside (**5**) and 6,8-di-C-α-L-arabinopyranosylapigenin (**7**) (see supplementary data **S5**). These structures were based on a user-defined *Viola* database, as hits were more pertinent and the final candidate had the highest total score, generally above 7.

The other compounds of interest belonged to the coumarin cluster (**Fig. 3**). Compounds **3** and **8** were connected with the identified glycosylated coumarins $C_{16}H_{18}O_{11}$ (compound **2**), confirmed by 1D and 2D NMR (see supplementary data **S6**). Thus, we applied a *de novo* dereplication based on the interpretation of mass loss to propose potential chemical structures.

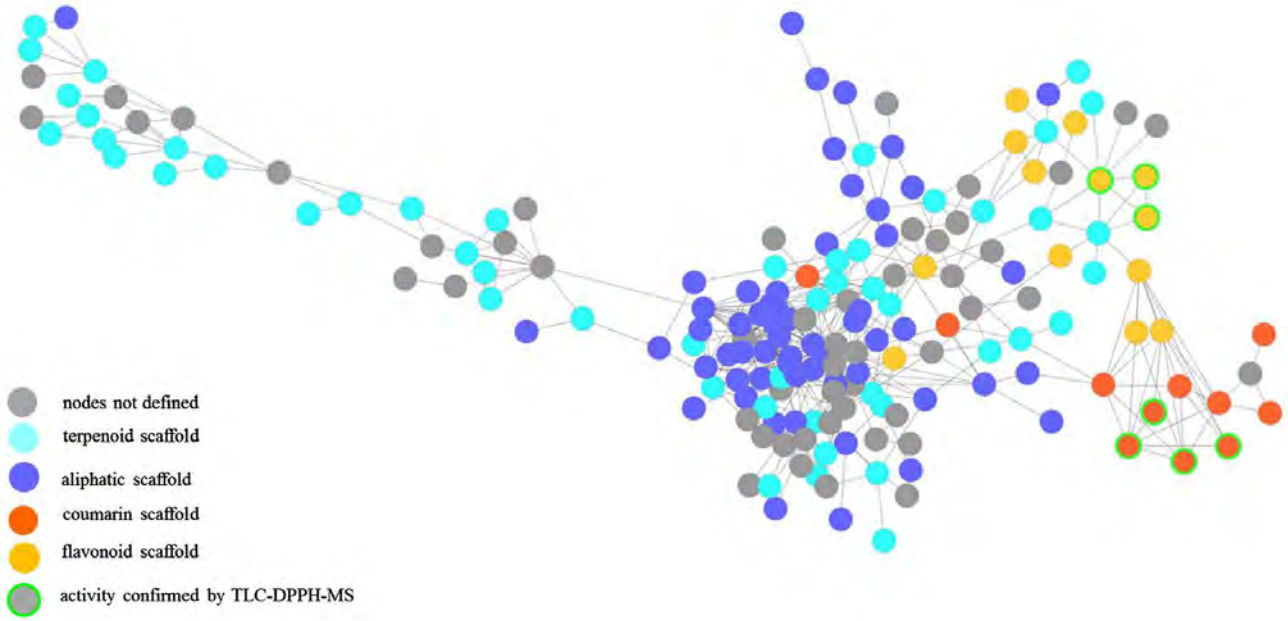


Fig. 2 Mass spectral similarity network of *Viola* ethanolic extracts with coloration based on the chemical class

Table 1 Summary of all dereplicated compounds

Rank*	RT (min)	[M-H]-	Molecular Formula	Error (mDa)	Hit NP databases**	Hit UNPD *Viola***	Main MS/MS fragments	UV λ_{max} (nm)	Potential Id[a,b,c]	MSI level	Chemical class	Biological source	OPLS coefficient	TLC-DPPH assay	Ref
1	3.17	385.0771	$C_{16}H_{18}O_{11}$	0.4349	2	0	206.9150 222.0869 370.1008	260 326	7-Methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside[a,b]	1	Coumarin	*Tetraphis pellucida*	-13.03	Detected	(Jung et al. 1995)
2	3.77	385.0770	$C_{16}H_{18}O_{11}$	0.7349	2	0	207.9790 223.1457 370.1048	258 325	7-Methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[a,b]	1	Coumarin	*Tetraphis pellucida*	-12.04	Detected	(Jung et al. 1995)
3	3.00	547.1304	$C_{22}H_{28}O_{16}$	1.1584	0	0	208.0135 223.0087 385.1274	268 320	7,8-dimethoxy-5,6,8-trihydroxycoumarin-8-β glucopyranosyl-glucopyranoside[c]	2	Coumarin	-	-7.15	Detected	-
4	4.24	563.1404	$C_{26}H_{28}O_{14}$	0.5291	88	3	353.0987 383.2114 443.1452 473.0864	271 333	schaftoside[b]	2	Flavonoid	numerous plants	-6.08	Detected	(Chen Xie et al. 2003)
5	3.96	593.1501	$C_{27}H_{30}O_{15}$	1.2938	154	2	353.0222 383.19O3 473.0770 503.1120	272 332	isovitexin 6"-*O*-β-D glucopyranoside[b]	2	Flavonoid	*Viola tricolor*	-5.78	Detected	(Vukics, Ringer, et al. 2008)
6	4.49	371.099	$C_{16}H_{20}O_{10}$	0.5704	11	8	120.976 231.124 248.972	265 320	8-*C*-Glucosyl-3,4-dihydro-5,7-dihydroxy-6-methoxycoumarin[b]	2	Coumarin	*Diceratella elliptica*	-5.13	Not detected	(Marzouk et al. 2012)
7	4.48	533.1296	$C_{25}H_{26}O_{13}$	0.8644	25	1	353.1044 383.2066 443.1201 473.1409	271 331	6,8-Di-*C*-α-L-arabinopyranosylapigenin[b]	2	Flavonoid	*Viola yedoensis*	-3.61	Detected	(Chen Xie et al. 2003)
8	3.81	399.0922	$C_{17}H_{20}O_{11}$	0.2850	3	0	207.0253 221.9839 237.0366	270 320	7,8-Dimethoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[a,c]	1	Coumarin	-	-0.02	Detected	-

*Ranking based on OPLS regression coefficients, served as compound number

**Hits found when interrogating natural product databases (UNPD, KnapSack, CheBI, and PlantCyc) in MS-FINDER

***Hits found when interrogating the user-defined database (UNPD and DNP) composed of known compounds in *Viola* genus

[a] Confirmed by NMR

[b] Top ranked hit determined by *in silico* MS/MS fragmentation with MS-FINDER

[c] Dereplicated with molecular networking

We interpreted a difference of 162 Da between **3** and **2** as a glucose substituent. We suggest a new 8-*O*-diglycosylated coumarin structure for **3**, based on the MS/MS spectrum and the loss of a 324.12 Da, correlated to two linked glucose units (see supplementary data **S5**). Regarding **8**, a difference of 14 Da was indicative of a di-methoxy-8-*O*-glycosylated coumarin.

Compound **6** was tentatively identified as an 8-*C*-glycolsylated coumarin from comparison with *in silico* fragmentation data using MS-FINDER local databases.

3.5. Validation of the model

To validate our OPLS model, we used a TLC-DPPH-MS assay for fractions F2 (**Fig. 4 a**) and F3 (**Fig. 4 b**). Overall, the four coumarins (**1**, **2**, **3** and **8)** and the three flavonoids (**4**, **5** and **7**) were identified and matched with the OPLS ranking.

We then tested purified compounds **1** and **2** for their respective scavenging activities against DPPH (plain bar plots) and superoxide radicals (striped bar plots) (**Fig. 5**). Trolox and rutoside were used as references for both assays. The difference in radical scavenging capacity of Trolox against DPPH and superoxide radicals can be easily explained by considering the rate constants of the scavenging reactions : $k_{Trolox/DPPH} = 1.1\times10^4\ M^{-1}s^{-1}$ (Friaa & Brault, 2006) and $k_{Trolox/O2\bullet-} = 0.1\ M^{-1}s^{-1}$ (Cabelli & Bielski, 1986).

The radical scavenging properties of compound **1** (97% pure) were confirmed by comparison with the active controls, Trolox (DPPH IC_{50}: **1** 9 ± 2 mg/L vs. Trolox 7 mg/L), and rutoside (ESR double integration: **1** 0.88 ± 0.01 vs, rutoside 1.78± 0.10).
For compound **2** (90% pure), the DPPH IC_{50} of 49 ± 1 mg/L and ESR double integration of 15 ± 1, suggests this compound is a less active radical scavenger than compound **1**, meeting the statistical ranking.

4. DISCUSSION

The objective of this work was to characterise redox active metabolites from the leaves of *Viola alba* subsp. *dehnhardtii* together with *de novo* dereplication of active compounds and the establishment of a non-volatile metabolite fingerprint.

In order to decipher the redox potential in few steps, we correlated DPPH results to UHPLC-HRMS fingerprints by OPLS regression. This method allowed the ranking of detected feature according to their potential radical scavenging properties. In addition, a mass spectral similarity network allowed *de novo* dereplication of top ranked features, based on the acquisition of UHPLC-HRMS profiles in data dependent analysis mode. This provided accurate mass-to-charge ratio for molecular formula determination along with MS/MS fragments used for peak assignments. NI mode MS and MS/MS spectra were mainly used as they were of better quality than PI spectra, thus improving *in silico* identification using MS-FINDER. Processing with MS-DIAL allowed the acquisition of a clean peak list with deconvoluted MS/MS data for each feature. This meant we could easily remove adducts and other ionisation artefacts from files uploaded to MetamapR. This resulted in a cleaner molecular network compared with uploading raw data files. The mass spectral similarity network allowed the organisation of features in clusters and dereplication results highlighted mainly coumarins in SPE fraction F2 and mainly flavonoids in fraction F3. These scaffolds are known for having several biological activities, including redox activity (Procházkoá et al., 2011) (Gacche & Jadhav, 2012). We identified a mixture of *C*- glycosylated flavonoids along with *O*-glycosylated coumarins. The enrichment of these molecules in the two fractions explains the improved antioxidant properties of the fractions over the crude extracts. Although some of the compounds, e.g., compounds **5** and **7,** are well-known in *Viola* genus (Vukics, Kery, Bonn, & Guttman, 2008) (C. Xie, Veitch, Houghton, & Simmonds, 2003), the other compounds we identified have not previously been described in this genus.
The protective effect of phenolic coumarins against oxidative damage depends on the hydrogen-donating capacity of hydroxyl groups (Kostova, 2006) (Borges, Roleira, Milhazes, Santana, & Uriarte, 2005) and then on their oxidisability. Previous structure-activity relationship studies have identified the importance of the number and location of the phenolic hydroxyl groups (Parshad et al., 2016). As compounds **1** and **2**

have the same substituents, energy minimisation calculations were carried out using the

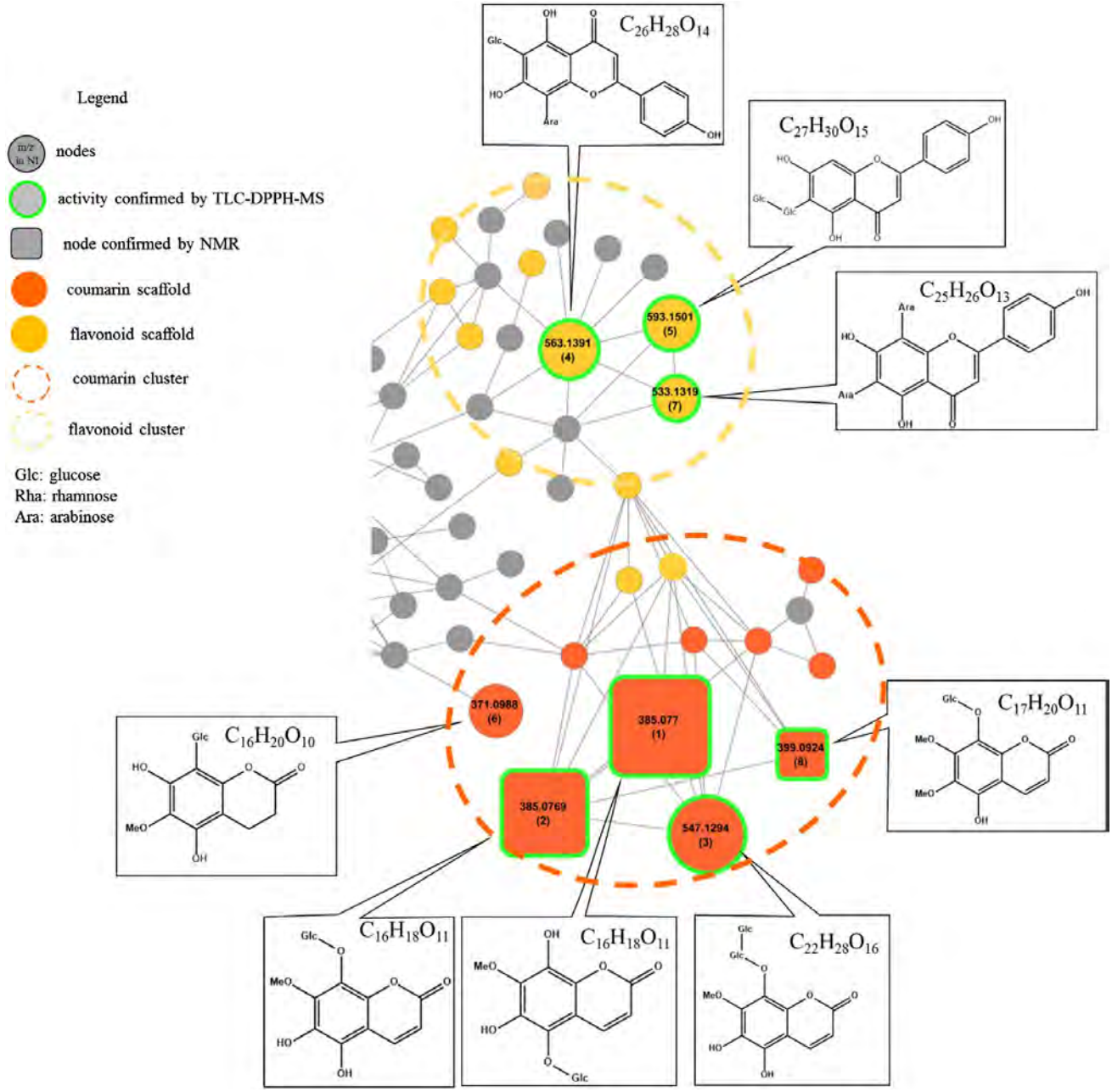


Fig. 3 Expansion on the redox-active compounds clusters. Putative coumarin and flavonoids scaffold correspond to top ranked hits from MS-FINDER. Node size was emphased based on OPLS coefficient value. Compound number is indicated in brackets in each node in correspondence to Table 1

MM2 forcefield (Chemdraw 3D, Cambridge Soft, USA) to understand the difference between their radical reducing capacity. These calculations demonstrate that the hydroxyl group in position 6 in compound **1** was the most readily oxidisable into $OH^{\bullet+}$ and thus responsible for the strong antioxidant capacity measured for **1** compared with **2**. Purification of coumarins $C_{16}H_{18}O_{11}$ **1** and **2** tested in DPPH and ESR redox assays confirmed the workflow used in this study: as predicted by the model, compound **1** appeared as a reducer that is a good antioxidant compound *in vitro*.

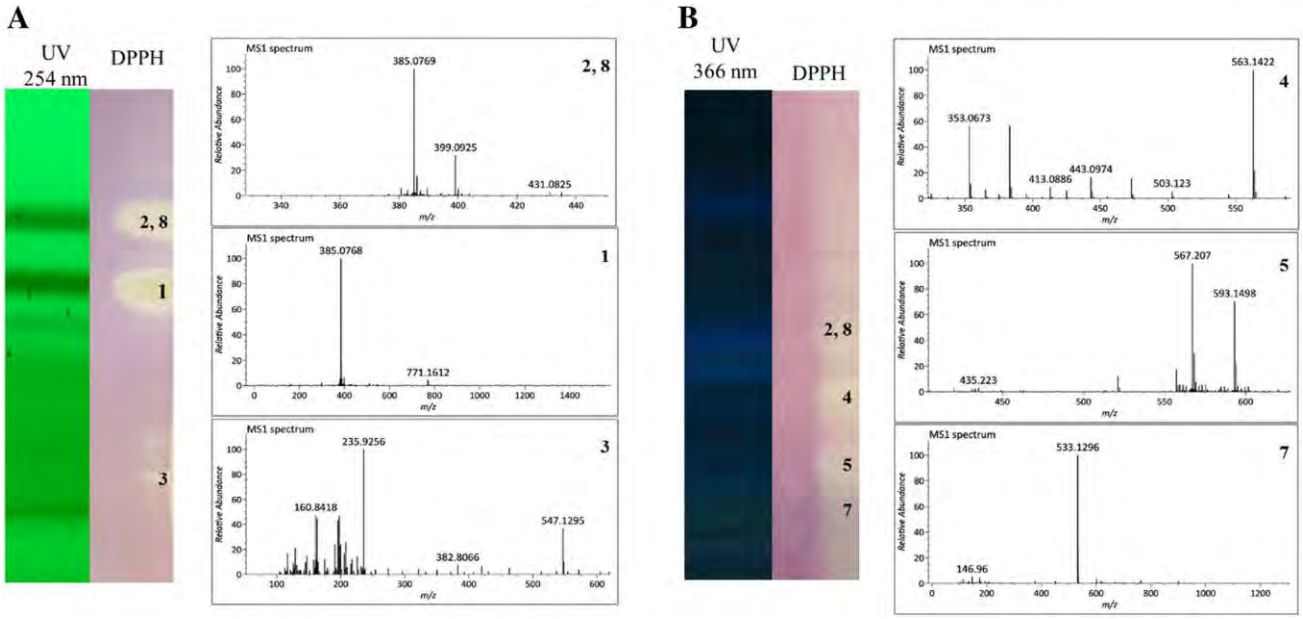


Fig. 4 DPPH TLC–MS dereplication of redox active compounds of F2 (**a**) and F3 (**b**)

Our statistical analysis was validated by TLC-DPPH-MS assays, with detection of seven DPPH-active compounds ranked in top positions of the analysis. A big advantage of our approach is the limited manipulation steps (crude extraction followed by a single SPE fractionation) and low sample demand. For instance, 1 g of dried plant material should be sufficient to conduct the entire workflow. The workflow could easily be implemented to any kind of extract with other biological activities. However, ranking results of variables may not accurately reflect the biological activity. Ranked positions may vary depending on scaling and normalisation processes. The first twenty ranked variables should be considered for thorough analysis. Another drawback of this method concerns the correlation between biological assay results (Y) and the detection method used to generate variables (X). In our case, we only took into account ionisable compounds in electrospray ionisation (ESI) mode for OPLS modelling. The ionisation process is compound dependent and does not reflect a quantitative measure of each detected feature. We could reach more realistic models by using universal-type detectors such as ELSD (Evaporating Light Scattering Detector) or CAD (Charged Aerosol Detector). While ^{1}H- or ^{13}C-NMR could be an interesting alternative, the lack of sensitivity of this method is not compatible with complex samples like crude natural extracts.

Finally, the framework of our dereplication strategy capitalised on UHPLC-HRMS-MS/MS profiles, using agreement between modelled and experimental MS/MS fragmentation patterns and mass spectral similarity network to propagate assignments of clustered features. As illustrated with the coumarin cluster, using the molecular network and propagating from a known compound allows identification of others belonging to the same class by interpreting the neutral mass loss.

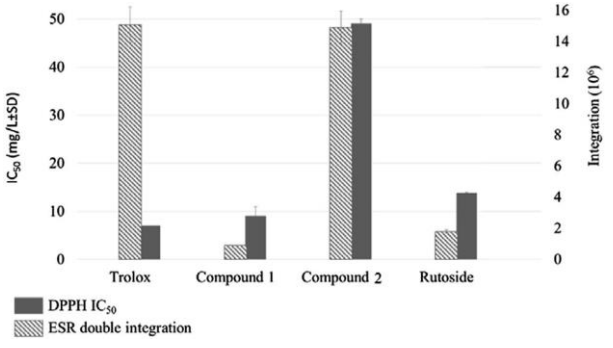


Fig. 5 Capacity of purified compounds **1** and **2** to scavenge free radicals

5. Conclusion

A UHPLC-HRMS based metabolomic study combined with molecular networking tools allowed the dereplication of redox-active metabolites present in the violet of Toulouse. These compounds were mainly flavonoid and coumarin derivatives. Our approach highlighted two main metabolite clusters displaying functional groups with redox-active properties. We identified seven compounds

of interest, of which five were found through MS/MS fragmentation and comparison with references found in databases. Two unknown coumarins were *de novo* dereplicated through a molecular networking approach. Overall, the workflow proposed here allowed early identification of redox active compounds within a complex mixture with limited effort and crude materials.

In this work, it was chosen to correlate redox-active properties with LC-MS profiles to identify compounds of interest but more generally, provided variability is brought to the data set, this method can be implemented in any study of other chemical and biological properties employing different types of assays.

Acknowledgments The authors would like to thank the Regional Council Midi-Pyrénées for sponsoring this project (Project CLE 13053062) as well as the Toulouse Municipal Greenhouses for making the violet collection available to its collaborators and Iroshi Tsugawa for our fruitful correspondence.

Funding: This study was funded by the Regional Council Midi-Pyrénées (Project CLE 13053062).

Author Contributions G. Marti proposed the study and designed the experiments. J. Chervin performed the extraction of crude material, the radical scavenging assays under the supervision of P. Perio, processed the LC-HRMS profiles acquired by N. Martins-Froment and performed MVA and molecular networking analysis under the supervision of G. Marti. J. Chervin analysed the results and wrote the manuscript which was reviewed by all the collaborators of the work.

Compliance with ethical standards

Research involving human participants and/or animals: No human participants or animals were involved in this study.

Informed consent: No human participants were involved in this study

Conflict of interest: The authors declare that they have no conflict of interest.

References

Aceto, S., Caputo, P., Cozzolino, S., Gaudio, L., & Moretti, A. 1999. Phylogeny and evolution of Orchis and allied genera based on ITS DNA variation: morphological gaps and molecular continuity. ***Molecular Phylogenetics and Evolution***, 13(1): 67–76.

Ahmad, H., Khan, S. M., Ghafoor, S., & Ali, N. 2009. Ethnobotanical Study of Upper Siran. ***Journal of Herbs Spices & Medicinal Plants***, 15: 86–97.

Allard, P. M., Péresse, T., Bisson, J., Gindro, K., Marcourt, L., et al. 2016. Integration of Molecular Networking and In-Silico MS/MS Fragmentation for Natural Products Dereplication. ***Analytical Chemistry***, 88(6): 3317–3323.

Álvarez, I. 2003. Ribosomal ITS sequences and plant phylogenetic inference. ***Molecular Phylogenetics and Evolution***, 29(3): 417–434.

Anizon, J.-Y., Lamaire, B., & Surbled, M. 2003. Extraction assistée par micro-ondes. ***Techniques de l'Ingénieur***, 1–10.

Ansari, M., & Emami, S. 2016. β-Ionone and its analogs as promising anticancer agents. ***European Journal of Medicinal Chemistry***, 123: 141–154.

Arnold, M. L. 2006. ***Evolution through genetic exchange***. Oxford ; New York: Oxford University Press.

Ayouni, K., Berboucha-rahmani, M., Kim, H. K., Atmani, D., Verpoorte, R., et al. 2016. angustifolia leaf and stem bark extracts. ***Industrial Crops & Products***, 1–13.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., et al. 1995. The ITS Region of Nuclear Ribosomal DNA: A Valuable Source of Evidence on Angiosperm Phylogeny. ***Annals of the Missouri Botanical Garden***, 82(2): 247.

Ballard, H. E., Sytsma, K. J., & Kowal, R. R. 1999. Shrinking the violets: phylogenetic relationships of infrageneric groups in Viola (Violaceae) based on internal transcribed spacer DNA sequences. ***Systematic Botany***, 23(4): 439–458.

Barbeta, B. L., Marshall, A. T., Gillon, A. D., Craik, D. J., & Anderson, M. A. 2008. Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. ***Proceedings of the National Academy of Sciences***, 105(4): 1221–1225.

Barry, D. G., Daly, N. L., Clark, R. J., Sando, L., & Craik, D. J. 2003. Linearization of a Naturally Occurring Circular Protein Maintains Structure but Eliminates Hemolytic Activity †. ***Biochemistry***, 42(22): 6688–6695.

Bartle, K. D., & Myers, P. 2002. History of gas chromatography. ***Trends in Anlytical Chrmistry***, 21(02): 547–557.

Bedoukian, P. Z. 1982. Violet Fragrance Compounds. ***Fragrance Chemistry: The Science of the Sense of Smell*** (Academic Press Inc.): 285–315. London: Ernst T. Theimer.

Beille, L. 1909. Famille des Violacées. ***Botanique pharmaceutique***, vol. II: 411–415. A. Maloine.

Beirnaert, C., Meysman, P., Vu, T. N., Hermans, N., Apers, S., et al. 2018. speaq 2.0: A complete workflow for high-throughput 1D NMR spectra processing and quantification. (D. Schneidman, Ed.)***PLOS Computational Biology***, 14(3): e1006018.

Benaissi, K. 2013. Le CO2 supercritique appliqué à l'extraction végétale. ***Techniques de l'Ingénieur***, 1–16.

Benjamini, Y., & Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. ***Journal of the Royal Statistical Society. Series B (Methodological)***, 289–300.

Bhatt, V. P., & Negi, G. C. S. 2006. Ethnomedicinal plant resources of Jaunsari tribe of Garhwal Himalaya, Uttaranchal. ***Indian Journal of Traditional Knowledge***, 5(3): 331–335.

Blomberg, J., Schoenmakers, P. J., Beens, J., & Tijssen, R. 1997. Compehensive two-dimensional gas chromatography (GC$\times$ GC) and its applicability to the characterization of complex (petrochemical) mixtures. ***Journal of High Resolution Chromatography***, 20(10): 539–544.

Bonnier, G., & Douin, R. 1990. ***La Grande Flore en couleurs***, vol. 3. Belin.

Borges, F., Roleira, F., Milhazes, N., Santana, L., & Uriarte, E. 2005. Simple Coumarins and Analogues in Medicinal Chemistry : Occurrence , Synthesis and Biological Activity. ***Current Medicinal Chemistry***, 12: 887–916.

Burgot, G., & Pellerin, F. 2003. Microextraction en phase solide (SPME). ***Techniques de l'ingénieur***, 33(0): 0–20.

Burketova, L., Trda, L., Ott, P. G., & Valentova, O. 2015. Bio-based resistance inducers for sustainable plant protection against pathogens. ***Biotechnology Advances***, 33(6): 994–1004.

Cabelli, D. E., & Bielski, B. H. J. 1986. STUDIES OF THE REACTIVITY OF TROLOX WITH Mn3+/Fe 3÷ COMPLEXES BY PULSE RADIOLYSIS. ***Journal of Free Radicals in Biology & Medicine***, 2: 71–75.

Cai Ke, J. 2007. Violaceae. ***Flora of China*** (Science Press (Beijing) & Missouri Botanical Garden (St. Louis)), vol. 13: 72–111. Wu Zhengyi, Peter H. Raven & Hong Deyuan.

Calderón-Santiago, M., Fernández-Peralbo, M. A., Priego-Capote, F., & Luque de Castro, M. D. 2016. MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes. ***Metabolomics***, 12(3): 1–12.

Cao, J., Yin, C., Qin, Y., & Cheng, Z. 2013. Chemical Constituents of Viola yedoensis and Their Antioxidant Activity. ***Chinese Journal of Experimental Traditional Medical Formulae***, 19: 77–81.

Caude, M., & Jardy, A. 2016. ***Chromatographie en phase liquide - Théorie et méthodes de séparation Chromatographie en phase liquide Théorie et méthodes de séparation***, 33(0).

Cawoy, H., Mariutto, M., Henry, G., Fisher, C., Vasilyeva, N., et al. 2014. Plant Defense Stimulation by Natural Isolates of *Bacillus* Depends on Efficient Surfactin Production. ***Molecular Plant-Microbe Interactions***, 27(2): 87–100.

Cennamo, P., Guacchio, E. D., Jury, S. L., & Caputo, P. 2011. Molecular markers in Viola L. subsect. Viola : Application and taxonomic implications for the identification of dubious herbarium specimens. ***Plant Biosystems - An International Journal Dealing with All Aspects of Plant Biology***, 145(2): 306–323.

Chauchard, B., Munzinger, J., Henry, M., & Cedex, N. 2013. L'étude caryologique de deux cultivars de Violettes odorantes remet en cause leur origine taxonomique. ***Acta Botanica Gallica***, 150(March 2014): 85–93.

Chen, B., Colgrave, M. L., Daly, N. L., Rosengren, K. J., Gustafson, K. R., et al. 2005. Isolation and Characterization of Novel Cyclotides from

Viola hederaceae: Solution structure and anti-HIV activity of vhl-1, a leak-specific expressed cyclotide. ***Journal of Biological Chemistry***, 280(23): 22395–22405.

Chervin, J., Perio, P., Martins-Froment, N., Pharkeovilay, C., Reybier, K., et al. 2017. Dereplication of natural products from complex extracts by regression analysis and molecular networking: case study of redox-active compounds from Viola alba subsp. dehnhardtii. ***Metabolomics***, 13(8). https://doi.org/10.1007/s11306-017-1227-6.

Clausen, J. 1927. Chromosome number and the relationship of species in the genus Viola. ***Annals of Botany***, 41(164): 677–714.

Conesa, M. à., Mus, M., & Rosselló, J. A. 2008. Hybridization between insular endemic and widespread species of Viola in non-disturbed environments assessed by nuclear ribosomal and cpDNA sequences. ***Plant Systematics and Evolution***, 273(3–4): 169–177.

Costa, R., Fanali, C., Pennazza, G., Tedone, L., Dugo, L., et al. 2015. Screening of volatile compounds composition of white truffle during storage by GCxGC-(FID/MS) and gas sensor array analyses. ***LWT - Food Science and Technology***, 60(2): 905–913.

Cotton, J. 2015. ***Application de l'analyse métabolomique à la détection ciblée et globale de contaminants organiques dans des matrices agroalimentaires et environnementales par spectrométrie de masse à ultra-haute résolution***. Université Pierre et Marie Curie Paris 6.

Cox, D. G., Oh, J., Keasling, A., Colson, K. L., & Hamann, M. T. 2014a. The Utility of Metabolomics in Natural Product and Biomarker Characterization. ***Biochimica et Biophysica Acta***, 1840(12): 3460–3474.

Cox, D. G., Oh, J., Keasling, A., Colson, K. L., & Hamann, M. T. 2014b. The utility of metabolomics in natural product and biomarker characterization. ***Biochimica et Biophysica Acta***, 1840(12): 3460–3474.

Craik, D. J. 2012. Host-Defense Activities of Cyclotides. ***Toxins***, 4: 139–156.

Cu, J. Q., Perineau, F., & Gaset, A. 1992. Volatile components of violet leaves. ***Phytochemistry***, 31(2): 571–573.

de Hoffmann, E., & Stroobant, V. 2005. ***Spectrométrie de masse*** (3e édition). Dunod.

De Lamarck, J.-B. P. A. de M., & De Candolle, A. P. 1805. ***Flore française, ou description succinctes de toutes les plantes qui croissent naturellement en France***, vol. 4. Paris: Desray.

Dong, M. W., & Zhang, K. 2014. Ultra-high-pressure liquid chromatography (UHPLC) in method development. ***TrAC Trends in Analytical Chemistry***, 63: 21–30.

Dudareva, N., Klempien, A., Muhlemann, J. K., & Kaplan, I. 2013. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. ***New Phytologist***, 198: 16–32.

Dunn, W. B., & Ellis, D. I. 2005. Metabolomics: Current analytical platforms and methodologies. ***Trends in Analytical Chemistry***, 24(4): 285–294.

Erben, M. 1996. The significance of hybridization on the forming of species in the genus Viola. ***Bocconea***, 5(1): 113–118.

Eriksson, L., Trygg, J., & Wold, S. 2008. CV-ANOVA for significance testing of PLS and OPLS® models. ***Journal of Chemometrics***, 22(11–12): 594–600.

Favier, A. 2003. ***Le stress oxydant***. L'actualité chimique.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. ***Evolution***, 39(4): 783–791.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. 1990. Electropsray ionization-principles and practice. ***Mass Spectrometry Reviews***, 9: 32–70.

Fernandez, X., Breme, K., & Varlet, V. 2009. Couplage chromatographie en phase gazeuse / olfactométrie. ***Technique de l'Ingénieur***, 33(Base documentaire : Gestion des odeurs et des nuisances olfactives): 26.

Fernandez, X., Cabrol-bass, D., Fernandez, X., & Cabrol-bass, D. 2007. Analyse des arômes. ***Technique de l'Ingénieur***, 33(0): 1–22.

Ferrari, G., Donsı, F., & Pataro, G. 2010. Applications of Pulsed Electric Field Treatments for the Enhancement of Mass Transfer from Vegetable Tissue. ***Food Eng Rev (2010)***, 2: 109–130.

Ferreira, S. 2004. Doehlert matrix: a chemometric tool for analytical chemistry - review. ***Talanta***, 63(4): 1061–1067.

Finkelstein, E., Rosen, G. M., & Rauckman, E. J. 1980. Spin trapping of superoxide and hydroxyl radical: practical aspects. ***Archives of Biochemistry and Biophysics***, 200(1): 1–16.

Flamini, G. 2007. Flavonoids and other compounds from the aerial parts of Viola etrusca. ***Chemistry & Biodiversity***, 4(2): 139–44.

Flavonoids as antioxidants: determination of radical-scavenging efficiencies. n.d. .

Forman, H. J., & Davies, K. J. A. 2014. How Do Nutritional Antioxidants Really Work: Nucleophilic Tone and Para-Hormesis Versus Free Radical Scavenging in vivo. ***Free Radical Biology and Medicine***, 66: 24–35.

Friaa, O., & Brault, D. 2006. Kinetics of the reaction between the antioxidant Trolox and the free radical DPPH • in semi-aqueous solution. ***Organic & Biomolecular Chemistry***, 4: 2417–2423.

Gacche, R. N., & Jadhav, S. G. 2012. Antioxidant Activities and Cytotoxicity of Selected Coumarin Derivatives: Preliminary Results of a Structure-Activity Relationship Study Using Computational Tools. ***Journal of Experimental and Clinical Medicine***, 4(3): 165–169.

Gerlach, S. L., Rathinakumar, R., Chakravarty, G., Göransson, U., Wimley, W. C., et al. 2010. Anticancer and chemosensitizing abilities of cycloviolacin O2 from Viola odorata and psyle cyclotides from Psychotria leptothyrsa. ***Biopolymers***, 94(5): 617–625.

Goff, S. A. 2006. Plant Volatile Compounds: Sensory Cues for Health and Nutritional Value? ***Science***, 311(5762): 815–819.

Goudable, J., & Favier, A. 1997. Radicaux libres oxyénés et antioxydants. ***Nutur Clin Métabol***, 11: 115–120.

Gran, L., Sandberg, F., & Sletten, K. 2000. Oldenlandia affinis (R&S) DC A plant containing uteroactive peptides used in African traditional medicine. ***Journal of Ethnopharmacology***, 70: 197–203.

Grapov, D., Wanichthanarak, K., & Fiehn, O. 2015a. MetaMapR: pathway independent metabolomic network analysis incorporating unknowns. ***Bioinformatics***, 31(16): 2757–2760.

Grapov, D., Wanichthanarak, K., & Fiehn, O. 2015b. MetaMapR : pathway independent metabolomic network analysis incorporating unknowns. ***Bioinformatics***, 31(16): 2757–2760.

Grasse, M.-C., Ellena, J.-C., & Peyron, L. 2009a. Survol historique de la culture des Violettes et de leur valorisation à Tourettes-sur-Loup. ***Un jour, une plante : Les violettes en parfumerie***.

Grasse, M.-C., Ellena, J.-C., & Peyron, L. 2009b. Utilisation des feuilles de violette à Grasse. ***Un jour, une plante : Les violettes en parfumerie***.

Gruber, C. W., Čemažar, M., Anderson, M. A., & Craik, D. J. 2007. Insecticidal plant cyclotides and related cystine knot toxins. ***Toxicon***, 49(4): 561–575.

Gustafson, K. R., McKee, T. C., & Bokesch, H. R. 2004. Anti-HIV cyclotides. ***Current Protein and Peptide Science***, 5: 331–340.

Haleng, J., Pincemail, J., Defraigne, J.-O., Charlier, C., & Chapelle, J.-P. 2007. Le stress oxydant. ***Revue Médicale de Liège***, 62(10): 628–38.

Hammami, I., Kamoun, N., & Rebai, A. 2011. Biocontrol of Botrytis cinerea with essential oil and methanol extract of Viola odorata L . flowers. ***Archive of Applied Science Research***, 3(5): 44–51.

Hanem Fathy, K. 2012. Prospect of Botanical Biopesticides in Insect Pest Management. ***Pharmacologia***, 3(12): 641–656.

Harbourne, N., Marete, E., Jacquier, J. C., & O'Riordan, D. 2009. Effect of drying methods on the phenolic constituents of meadowsweet (Filipendula ulmaria) and willow (Salix alba). ***LWT - Food Science and Technology***, 42(9): 1468–1473.

Harman, D. 1992. Free radical theory of aging. ***Mutation Research/DNAging***, 275(3–6): 257–266.

Hasegawa, M., Kishino, H., & Yano, T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. ***Journal of Molecular Evolution***, 22(2): 160–174.

Hashempour, H., Koehbach, J., Daly, N. L., Ghassempour, A., & Gruber, C. W. 2013. Characterizing circular peptides in mixtures: sequence fragment assembly of cyclotides from a violet plant by MALDI-TOF/TOF mass spectrometry. ***Amino Acids***, 44(2): 581–595.

Hekking, W. H. A. (Ed.). 1988. Violaceae ; P. 1. Rinorea and Rinoreocarpus. ***Flora Neotropica***: 1–207. New York: New York Botanical Garden.

Herrmann, A., Burman, R., Mylne, J. S., Karlsson, G., Gullbo, J., et al. 2008. The alpine violet, Viola biflora, is a rich source of cyclotides with potent cytotoxicity. ***Phytochemistry***, 69: 939–952.

Hubert, J., Berger, M., & Dayde, J. 2008. Effects of fermentation on the phytochemical composition and antioxidant properties of soy germ. ***Food Chemistry 109***, 109: 709–721.

Ibrahim, R. M. 2017. Immunological, Cytogenetic and Hepatoprotective Effect of Viola odorata Methanolic Extract on Methotrexate Albino

Male Mice. ***Journal of Biotechnology Research Center***, 11(2): 47–55.

Ireland, David C., Colgrave, M. L., & Craik, D. J. 2006. A novel suite of cyclotides from *Viola odorata* : sequence variation and the implications for structure, function and stability. ***Biochemical Journal***, 400(1): 1–12.

Ireland, David C., Colgrave, M. L., Nguyencong, P., Daly, N. L., & Craik, D. J. 2006. Discovery and Characterization of a Linear Cyclotide from Viola odorata: Implications for the Processing of Circular Proteins. ***Journal of Molecular Biology***, 357: 1522–1535.

James W., B., Mark W., C., Maarten J. M., C., Michael F., F., Walter S., J., et al. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. ***Botanical Journal of the Linnean Society***, 181(1): 1–20.

Jennings, C., West, J., Waine, C., Craik, D., & Anderson, M. 2001. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from Oldenlandia affinis. ***Proceedings of the National Academy of Sciences***, 98(19): 10614–10619.

Jeong, Y. H., Oh, Y.-C., Cho, W.-K., Shin, H., Lee, K. Y., et al. 2016. Anti-inflammatory effects of Viola yedoensis and the application of cell extraction methods for investigating bioactive constituents in macrophages. ***BMC Complementary and Alternative Medicine***, 16: 1–16.

Jukes, T. H., & Cantor, C. R. 1969. Evolution of protein molecules. ***Mammalian Protein Metabolism*** (Academic Press), vol. III: 21–132. Munro H. N.

Jung, M., Geiger, H., Zinsmeistert, H. D., & Botanik, F. 1995. ***Tri- and tetrahydroxycoumarin derivatives from tetraphis pellucida****, 39(2): 379–381.

Kanchanapoom, T. 2007. Aromatic diglycosides from Cladogynos orientalis. ***Phytochemistry***, 68(5): 692–696.

Kanti Das, T., Wati, M. R., & Fatima-Shad, K. 2014. Oxidative Stress Gated by Fenton and Haber Weiss Reactions and Its Association With Alzheimer's Disease. ***Archives of Neuroscience***, 2(3): 1–8.

Katajamaa, M., Miettinen, J., & Ores, M. 2006. MZmine : toolbox for processing and visualization of mass spectrometry based molecular profile data ˇ ic. ***Bioinformatics***, 22(5): 634–636.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. ***Journal of Molecular Evolution***, 16: 111–120.

Kind, T., & Fiehn, O. 2016. Phytochemistry Letters Strategies for dereplication of natural compounds using high-resolution tandem mass spectrometry. ***Phytochemistry Letters***. https://doi.org/10.1016/j.phytol.2016.11.006.

Kostova, I. 2006. Synthetic and Natural Coumarins as Antioxidants. ***Mini-Reviews in Medicinal Chemistry***, 6: 365–374.

Le Floch, F., Tena, M. T., R??os, A., & Valc??rcel, M. 1998. Supercritical fluid extraction of phenol compounds from olive leaves. ***Talanta***, 46(5): 1123–1130.

Lee, S.-W., Chen, Z.-T., & Chen, C.-M. 1993. The 4α-Methylsterol Violasterol A from Viola Formosana Hayata. ***Journal of the Chinese Chemical Society***, 40(3): 305–307.

Leopoldini, M., Marino, T., Russo, N., & Toscano, M. 2004. Antioxidant Properties of Phenolic Compounds: H-Atom versus Electron Transfer Mechanism. ***The Journal of Physical Chemistry A***, 108(22): 4916–4922.

Levin, D. A., Francisco-Ortega, J., & Jansen, R. K. 1996. Hybridization and the extinction of rare plant species. ***Conservation Biology***, 10(1): 10–16.

Liochev, S. I., & Fridovich, I. 2002. The Haber-Weiss cycle—70 years later: an alternative view. ***Redox Report***, 7(1): 55–57.

Little, J. R., & McKinney, L. E. 2016. Violaceae Batsch. ***Flora of North America***, vol. 6. Oxford University Press.

Liza, M. S., Abdul Rahman, R., Mandana, B., Jinap, S., Rahmat, A., et al. 2012. Supercritical fluid extraction of bioactive flavonoid from Strobilanthes crispus (Pecah kaca) and its comparison with solvent extraction. ***International Food Research Journal***, 19(2): 503–508.

López-alarcón, C., & Denicola, A. 2013a. Evaluating the antioxidant capacity of natural products : A review on chemical and cellular-based assays. ***Analytica Chimica Acta***, 763: 1–10.

López-alarcón, C., & Denicola, A. 2013b. Evaluating the antioxidant capacity of natural products : A review on chemical and cellular-based assays. ***Analytica Chimica Acta***, 763: 1–10.

Luzzatto-Knaan, T., Melnik, A. V., & Dorrestein, P. C. 2015. Mass spectrometry tools and workflows for revealing microbial chemistry. ***Analyst***, 140(15): 4949–4966.

Mahieu, N. G., & Patti, G. J. 2017. Systems-Level Annotation of a Metabolomics Data Set

Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. ***Analytical Chemistry***, 89(19): 10397–10406.

Malécot, V., Marcussen, T., Munziger, J., Yockteng, R., & Henry, M. 2007a. On the origin of the sweet-smelling Parma violet cultivars (Violaceae): wide intraspecific hybridization, sterility, and sexual reproduction. ***American Journal of Botany***, 94(1): 29–41.

Malécot, V., Marcussen, T., Munziger, J., Yockteng, R., & Henry, M. 2007b. On the origin of the sweet-smelling Parma violet cultivars (Violaceae): wide intraspecific hybridization, sterility, and sexual reproduction. ***American Journal of Botany***, 94(1): 29–41.

Malien-Aubert, C., Dangles, O., & Amiot, M. J. 2001. Color Stability of Commercial Anthocyanin-Based Extracts in Relation to the Phenolic Composition. Protective Effects by Intra- and Intermolecular Copigmentation. ***Journal of Agricultural and Food Chemistry***, 49(1): 170–176.

Mandal, V., Mohan, Y., & Hamalatha, S. 2007. Microwave Assisted Extraction - An Innovative and Promising Extraction Tool for Medicinal Plant Research. ***Pharmacognosy Reviews***, 1(1): 1–18.

Marcussen, T., Karlsson, T., Jonsell, B., & Wind, P. 2010. Violaceae. ***Flora Nordica***, vol. 6: 12–52. Stockholm: The Bergius Foundation.

Marcussen, T., Yousefi, N., & Mehrvarz, S. S. 2012. Anatomical studies on selected species of Viola (Violaceae). ***Nordic Journal of Botany***, 30(4): 461–469.

Martucci, M. E. P., De Vos, R. C. H., Carollo, C. A., & Gobbo-Neto, L. 2014. Metabolomics as a Potential Chemotaxonomical Tool: Application in the Genus Vernonia Schreb. (J. Chamani, Ed.)***PLoS ONE***, 9(4): e93149.

Marzouk, M. M., Elkhateeb, A., Ibrahim, L. F., Sameh, R., & Kawashty, S. A. 2012. Two Cytotoxic Coumarin Glycosides from the Aerial Parts of Diceratella elliptica (DC .) Jonsell Growing in Egypt. ***Rec. Nat. Prod.***, 3: 237–241.

Masrani, K. V., Rama, H. S., & Bafna, S. L. 1974. Ultraviolet absorption spectra: Some substituted coumarins. ***Journal of Applied Chemistry and Biotechnology***, 24(6): 331–341.

Mereda JR., P., Hodalova, I., & Kucera, J. 2011. Genetic and morphological variation in Viola suavis s.l. (Violaceae) in the western Balkan Peninsula : two endemic subspecies revealed. ***Systematics and Biodiversity***, 9: 211–231.

Michener, C. D., & Sokal, R. R. 1957. A quantitative approach to a problem in classification. ***Evolution***, 11(2): 130–162.

Milman, B. L., & Zhurkovich, I. K. 2016. Mass spectral libraries: A statistical review of the visible use. ***TrAC Trends in Analytical Chemistry***, 80: 636–640.

Molnár, P., & Szabolcs, J. 1980. Occurrence of 15-cis-violaxanthin in Viola tricolor. ***Phytochemistry***, 19(4): 623–627.

Moon, H.-I., Jung, J.-C., & Lee, J. 2007. Antiplasmodial activity of triterpenoid isolated from whole plants of Viola genus from South Korea. ***Parasitology Research***, 100(3): 641–644.

Muhammad, N., Saeed, M., Awen, A., & Khan, H. 2012. Ethnomedicinal , phytochemical and pharmacological profile of genus Viola. ***Phytopharmacology***, 3(1): 214–226.

Mullis, K., Faloona, F., Scharf, S., Saiki, R. K., Horn, G. T., et al. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. ***Cold Spring Harbor symposia on quantitative biology***, 51: 263–273. Cold Spring Harbor Laboratory Press.

Murray, A. P., Nacional, U., Rodriguez, S., & Nacional, U. 2015. Chemical Constituents and Biological Activities of Plants from the Genus Viola. ***Chemistry & Biodiversity***, 12: 1777–1808.

Naruya, S., & Masatoshi, N. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. ***Molecular Biology and Evolution***, 4(4): 406–425.

Nejia, H. 2013. ***Extraction et purification de substances naturelles : comparaison de l'extraction au CO2-supercritique et des techniques conventionnelles***. Université de Toulouse.

Nepveu, F. 2008. ***APPRT- Région Midi-Pyrénées VALorisation d ' Agro-Produits : Identification de Marqueurs spécIfiques par leurs Propriétés redox Porteur : UPS***: 1–39. Document interne, Toulouse: Université Paul Sabatier.

Nguyen, H. T., Vo, B. T., Nguyen, L. H., Bernad, J., Alaeddine, M., et al. 2013. Extracts of Crinum latifolium inhibit the cell viability of mouse lymph oma cell line EL4 and induce activation of anti-tumour activity of macrophages in vitro. ***Journal of Ethnopharmacology***, 149(1): 75–83.

Niederbacher, B., Winkler, J. B., & Schnitzler, J. P. 2015. Volatile organic compounds as non-invasive markers for plant phenotyping.

Journal of Experimental Botany, 66(18): 5403–5416.

Nkhili, E., Tomao, V., El Hajji, H., El Boustani, E.-S., Chemat, F., et al. 2009. Microwave-assisted water extraction of green tea polyphenols. ***Phytochemical Analysis***, 20(5): 408–415.

Ohashi, Y., & Ohshima, M. 1992. Stress-induced expression of genes for pathogenesis-related proteins in plants. ***Plant and Cell Physiology***, 33(7): 819–826.

Overy, S. A. 2004. Application of metabolite profiling to the identification of traits in a population of tomato introgression lines. ***Journal of Experimental Botany***, 56(410): 287–296.

Özel, M. Z., Göğüş, F., & Lewis, A. C. 2006. Comparison of direct thermal desorption with water distillation and superheated water extraction for the analysis of volatile components of Rosa damascena Mill. using GCxGC-TOF/MS. ***Analytica Chimica Acta***, 566(2): 172–177.

Papet, Y., Brunet, B., & Mura, P. 2010. Headspace (HS) et micro-extraction en phase solide (SPME). Théorie et applications. ***Annales de Toxicologie Analytique***, 22(2): 75–79.

Parshad, B., Duraisamy, A. J., Saini, S., Yadav, P., Vats, P., et al. 2016. Synthesis and SAR Study of Antioxidant Potential of Polyhydroxy Coumarin Derivatives. ***Medicinal Chemistry***, 6(7): 506–514.

Parsley, N. C., Kirkpatrick, C. L., Crittenden, C. M., Rad, J. G., Hoskin, D. W., et al. 2018. PepSAVI-MS reveals anticancer and antifungal cycloviolacins in Viola odorata. ***Phytochemistry***, 152: 61–70.

Perez de Souza, L., Naake, T., Tohge, T., & Fernie, A. R. 2017. From chromatogram to analyte to metabolite. How to pick horses for courses from the massive web resources for mass spectral plant metabolomics. ***GigaScience***, 6(7): 1–20.

Pichersky, E. 2006. Biosynthesis of Plant Volatiles: Nature's Diversity and Ingenuity. ***Science***, 311(5762): 808–811.

Pranting, M., Loov, C., Burman, R., Goransson, U., & Andersson, D. I. 2010. The cyclotide cycloviolacin O2 from Viola odorata has potent bactericidal activity against Gram-negative bacteria. ***Journal of Antimicrobial Chemotherapy***, 65(9): 1964–1971.

Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., et al. 1987. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. ***Science***, 238(4825): 336–341.

Procházkoá, D., Boušová, I., & Wilhelmová, N. 2011. Antioxidant and prooxidant properties of flavonoids. ***Fitoterapia***, 82(4): 513–523.

Prost, I., Dhondt, S., Rothe, G., Vicente, J., Rodriguez, M. J., et al. 2005. Evaluation of the Antimicrobial Activities of Plant Oxylipins Supports Their Involvement in Defense against Pathogens. ***Plant Physiology***, 139(4): 1902–1913.

Pruitt, K. D. 2004. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. ***Nucleic Acids Research***, 33(Database issue): D501–D504.

Qin, Y., Yin, C., & Cheng, Z. 2013. A New Tetrahydrofuran Lignan Diglycoside from Viola tianshanica Maxim. ***Molecules***, 18(12): 13636–13644.

Quideau, S., Deffieux, D., Douat-Casassus, C., & Pouységu, L. 2011. Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. ***Angewandte Chemie International Edition***, 50(3): 586–621.

Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J. 2006. OPLS discriminant analysis : combining the strengths of PLS-DA and SIMCA classification. ***Journal of Chemometrics***, 20: 341–351.

Ribeiro, S. M., Porto, W. F., Silva, O. N., Santos, M. de O., Dias, S. C., et al. 2013. Plant Antifungal Peptides. ***Handbook of Biologically Active Peptides***: 169–179. Elsevier.

Rodrigues, V. H., Cardoso, C., Paixão, J. A., & Costa, M. M. R. R. 2007. 2,2,6,6-Tetramethyl-4-oxopiperidinium nitrate. ***Acta Crystallographica Section C Crystal Structure Communications***, 63(4): 243–245.

Rouadjia, I. 2009, October. ***Infos Matériaux : Le "champ des odeurs."*** Matériauthèque Ensad.

Ruzicka, L., Schinz, H., & Susz, B.-P. 1944. Veilchenriechstoffe.(14. Mitteilung). Zur Stereoisomerie von Hexen-(3)-ol-(1), Nonadien-(2, 6)-al-(1) und Nonadien-(2, 6)-ol-(1). ***Helvetica Chimica Acta***, 27(1): 1561–1569.

Saint-lary, L., Debrauwer, M. L., Gougeon, M. R., Paris, M. J., Thomas, M. O., et al. 2015. ***Evaluation de l'approche métabolomique pour l'authentification des extraits naturels utilisés dans le secteur arômes et parfums***. Université Nice Sophia-Antipolis.

Saldaña, G., Álvarez, I., Condón, S., & Raso, J. 2014. Microbiological Aspects Related to the

Feasibility of PEF Technology for Food Pasteurization. ***Critical Reviews in Food Science and Nutrition***, 54(11): 1415–1426.

Salem, M. A., Jüppner, J., Bajdzienko, K., & Giavalisco, P. 2016. Protocol: a fast, comprehensive and reproducible one-step extraction method for the rapid preparation of polar and semi-polar metabolites, lipids, proteins, starch and cell wall polymers from a single sample. ***Plant Methods***, 12: 1–15.

Sanger, F., & Coulson, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. ***Journal of Molecular Biology***, 94: 441–448.

Sangster, T., Major, H., Plumb, R., Wilson, A. J., & Wilson, I. D. 2006. A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis. ***The Analyst***, 131(10): 1075.

Segovia, F. J., Luengo, E., Corral-Pérez, J. J., Raso, J., & Almajano, M. P. 2015. Improvements in the aqueous extraction of polyphenols from borage (Borago officinalis L.) leaves by pulsed electric fields: Pulsed electric fields (PEF) applications. ***Industrial Crops and Products***, 65: 390–396.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., et al. 2003a. Cytoscape : A Software Environment for Integrated Models of Biomolecular Interaction Networks. ***Genome Research***, 13: 2498–2504.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., et al. 2003b. ***Cytoscape : A Software Environment for Integrated Models of Biomolecular Interaction Networks***, (Karp 2001): 2498–2504.

Shen, X.-Y., & Xie, C. 2009. Study on Antimicrobial Activity of Viola tianshanica Maxim.Extracts. ***Biotechnology***, 19: 71–73.

Sies, H. 1985. Oxidative Stress : Introductory Remarks. ***Oxidative stress***.

Skowyra, M., Calvo, M. I., Gallego, M. G., Azman, N. A. M., & Almajano, M. P. 2014. Characterization of Phytochemicals in Petals of Different Colours From Viola × wittrockiana Gams. and Their Correlation With Antioxidant Activity. ***Journal of Agricultural Science***, 6(9): 93–105.

Smedsgaard, J., & Nielsen, J. 2005. Metabolite profiling of fungi and yeast: from phenotype to metabolome by MS and informatics. ***Journal of Experimental Botany***, 56(410): 273–286.

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., & Siuzdak, G. 2006. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. ***Analytical Chemistry***, 78(3): 779–787.

Sun, Y., Du, L., Zhou, L., Miao, F., Yang, X., et al. 2011. Study on antibacterial active components from Viola yedoensis. ***Zhongguo Zhong Yao Za Zhi.***, 36(19): 2666–2671.

Swofford, D. L. 2001. ***PAUP*: Phylogenetic Analysis Using Parsimony (and other methods) 4.0.b5***. Sinauer Associates, Sunderland, Massachusetts.

Tholl, D., Boland, W., Hansel, A., Loreto, F., Röse, U. S. R., et al. 2006. Practical approaches to plant volatile analysis. ***The Plant Journal***, 45(4): 540–560.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., et al. 2016. Hybridization and extinction. ***Evolutionary Applications***, 9(7): 892–908.

Trabi, M. 2004. Tissue-Specific Expression of Head-to-Tail Cyclized Miniproteins in Violaceae and Structure Determination of the Root Cyclotide Viola hederacea root cyclotide1. ***The Plant Cell***, 16(8): 2204–2216.

Tsuagawa, H. et al. 2015. MS-DIAL: Data Independent MS/MS Deconvolution for Comprehensive Metabolome Analysis. ***Nat Methods***, 12(6): 523–526.

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., et al. 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. ***Nature Methods***, 12: 523–526.

Tsugawa, H., Kind, T., Nakabayashi, R., Yukihira, D., Tanaka, W., et al. 2016a. Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. ***Analytical Chemistry***. https://doi.org/10.1021/acs.analchem.6b00770.

Tsugawa, H., Kind, T., Nakabayashi, R., Yukihira, D., Tanaka, W., et al. 2016b. Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. ***Analytical Chemistry***. https://doi.org/10.1021/acs.analchem.6b00770.

Tutin, T. G., Moore, D. M., Valentine, D. H., Walters, S. M., & Webb, D. A. 1968. Violaceae. ***Flora Europaea***, vol. 2: 270–275. Cambridge University Press.

Uhde, G., & Ohloff, G. 1972. Parmon, eine phantomverbindung im veilchenblütenöl. ***Helvetica Chimica Acta***, 55(7): 2621–2625.

van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J. 2006. Centering, scaling, and transformations: improving the biological information content of metabolomics data. ***BMC Genomics***, 7(1): 142.

Vercammen, J., Sandra*, P., Baltussen, E., Sandra, T., & David, F. 2000. Considerations on Static and Dynamic Sorptive and Adsorptive Sampling to Monitor Volatiles Emitted by Living Plants. ***Journal of High Resolution Chromatography***, 23(9): 547–553.

Vergnes, S., Ladouce, N., Fournier, S., Ferhout, H., Attia, F., et al. 2014. Foliar treatments with Gaultheria procumbens essential oil induce defense responses and resistance against a fungal pathogen in Arabidopsis. ***Frontiers in Plant Science***, 5(477): 1–8.

Villaverde, J. J., Sandín-España, P., Sevilla-Morán, B., López-Goti, C., & Alonso-Prados, J. L. 2016. Biopesticides from Natural Products: Current Development, Legislative Framework, and Future Trends. ***BioResources***, 11(2). https://doi.org/10.15376/biores.11.2.Villaverde.

Vukics, V. 2009. ***Antioxidant Flavonoid Glycosides in Viola Tricolor L.*** SEMMELWEIS UNIVERSITY.

Vukics, V., Kery, A., Bonn, G. K., & Guttman, A. 2008. Major flavonoid components of heartsease (Viola tricolor L.) and their antioxidant activities. ***Analytical and Bioanalytical Chemistry***, 390(7): 1917–1925.

Vukics, V., Ringer, T., Kery, A., Bonn, G. K., & Guttman, A. 2008. ***Analysis of heartsease (Viola tricolor L .) flavonoid glycosides by micro-liquid chromatography coupled to multistage mass spectrometry***, 1206: 11–20.

Wang, C. K., Colgrave, M. L., Gustafson, K. R., Ireland, D. C., Goransson, U., et al. 2007. Anti-HIV cyclotides from the Chinese medicinal herb Viola yedoensis. ***Journal of Natural Products***, 71: 47–52.

Wardencki, W., Michulec, M., & Curylo, J. 2004. A review of theoretical and practical aspects of solid-phase microextraction in food analysis. ***International Journal of Food Science and Technology***, 39(7): 703–717.

Wätjen, W., Michels, G., Steffan, B., Niering, P., Chovolou, Y., et al. 2005. Low Concentrations of Flavonoids Are Protective in Rat H4IIE Cells Whereas High Concentrations Cause DNA Damage and Apoptosis. ***The Journal of Nutrition***, 135(3): 525–531.

Watson, J. D., Gilman, M., Witkowski, J., & Zoller, M. 1994a. La réaction de polymérisation en chaîne. ***ADN Recombinant***, 33(0): 79–98.

Watson, J. D., Gilman, M., Witkowski, J., & Zoller, M. 1994b. La réaction de polymérisation en chaîne. ***ADN Recombinant***, 33(0): 79–98.

White, T. J., Bruns, T., Lee, S., & Taylor, J. 1990. Amplification and Direct Sequencing of Fungal Ribosomal Rna Genes for Phylogenetics. ***PCR Protocols***, (January 2016): 315–322.

Willmer, P. 2011. The Timing and Patterning of Flowering. ***Pollination and Floral Ecology***: 483–502. Princeton: Princeton University Press.

Witkowska-Banaszczak, E., Bylka, W., Matławska, I., Goślińska, O., & Muszyński, Z. 2005. Antimicrobial activity of Viola tricolor herb. ***Fitoterapia***, 76(5): 458–61.

Wodehouse, R. P. 1935. ***Pollen grains***. Mcgraw-Hill Book Company, Inc; New York; London.

Wolfender, J., Marti, G., & Queiroz, E. F. 2010. Advances in Techniques for Profiling Crude Extracts and for the Rapid Identification of Natural Products : Dereplication , Quality Control and Metabolomics. ***Current Organic Chemistry***, 14: 1808–1832.

Wolfender, J.-L., Marti, G., Thomas, A., & Bertrand, S. 2015a. Current approaches and challenges for the metabolite profiling of complex natural extracts. ***Journal of Chromatography A***, 1382: 136–164.

Wolfender, J.-L., Marti, G., Thomas, A., & Bertrand, S. 2015b. Current approaches and challenges for the metabolite profiling of complex natural extracts. ***Journal of Chromatography A***, 1382: 136–164.

Wolfender, J.-L., Queiroz, E. F., & Hostettmann, K. 2006. The importance of hyphenated techniques in the discovery of new lead compounds from nature. ***Expert Opinion on Drug Discovery***, 1(3): 237–260.

Worley, B., & Powers, R. 2013. Multivariate Analysis in Metabolomics. ***Current Metabolomics***, 1(1): 92–107.

Xiao, J. F., Zhou, B., & Ressom, H. W. 2012. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. ***TrAC - Trends in Analytical Chemistry***, 32: 1–14.

Xie, Chen, Veitch, N. C., Houghton, P. J., & Simmonds, M. S. J. 2003. Flavone C-

glycosides from Viola yedoensis MAKINO. ***Chemical & Pharmaceutical Bulletin***, 51(10): 1204–7.

Xie, C., Veitch, N. C., Houghton, P. J., & Simmonds, M. S. J. 2003. Flavone C-Glycosides from Viola yedoensis Makino. ***Chem. Pharm. Bull.***, 51(10): 1204–1207.

Yang, J., Guo, J., & Yuan, J. 2008. In vitro antioxidant properties of rutin. ***LWT - Food Science and Technology***, 41(6): 1060–1066.

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., Bruns, N., Glukhov, E., Wodtke, A., Felicio, R. D., et al. 2013. Molecular Networking as a Dereplication Strategy. ***Journal of Natural Products***, 76: 1686–1699.

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., Bruns, N., Glukhov, E., Wodtke, A., Felicio, R. De, et al. 2013. Molecular Networking as a Dereplication Strategy. ***Journal of Natural Products***, 76: 1686–1699.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. ***Journal of Molecular Evolution***, 39(3): 306–314.

Yeshak, M. Y., Burman, R., Eriksson, C., & Göransson, U. 2012. Optimization of cyclotide extraction parameters. ***Phytochemistry Letters***, 5: 776–781.

Yockteng, R., Jr, H. E. B., Mansion, G., Dajoz, I., & Nadot, S. 2003. ***Relationships among pansies (Viola section Melanium) investigated using ITS and ISSR markers***, 241: 153–170.

Yoo, K.-O., & Jang, S.-K. 2010. Infrageneric relationships of Korean Viola based on eight chloroplast markers. ***Journal of Systematics and Evolution***, 48(6): 474–481.

Zhang, B., Huang, Y., Zhang, Q., Liu, X., Li, F., et al. 2014. Fragrance discrimination of Chinese Cymbidium species and cultivars using an electronic nose. ***Scientia Horticulturae***, 172: 271–277.

Zhou, H. Y., Hong, J. L., Shu, P., Ni, Y. J., & Qin, M. J. 2009. A new dicoumarin and anticoagulant activity from Viola yedoensis Makino. ***Fitoterapia***, 80(5): 283–285.

Deciphering the phylogeny of violets based on multiplexed genetic and metabolomic approaches

Justine Chervin[1,2,3], Thierry Talou[1], Marjorie Audonnet[2], Bernard Dumas[2], Laurent Camborde[2], Marie-Thérèse Esquerré-Tugayé[2], Christophe Roux[2], Guillaume Cabanac[4], Guillaume Marti[3*].

[1] Laboratoire de Chimie Agro-industrielle, LCA, Université de Toulouse, INRA, Toulouse, France ; *justine.chervin@ensiacet.fr*, *thierry.talou@ensiacet.fr*,

[2] Laboratoire de Recherche en Sciences Végétales [LRSV], UMR 5546, UPS/CNRS, Toulouse, France ; *dumas@lrsv.ups-tlse.fr*, *esquerre@lrsv.ups-tlse.fr*, *camborde@lrsv.ups-tlse.fr*, *roux@lrsv.ups-tlse.fr*, *marjorie.audonnet@laposte.net*

[3] UMR 152 PharmaDEV, Université de Toulouse, IRD, UPS, Toulouse, France; *guillaume.marti@univ-tlse3.fr*,

[4] UMR 5505 IRIT, Université de Toulouse, UPS, Toulouse, France ; *guillaume.cabanac@univ-tlse3.fr*

*Guillaume Marti +33 (0)5 62 25 68 46

Abstract

Molecular phylogenetics based on nucleotide sequence comparisons has profoundly influenced plant taxonomy. A comprehensive chemotaxonomical approach based on GC-MS and UHPLC-HRMS profiling was evaluated for its ability to characterize a large collection of plants all in the violet family Violoaceae (n = 111) and thus decipher the taxonomy. A thorough identification of violets is challenging due to their natural hybridization and phenotypic variability. Phylogenetic inference performed on ribosomal internal transcribed spacer sequences using maximum likelihood and neighbor-joining distance methods allowed the clear identification of 58% of the collection. Metabolomics approaches with multivariate data analysis were performed on SPME/GC-MS chromatograms of volatile compounds emitted by fresh mature flowers and on UHPLC-HRMS/MS leaf extracts for non-volatile compounds. Interestingly, molecular and biochemical approaches provided separate classifications while highlighting several common clusters. The profiling of secondary metabolites was proved most suitable for the classification of hundreds of extracts. The combination of phylogenetic and chemotaxonomic approaches, allowed the classification of 96% of the entire collection. A correlation network revealed specific chemotaxonomic biomarkers, in particular flavonoids, coumarins and cyclotides. Overall, our pioneering approach could be useful to solve misclassification issues within collections of close plant species.

Highlight

This pioneering approach combining genetic-based phylogeny and comprehensive chemotaxonomy studies through multivariate data analysis allowed the resolution of misclassification issues within violet species with the highlight of biomarkers.

Keywords:

Biomarkers; Chemotaxonomy; GC-MS; Metabolomics, Phylogeny; UHPLC-HRMS; Violets.

Abbreviations

Diode array detector (DAD)

Electrospray iononisation (ESI)

Headspace solid-phase microextraction (HS-SPME)

Internal transcribed spacer (ITS)

Orthogonal partial least square discriminant analysis (OPLS-DA)

Pair of bases (pb)

Polymerase Chain Reaction (PCR)

Principal component analysis (PCA)

Quality control (QC)

1. Introduction

The Violaceae family of plants contains 23 genera and approximately 900 species distributed all over the world. Its three largest genera represent 98% of its species, including *Viola* with approximately 600 species. This genus is separated into two main groups (sections) widely distributed in the Northern Hemisphere (Ballard, Sytsma, and Kowal 1999): violets in the *Viola* section and pansies in the *Melanium* section (Yockteng et al. 2003). *Viola* are perennial herbs with very short stems, often stolons and typically zygomorphous flowers.

The delimitation of *Viola* species into sections and subsections can be problematic in particular when using morphological characteristics. This is due to hybridization, which is relatively common in vascular plants and has a large impact on speciation events (Erben 1996). Such hybridization can contribute to rapid diversification of many plant lineages with karyotypic and genomic changes (Arnold 2006) but can also lead to the demise of rare species (Todesco et al. 2016). Hybridization may reduce a population's growth rate by the production of hybrid seed, which is produced at the expense of conspecific (pure) seed when they competed with them for resources or they possess better vigor and fertility than parents, for instance (Levin, Francisco-Ortega, and Jansen 1996). In the past, taxonomists have studied the phylogeny of *Viola* based on anatomical characteristics (Marcussen, Yousefi, and Mehrvarz 2012), chromosome numbers (Clausen 1927), nuclear ribosomal sequences (Ballard, Sytsma, and Kowal 1999) (Mereda JR., Hodalova, and Kucera 2011) and chloroplast DNA as complementary tools in particular to detect maternal lineages (Cennamo and al. 2011). Within nuclear ribosomal sequences, internal transcribed spacer (ITS) regions are highly popular in phylogenetic studies (Álvarez 2003) and have been efficiently used to differentiate plant species. Indeed, they are easily amplified by polymerase chain reaction (PCR) using universal primers (Baldwin et al. 1995) and the nucleotide sequences of ITS1 and ITS2 regions are highly polymorphic between species.

The French National collection of violets held in Toulouse (France) comprises 111 plants individuals from all over the world (Asia, America, and Europe), each characterized by its phenotypic traits like flower shape and color and to a lesser extent leaf shape. Although they all possess vernacular names, only 21% have been scientifically identified with respect to taxa. Their rigorous classification was therefore necessary to allow better identification and discrimination. A first strategy was to identify species and cultivars according to ITS regions known to be conserved. However, due to hybridization between violets, distinction of individuals based on phenotype and ITS alone is often insufficient and complementary methods such as chloroplast genome comparison are necessary. To improve the resolution of the genotypic characterization of our collection, we conducted chemotaxonomic studies using a metabolomics approach with the aim of facilitating the analyses of large datasets via the association of potential biomarkers to specific genotypes (Cox, Oh, Keasling, Colson, & Hamann, 2014b). Our first chemotaxonomic study investigated volatiles of fresh native flowers at maturity using headspace solid-phase microextraction coupled to gas chromatography with mass spectrometry (HS-SPME-GC-MS). The second analyzed the secondary metabolites of ethanolic leaf extracts by ultra-high performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS).

2. Results

2.1 Phylogenetic analysis

The whole ITS region (500 bp) was amplified for the 85 plant individuals of the French collection of violets and sequenced. In this way, 80 well-defined sequences were obtained. Other violets were either not sequenced (N = 26) or presented non-exploitable sequences (N = 9), likely due to the presence of heterozygotes and intra-individual polymorphism as suggested by the presence of double peaks for some nucleotide sites. For 31 violets, only a few double peaks along the whole sequence were observed, thus the two sequences were studied separately by manually substituting nucleotides and identification performed for both. In each case, the same species but sometimes different isolates were obtained, revealing the significant hybridization between violets of the same species. Sequence alignment (see **supplementary Fig. S1**) highlighted few differences: twenty nucleotide sites were polymorphic in ITS 1 and 12 in ITS 2. Alignment of experimental sequences was made with references of GenBank database from the National Center for Biotechnology Information (NCBI), a curated non-redundant public nucleotide sequence database used internationally as a standard for genome annotation (Pruitt 2004) used in previous pansies (Yockteng and al. 2003) and violet (Conesa, Mus, and Rosselló 2008) studies. It provided a similarity score and a no-doubt-identification was decided above 99%. This was the case for 64 violets (58% of the whole collection): two were referenced as *V. subs Rostratae* Kupffer (namely *V. labradorica* and *V. grypoceras*) (Marcussen and al. 2010), two as *V. alba* Besser, six as *V. suavis* Bieberstein, twelve as *V.* sp Hearn cult 33, one as *V. verecunda* A Gray, one as *V. mandshurica* Becker and forty as *V. odorata* Linneaus.

Clear defined sequences (N = 80) were compared by a distance based reconstruction method to identify clades with a bootstrap threshold value of 80%. Six clades of various sizes were highlighted and seemed to correlate with specific species (**supplementary Fig. S2**). Fifty-five percent of the sequences grouped together in one clade, around 15% in each of two clades and 5% in each of the remaining three. Therefore, while the collection presented some genotype variability, the majority of the violets were found to possess similar DNA sequences.

Comparisons were then made with references found in GenBank nucleotides database to reconstruct maximum parsimony (**Fig. 1 A**) and the maximum likelihood phylogenetic trees, the latter built with few selected sequences of distinct clades for visual quality (**Fig. 1 B**). These references were: *Viola alba isolate* ALB184 V6 (GenBank: EU413916), *Viola sp* Hearn cult33 (GenBank: DQ521291), *Viola suavis isolate* SUW170 V17 (GenBank: EU413930), *Viola mandshurica* isolate TKM201586 (GenBank: KX394615), *Viola alba subsp dehnhardtii* clone 200781 (GenBank: EU430656), *Viola odorata* isolate ODO178 V11 (GenBank: EU413922), isolate ODO214 V8 (GenBank: EU413919), isolate ODO182 V3 (GenBank: EU413918). One *V. odorata* botanically identified obtained for the Museum of Natural History of Toulouse was also included and 99.7% correspondence with *V.* odorata from NCBI was obtained. Cladograms was thus obtained and *Rinorea ledermannii* as species of Violaceae was used to root the tree. As used previously, a bootstrap threshold value of 80% was fixed and same clades were highlighted in both trees with confirmed identifications made as *V. odorata* for the orange clade with 44 violets, *V. alba* for the purple clade with 5 violets, *V. sp* Hearn for the pink clade with 14 violets, *V. suavis* for the turquoise clade with 11 violets, and *V. subs Rostratae* for the red clade with 3 violets; an unidentified cluster in green with 4 violets was well separated from the other clades.

Fig. 1 Cladograms obtained with maximum parsimony (A) and maximum likelihood (B) method showing genetic relationship of selected experimental sequences and references based on ITS sequences. To improve the quality of the figure, only sequences of a few characteristic samples from each clade observed with distance model were selected for B. The outgroup is used to root the tree. Bootstrap values to study robustness are indicated in red (in colour).

It was quite interesting to also note certain similarities in terms of phenotypic traits of violets belonging to the same clade. A lavender and sometimes white double flower characterized *V. alba* **(supplementary data Fig. S3 a)** against mainly big dark purple and sometimes single white flowers for *V.* sp Hearn **(Fig. S3 b)**, small violet single flowers for *V. subs Rostratae* **(Fig.S3 c)** and single mauve flowers for *V. suavis* **(Fig. S3 d)**. Nevertheless, such phenotypes cannot be described as species-specific since various colors and shapes were observed for different violets all belonging to *V. odorata* **(Fig. S3 e)**, for example either lavender or white double flowers.

Besides, botanical studies based on morphological and anatomical (bracteoles and leaf epidermis) aspects were conducted on three characteristic violets belonging to each clade. For *V. subs Rostratae*, both *V. labradorica* and *V. grypoceras* species were described. All violets possess heart shape leave with either rounded or sharped extremity of various size. Average size of 3 to 4 cm were observed but

bigger and smaller leaves were respectively noted for *V. sp* Hearn and violets belonging to *V. subs Rostratae*. Average flower size is around 1 to 2 cm and bigger and smaller flowers were as well observed for *V. sp* Hearn and *V. subs Rostratae*.

Main anatomical divergences focused on the presence, shape and localization of glandular and non-glandular trichomes, the size of the anisocytic-type stomata bigger for *V. sp* Hearn and *V. labradorica* as well as the shape of adjoining cells, either rounded or epidermal. Presence of mucilage and hypertrophied epidermal cells were also of varying importance depending on the clade. Tables of resemblances and divergences can be noted in **supplementary Table S2**. This study allowed the validation of species identification by comparison with literature (Bonnier and Douin. 1990) (Tutin and al. 1968)

Regarding the identification, data could be more precise than that obtained here, for example violet of Toulouse being identified as *Viola alba* (purple cluster). Indeed, in the literature it is known as *Viola alba sups dehnhardtii* (Malécot, Marcussen, Munziger, Yockteng, & Henry, 2007b). In addition, violets belonging to *Viola odorata* were distributed among the three previously quoted isolates (ODO178, ODO182 and ODO214), however, their sequences were highly analogous with 99% and even 100% similarity. This is indicative of the insufficiency of the ITS sequence for precise identification and the need for further methods such as chloroplast DNA in such cases. Indeed, chloroplast DNA markers showed variability among *Viola* and allowed the distinction of sections and subsections as in the analysis of Korean *Viola* (Yoo and Jang 2010). Nevertheless, since the main goal of our study was to obtain a classification of violets into various clades limiting their identification to the species level was sufficient.

These genetic results were then used as a base model on which to apply a new strategy based on chemotaxonomic studies. From the dataset of violets analyzed by each of two methods (GC-MS and UHPLC-HRMS), genetically well-identified violets (N = 64) were used as training set on which to build statistical models. The unidentified violets were used as a prediction set to try to classify them and attribute a potential identification (**Fig. 2**).

Fig. 2 Strategy of classification based on training and prediction sets. Pie charts show the distribution of results of classification (in colour).

2.2 HS-SPME-GC-MS

2.2.1 HS-SPME protocol optimization

A second order Doehlert design (Ferreira 2004) was used to optimize the release and the trapping of volatiles emitted by fresh flowers at maturity. Two time parameters were studied: the time leading up to volatile release and the adsorption time of volatiles trapped onto the SPME fiber. This optimization was

undertaken with the violet of Toulouse at room temperature. Plants were kept under a bell-shaped hood to ensure non-degradation of the vegetal matter. Based on one report in the literature, the expertise domain was defined as: a time for volatile release between 20 and 80 min and an adsorption time between 15 and 45 min (Vercammen et al. 2000). The Doehlert matrix indicated nine experiments to perform with two repetitions at the center of the experimental domain (**supplementary Table S3**). Two responses were observed, the number of peaks and their global integration (**supplementary Tables S4 and S5)**. The repeatability of the plan was confirmed by the repetition of the same order of magnitude: number of peaks 13 ± 1 and integration 21 ± 3. Method optimization was sought using the maximum final parameter values: 65 min to release the volatiles and 37 min for the adsorption. This theoretical maximum was experimentally confirmed by the analyses of three different plants on three days (**supplementary Table S6)**. For routine analyses, a proposed technical maximum of 60 min and 30 min was validated upon similar results to those of the theoretical maximum with a standard deviation of only 2% regarding the peak integration. This practical optimum was therein conserved.

2.2.2 HS-SPME-GC-MS-based metabolomics approach

HS-SPME-GC-MS profiles of all 39 analyzed flowering plants afforded 82 features (*m/z*-RT pairs) in PI mode. As a preliminary step, principal component analysis (PCA) was applied as an exploratory tool to provide an unsupervised overview of the GC-MS volatile fingerprints. Use of binary data based on the presence or absence of peaks in each sample was preferred to avoid intensity variability (**Fig. 3**). Indeed, analyses of the same violet of Toulouse throughout one season (February-March) provided quantitative but not qualitative variations and the identification of the same profile with only peak intensity variations (**supplementary Fig. S4**). This PCA revealed a group composed of only violets belonging to *V. sp* Hearn that were well-separated from the four other clusters. This separation was dependent on PC1 while the other species were separated by PC2. This observation suggests that *V.* sp Hearn features volatile profiles that differ considerably from the other species which among themselves share common features.

Fig. 3 PCA plot of EI data of GC-MS flower extracts with binary data (in colour).

We then applied a supervised orthogonal partial least square discriminant analysis (OPLS-DA) in order to try to classify the unidentified violets by projection and thus extend the classification of our collection. A base model was built with only violets identified genetically (N = 19). The quality of model prediction was quite satisfactory (R2Y = 0.93, Q2Y = 0.505, CV-ANOVA *p*-value = 0.59) and a permutation test assessed its validity (**supplementary Fig. S5**). A better cluster pattern was achieved with well-defined and non-overlapping clusters relating to *V. subs Rostratae*, *V.* sp Hearn, *V. alba*, *V. odorata* and *V. suavis*, (**Fig. 4A**). Prediction of unidentified violets (N = 20) was then processed taking into account

95% of the data (**Fig. 4B**). The results were quite weak as only ten violets in twenty (i.e., 50%) were projected and classified into the defined clusters. After prediction, three clusters related to *V. alba*, *V. suavis* and *V. odorata* overlap, indicating their likely presentation of very similar profiles that prevents their net separation.

Fig. 4 OPLS-DA based model (A) and prediction set of unidentified violets (B). Well-projected violets are indicated by a star (in colour).

2.2.3 Identification of violet species-specific volatile biomarkers

Despite the poor classification of violets with volatile analyses, some species-specific biomarkers were highlighted through database interrogation (**Fig. 5**). A limonene profile (**supplementary Fig. S6 a**) was revealed for violets belonging to *V.* sp Hearn with the presence of some terpenes such as α-terpinene (**S6 b**). Two ionone profiles enabled a distinction between violets belonging to *V. alba* and *V. suavis*. Regarding *V. alba*, the three ionones classically found for violets were noted, i.e. α-ionone (**S6 c**), β-ionone (**S6 d**) and dihydro-β-ionone (**S7 e**), as well as one methoxybenzene named methylanisole (**S6 f**). This latter is absent for *V. suavis* but we suspect the presence of other ionone derivatives near the area of well-identified ionones as two new peaks appeared. Finally, violets identified as *V. subs Rostratae* have no quantitative volatiles according to the absence of chromatographic peaks, correlating to nearly no flower fragrance compared to the others. These profiles correlate well with our findings using the OPLS-DA base model, showing high levels of similarity between *V. suavis* and *V. alba* in view of the presence of ionones, in contrast to both *V. sp* Hearn, which while devoid of these compounds are rich in others, and *V. subs Rostratae* with no volatiles. These ionone profiles thus explain the well-defined clusters of *V.* sp Hearn and *V. subs Rostratae* and the tendency for the others to merge.

Fig. 5 Highlighting of specific volatile profiles with identification of characteristic volatile biomarkers. Stars indicate well-projected violets (in colour).

2.3 UHPLC-HRMS-based metabolomics approach

UHPLC-HRMS profiles of all 119 extracts (110 crude extracts, 1 botanical reference *V. odorata* and 8 QC samples prepared by pooling aliquots of all extracts) afforded 208 and 198 features in NI and PI modes respectively and 120 common features between both ionization modes.

As expected, PCA-X grouped QC near the plot center with violet extracts distributed around this central point (**Fig. 6**). Clusters correlated with species were highlighted according to genetic results. They matched with the five identified species: *V. subs Rostratae*, *V. suavis*, *V. alba, V. odorata*, and *V.* sp

Hearn. These results indicated then a stable chemical composition within aerial parts of violets belonging to the same species. Moreover, it can be noticed the formation of two main groups based on PC1 with on the one hand *V. odorata* and on the other hand *V. suavis*, *V. sp* Hearn, *V. alba* and *V. subs Rostratae*. According to this configuration, *V. alba* and *V. subs Rostratae* seem to present similar profiles compared to the others, especially *V. odorata* and *V. suavis* which are both well-separated from the other clusters suggesting they possess specific metabolites.

Fig. 6 PCA-X plot of ESI-NI/PI combined dataset of UHPLC-HRMS leaf extracts of violets (in colour).

After PCA-X, we applied OPLS-DA analysis in order to try to classify the unidentified violets by projection and extend the classification of the collection. A base model was first built with only genetically-identified violets (N = 62) (**Fig. 7A)**. This model was quite well-defined (R2Y = 0.962, Q2Y = 0.822, CV-ANOVA *p*-value < 0.001) and validated by a permutation test (**supplementary Fig. S7**). Prediction of unidentified violets (N = 46) was then processed, taking into account 95% of the data. Forty-two out of the forty-six unknown violets were putatively identified by projection and thereby affiliated to clear-defined clusters characterized by specific chemical profiles and related to specific species (**Fig. 7B**). Three violets were associated with *V. alba*, two with *V. subs Rostratae*, nine with *V.* sp Hearn, two with *V. suavis* and twenty-six with *V. odorata*. Four violets remained totally separate from these clusters and could therefore be neither classified nor putatively identified.
In our previous genetic results, some violets were not clearly identified due to similarity scores under 99% but were then affiliated to one clade through distance model phylogeny. Upon comparison with the UHPLC-HRMS results, predictions correlated with genetic affiliation thus reinforcing the putative identification of these violets which allowed reaching the classification of 96 % of the collection. The four undefined violets were not well correlated to any specific cluster and regarding genetic results, two were not sequenced and the others revealed clear sequence with low identification scores. Thus 4% of our collection of violets remain unidentified.

Fig. 7 OPLS-DA base model (A) and prediction set of unidentified violets (B) taking into account 95% of the data. Well projected violets are indicated by a star (in colour).

The stability of leaves metabolome over time was assessed by a kinetic study. Every two months, ten leaves of violet of Toulouse were collected and extracted as depicted in section 2.4. The PCA score plot (**supplementary data Fig. S8)** containing few selected extracts of each species representative of each cluster was built. All kinetic extracts of violet of Toulouse were distinctly projected within *V. alba* cluster. Thus, the model validity obtained at one precise moment of the season may be more generally interpreted as attested by these kinetic results.

To complete this study, phylogenetic reconstruction was tried by combining ITS sequence and LC-MS data through a parsimony method. Since no standardization method was found in the literature to treat LC-MS data, the binary matrix was built on the average intensity as threshold, independently for each variable. Thus superior values were indexed with 1 against 0 on the contrary. All data were combined and the resulting combined tree (**supplementary Fig. S9**) conducted to the observation of the majority of the groups identified in ITS phylogenetic tree. The only difference focused on the scission of *V. odorata* clade into two groups. Complementary hierarchical cluster analysis (HCA) of LCMS data highlighted the same organization as one branch related to *V. odorata* was noted with a separation into two groups (**supplementary Fig. S10**). These results could be then explained by the presence of different varieties within *V. odorata* which seem highlighted by combination of data.

2.4 Identification of UHPLC-HRMS chemotaxonomic biomarkers of violet species

In an attempt to take our classification further and identify potential chemotaxonomic biomarkers, we built a correlation network (**Fig. 8**). This was composed of a set of nodes corresponding to UHPLC-HRMS peaks and correlated by edges using Spearman correlation. This correlation was based on the relative intensity of each peak in each extract. Compounds having the same pattern would thus be correlated together.

The network obtained was composed of around 250 nodes and allowed us to observe, as expected, clear clusters of metabolites specific to identified species. Non-discriminant metabolites shared by different and sometimes all species were also found plotted at the center of the network (**supplementary Fig. S11**).

Fig. 8 Correlation network based on ESI-NI/PI combined dataset of UHPLC-HRMS performed on leaf extracts and identification of characteristic biomarkers. Color tag is based on species identification by genetic analysis. Node size was emphasized based on OPLS-DA coefficient value. Putative structural identification was based on HRMS and MS/MS spectra (in colour).

The identification of characteristic features of each species was undertaken using *in silico* fragmentation to compare against our own in-house *Viola* databases as well as local databases within MS-FINDER (**Table 1**). The acquisition of UHPLC–HRMS profiles in data-dependent analysis mode provided accurate mass-to-charge ratios for molecular formula determination along with MS/MS fragments used for peak assignments. Regarding the various databases used, our *Viola* databases provided more pertinent identifications than those based on natural products databases within MS-FINDER, which allowed a more general identification for which mainly chemical class was relevant. Every correspondence between experimental and *in silico* fragments is presented in **supplementary Fig. S12**. OPLS-DA coefficients were determined for every feature and for each identified species. Characteristic features were correlated to the highest OPLS-DA coefficient scores.

It is interesting to note the abundance of *C*- and *O*- glycosylated coumarin and flavonoid derivatives. *V. odorata* presents mainly cyclotides as biomarkers with $[M+3H]^{3+}$ showing a characteristic isotopic shape, whereas *V. alba*, *V. subs Rostratae*, *V. sp* Hearn and *V. suavis* seem well characterized by polyphenols, in particular flavonoids and coumarins specific to *V. alba*. As in the above-mentioned UHPLC-HRMS-based metabolomics approach, the identification of two groups was reinforced by the putative distinguished biomarkers.

Table 1: Summary of all dereplicated compounds with MS-FINDER

3. Discussion

We have described the taxonomic delimitation of the French national collection of violets held in the municipal greenhouses of Toulouse performed by combining their genotypic and chemotaxonomic profiles. Untargeted metabolomics approaches coupled to LC-MS (Martucci et al. 2014) as well as SPME GC-MS (Khalil, Fekry, and Farag 2017) phytochemical characterizations have already demonstrated their potential in taxonomic classification on different species. As explained previously regarding the species concept (Smedsgaard and Nielsen 2005), the classification and understanding of any species rely on three main axes to distinguish their phenotypic characteristics, ecological behavior and genome. However, these aspects are not always sufficient to well-differentiate closely related species or during the speciation process. Since the metabolome represents the final picture resulting from genomic and environmental interaction, metabolite profiling appears to offer the data required to bridge the gap left by genetic-based taxonomical studies performed alone. Implementation of such metabolite profiling should thus provide new clues with which to refine genomic-based classification. Genetic analyses based on the study of ITS sequences allowed the classification of 72% of the collection of violets into six different clades with distance-based phylogeny reconstruction. Nevertheless, only 58% of the collection clearly matched to one reference found in the NCBI database with a similarity score above 99%. To improve this classification, we undertook chemotaxonomic studies involving native flower volatile profiling obtained by HS-SPME/GC-MS and UHPLC-HRMS analyses of fresh leaf extracts. We found the volatile profiling unsuitable for discrimination of the complete collection, mainly due to the small number of features inducing a less significant variability among the species. In addition, this technique is quite sensitive, in particular to environmental conditions, and is long to perform (around two and half hours for one analysis). Moreover, profiling the volatile emitted by flowers is limited to the flowering period, which is only two winter months for the *Viola* genus. Analysis of a large collection is therefore laborious and was deemed unsuitable in our case. On the contrary, UHPLC-HRMS analyses worked well for our classification purposes. Indeed, the strategy used here based on genetic results to build a statistical model from which a prediction set was used for unidentified violets,

the classification rate rose to 96 % of the whole collection composed of 111 violets. Importantly, this technique is easy to implement, in our case requiring only an ethanolic extraction which lasted 30 minutes using around 1g of fresh sample (the equivalent of a dozen leaves in the case of *Viola* plantlets). Moreover, separations are of better quality and are achieved in a shorter time (around thirty minutes per sample) thereby reducing the cost of analysis (Dong and Zhang 2014) making it better suited to the analysis of a whole collection containing hundreds of samples. The overall good predictive quality of our LCMS-based models could be explained by the larger number of variables involved (around 250 in total). The main limitation remains the false discovery rate (Benjamini and Hochberg 1995); a common way of circumventing this is to employ the diagnostic tool CV-ANOVA (analysis of variance testing of cross-validated predicted residuals) to assess the reliability of the OPLS model. This tool displays the *p*-value indicating the probability level where a model with one value is the result of chance. A *p*-value lower than 0.05 is correlated to a significant model. Permutation tests are also often applied in validation procedure (Eriksson, Trygg, and Wold 2008).

Cross referencing with the genetic classification reinforced the putative identification of unknown violets. In this way, more than half of the collection was identified as *V. odorata*, around 18 % as *V.* sp Hearn, 5 % as *V. alba*, 7 % as *V. suavis* and 4 % as *V. subs Rostratae*.

While *V. alba, V. suavis* and *V. odorata* possess similar volatile profiles principally composed of ionones, they could be better differentiated by their non-volatile secondary metabolites from coumarines, flavonoids and various non-phenolic classes of compounds. On the contrary, while *V. sp* Hearn and *V. subs Rostratae* possess similar secondary metabolite markers within the flavonoid class, they could be accurately distinguished by their volatile profiles with the presence of terpenic compounds within the first group and the absence of quantitative volatiles from the second.

Metabolite profiling, as depicted in a previous study of tomato species (Overy 2004), allowed the examination of biological diversity with clear discrimination of violet species; this permitted their segregation through multivariate data analysis. The chemical profiling methods used in our study provided complementary information and contributed towards the establishment of a significant relationship with ITS phylogenetic analysis for *Viola* genus.

Consensus tree based on ITS sequences and binary LC-MS matrix validated the majority of the clades correlated to one species. However, for *V. odorata* a scission was observed conducting to two groups instead of the only one observed with ITS sequence. The first hypothesis was about the potential loss of information due to the binarization of LC-MS data. Indeed, there are continuous data with semi-quantitative information at one variable level. But HCA of LC-MS data highlighted the same phenomenon which could then be explained by the potential presence of several varieties. Comparison of both clades highlighted a highest concentration of cycloviolacin O9 and an unidentified cyclotide characterized by $[M+H]^{2+}$ = 1057.7927 (cf. Table 1) in the major group of *V. odorata*. Independent analysis of ITS data did not permit the clear identification of varieties but combination of information

conducted to this revelation thus reinforcing the complementarity of the methods and the relevance of metabolome profiles in classification.

One of the main limitations of our workflow concerns the non-referenced sequences. As depicted for the four unclassified violets (grey dots, Fig. 7b), two of them displayed exploitable ITS sequences, with no satisfactory matches within the NCBI bank (> 99%). According to our combined results, these violets could be new species which warrants further investigation. Another drawback of our procedure is the identification of metabolite biomarkers which is time-consuming and elusive despite recent advances in this area (Perez de Souza and al. 2017). Contrary to other "-omics" based approaches, such as genomics and proteomics, metabolomics is still in its infancy and requires protocol standardization to overcome this identification shortfall. Nevertheless, taxonomy can expect a strong complementary partnership between comprehensive metabolome profiling methods and genetic-based approaches in the near future.

4. Materials and methods

4.1 Plant material

Violet leaves were harvested in spring 2016 from the whole collection of violet plants maintained in the municipal greenhouses of Toulouse, France: two leaves for genetic analysis and ten for UHPLC-HRMS analysis. Directly after collection, they were immerged in liquid nitrogen to stop any enzymatic activity and then stored at -80°C. Genetic analysis was undertaken on frozen leaves; for UHPLC-HRMS extracts, the vegetable matter was lyophilized and ground into powder using a microtube homogenizer (BeadBug, 40W).

For volatile profiling, the whole native flowering plants were analyzed between February and April 2016 according to their blossoming quality.

4.2 DNA amplification, purification and sequencing

Genetic analyses were performed directly on a small portion (approx. 1mm²) of a frozen fresh leaf of the various plants without previous DNA extraction. Species discrimination was accomplished through amplification of the whole internal transcribed sequence region ITS1-5.8S-ITS2 by PCR (Mullis et al. 1986) using the KAPA3G Plant PCR kit (Kapabiosystems).

The forward primer ITS (GGAGAAGTCGTAACAAGGTTTCCG) (Aceto and al. 1999) and the reverse primer ITS (GCTCGCCGTTACTAGGGGAATC) were used to amplify the ITS region according to the manufacturer's recommendations. Briefly, a small sample of frozen leaf was mixed with 50 µl of 1X PCR buffer solution, containing 1 mM of $MgCl_2$ and 0.2 mM of dNTPs, 0.3 µM of each primer and 1 Unit of KAPA3G DNA polymerase, in sterile conditions under a fume hood, with filter tips, after a UV cycle of 20 min. Amplification was performed on a Mastercycler (Eppendorf), starting with an initial denaturation step of 10 min at 95 °C followed by 35 cycles composed of a denaturation step at 95 °C for 30 sec, followed by an annealing step at 62 °C for 20 sec, and an extension step at 72 °C for 30 sec.

PCR amplicons were resolved on a 1 % agarose gel, after 30 min at 100 V. DNA revelation was achieved under UV with ethidium bromide (EtBr) solution. PCR products were then purified with Promega Wizard® SV Gel and PCR Clean-Up System (Promega Corp., Madison, Wisconsin). Finally DNA was collected in 50 µL of nuclease-free water. Sanger sequencing was performed by GATC Biotech (Constance, German). Sequence identification was achieved by matching against references found within the National Center for Biotechnology Information (NCBI) database (Pruitt 2004).

4.3 Sequence alignment and phylogenetic inference

Phylogenetic analyses were performed using ITS1-5.8S-ITS2 sequences of references available from GenBank (see **supplementary Table S1**). Alignment of ITS sequences was performed using the ClustalX program within the CLC Main Workbench software version 7.8.1 (Qiagen Aarhus A/S) and then manually adjusted. Two inferences were then performed: a distance analysis by Neighbor Joining (using Kimura 80 as distance setting) to define sequence similarity, and a phylogenetic analysis using the maximum likelihood method. Maximum parsimony trees of ITS sequences as well as combined ITS and metabolomics data were also built using PAUP* 4.0 software (Swofford 2001). Searches were heuristic with the parsimony criterion, with MulTrees and ACCTRAN options in effect and TBR branch swapping algorithm. Bootstrap values were derived from 1000 replicates and were calculated for both inferences to define clade robustness. The LCMS data were binarized according to mean peak intensity with a value of "1" for intensity above mean and "0" on the contrary. Outgroups composed of sequences of *Viola* species from section *Melanium* (pansies), the different genus *Rinorea* from Violaceae and *Malpighia* from Malpighiaceae were also introduced to obtain rooted trees.

4.4 Leaf metabolite extraction

Metabolites were extracted by adding ten volumes of 80% ethanol to the powdered material (80.0 ± 0.2 mg). The solutions were sonicated in a bath (Fisher Scientific, Illkirch, France) at room temperature for 10 min, then centrifuged for 2 min at 5°C and 14 000 rpm. The supernatant was put aside and this procedure then repeated once on the residue with fresh solvent. In order to discard chlorophyll pigments, the combined liquid extracts were mixed with silica-C_{18} in final weight proportions of 4 silica powder per 1 initial plant material. The mixture was centrifuged and the resulting liquid extract put in an UHPLC-HRMS vial for subsequent UHPLC-HRMS analysis.

4.5 UHPLC-HRMS profiling

UHPLC-HRMS analyses were performed with diode array detector (DAD) on a UHPLC-LTQ Orbitrap XL instrument (Ultimate 3000, Thermo Fisher Scientific, Hemel Hempstead, UK) as previously reported (Chervin and al. 2017). Briefly, the LC–MS system was run using a Acquity UPLC BEH C18 column (100 × 2.1 mm i.d., 1.7 µm, Waters, MA, USA) equipped with a guard column. The mobile phase A (MPA) was water with 0.1% formic acid (FA) and mobile phase B (MPB) was acetonitrile with 0.1%

FA. The solvent gradient was: 0 min, 95% MPA; 0.5 min 95% MPA; 12 min, 5% MPA; 15 min, 5% MPA, 15.5 min, 95% MPA; 19 min, 95% MPA. The flow rate was 0.3 mL/min, the column temperature was set to 40 °C and injection volume fixed to 2 µL. The UV detection was performed by a diode array detector (DAD) from 210 to 400 nm. Mass detection was performed using an electrospray source (ESI) in positive ionization (PI) and negative ionization (NI) modes. The mass scanning range was m/z 100–2000 Da. Each full MS scan was followed by data dependent MS/MS on the three most intense peaks.

4.6 Native plant volatiles emission

Analysis of the volatiles of fresh mature flowers from native plants was performed using HS-SPME coupled to GC-MS. The whole plant was analyzed to ensure that no fragile vegetal matter was destroyed in preparation. Optimization of the protocol was undertaken on the violet of Toulouse with use of a Doelhert experimental design and the following parameters were obtained at ambient temperature: time leading up to volatile release of 1 hour under a bell for followed by 30 min of adsorption on the divinylbenzene/carboxen/ polydimethylsiloxane (DVB/CAR/PDMS) SPME fiber (Supelco, Oakville, ON, Canada). The fiber was retracted and transferred to the GC-MS injection port at 250°C with desorption time of 5 min.

4.7 GC-MS profiling

GC-MS analyses were performed on an Agilent Technologies GC system 7890B equipped with HP 5 MS column (30 m x 0.25 mm id x 0.25 µm film thickness; Agilent, Santa Clara, USA) and coupled to Agilent Technologies 5977B MSD mass chromatograph. The temperature of the injector was maintained at 250°C. The following temperature gradient of the oven was used: 60°C for 10 min, then increase to 130°C at 4°C/min, then to 150°C at 2°C/min and finally ramped at a rate of 25°C/min to 250°C and maintained for 10 min. Helium was the carrier gas with flow rate of 1 mL/min. Splitless injection was used and no solvent delay was fixed. The mass spectrometer detector was operated in EI mode at 70 eV. The scan range was set at 50 to 550 Da.

4.8 Data processing

The GC-MS and UHPLC-HRMS raw data were converted to abf files (*Reifycs* Abf Converter) and processed with MS-DIAL version 2.74 (Tsuagawa 2015) for mass signal extraction respectively between 50 and 600 Da from 4 to 40 min and between 100 and 1500 Da from 0 to 14 min.

For LC-HRMS/MS, respective MS1 and MS2 tolerance were set to 0.01 and 0.2 Da in centroid mode. The optimized detection threshold was set to 4×10^5 for NI and $1.6\text{x}10^5$ for PI concerning MS1 and 10 for MS2 in both cases. Adducts and complexes were identified to exclude them from the final peak list. Finally, the peaks were aligned on a quality control (QC) reference file with a retention time tolerance of 0.1 min and a mass tolerance of 0.025 Da. Identification was performed with a local database built on MS-FINDER model (Tsugawa and al. 2016) and based on *Viola* genus.

Regarding GC-MS, the optimized detection threshold was set to 500. Finally, the peaks were aligned on the reference file of the violet of Toulouse with a retention time tolerance of 0.075 min. Identification was based on retention time matching with NIST MS Search 2.2 (National Institute of Standards and Technology, Gaithersburg, MD, USA) mass spectral database and EI-MS databases proposed with MS-DIAL.

The resulting peak lists were then exported to comma-separated value (CSV) format prior to multivariate data analysis (MVA) using SIMCA-P+ (version 14.0, Umerics, Umea, Sweden).

For cyclotide detection, positive profiles were processed with MZmine 2.29. A peak list was built and deconvoluted. The identification of isotopes was made with a maximum charge fixed at 5, a m/z tolerance of 10 ppm and a retention time tolerance of 0.2 min. The representative isotope selected was the lowest.

4.9 Statistical analysis

Both UHPLC-HRMS CSV files of NI and PI mode were combined using MScombine package (Calderón-Santiago et al., 2016). The resulting CSV and GC-MS files were then separately imported into SIMCA-P+. For MVA, all data were log transformed and pareto scaled. The unsupervised principal component analysis (PCA) allowed the visualization of the global organization of the samples. The supervised orthogonal partial least square discriminant analysis (OPLS-DA) was made with the names of species as qualitative Y input. Clearly genetically identified plant individuals (similarity score > 99%) were the starting point from which to build OPLS-DA models. Unidentified violets were predicted by projection in potential observable clusters. Coefficient scores were used to rank variables according to their chemotaxonomic biomarker potential. For each model, a leave-one-subject-out cross-validation was performed to assess the model fit. The validity of the discriminant model was verified using permutation tests (Y-scrambling).

4.10 Identification of significant features

Using the OPLS-DA regression analysis results, molecular formulae and structural identification of significant features were calculated with MS-FINDER 2.22 (Tsugawa et al. 2016) based on our previously described model (Chervin et al., 2017). An in-house database based on *Viola* matches inside the Dictionary of Natural Products (DNP, CRC press, v25:2) extended to the chemical classes and of local databases within MS-FINDER (KNApSAck, PlantCyc and UNPD) were used. For each compound, the results afforded several candidates and ranked them according to their similarity score, which was based on comparison between experimental MS/MS fragments and *in silico* spectra of candidates.

For cyclotide identification, a homemade database based on cyclotides found in the DNP was built, clarifying m/z of $[M+3H]^{3+}$ and $[M+4H]^{4+}$, and imported within MZmine with a *m/z* tolerance of 10 ppm.

4.11 Correlation network

The text file format exported from MS-DIAL was cleaned-up by eliminating adducts and keeping only peaks with MS/MS data before importation into MetamapR (version 1.4.0) (Grapov et al. 2015). A correlation network was created using the Spearman model and a p-value set to 5.10^{-6}. The calculated edge list was then downloaded and processed with Cytoscape 2.8.3 (Shannon et al. 2003). An attribute file containing all processed information was imported to improve network visualization and interpretation. The mean average intensity of each peak was added to specify the relative part of each compound per species with the use of the plugin "Color Nodes".

5. Acknowledgements

The authors would like to thank the Regional Council Occitanie for sponsoring this project (Project CLE 13053062), as well as the municipal greenhouses of Toulouse for making the violet collection available to its collaborators. The authors also would like to thank Professor Isabelle Fourasté for her help in botanical identification of Violets traits.

6. Funding

This work was supported by the Regional Council Occitanie (Project CLE 13053062)

7. Supplementary data

Tables S1. Sequence references from Genbank
Figure S1. Sequence alignment of experimental sequences and *Viola* references
Figure S2. Cladogram based on a distance model with 80 well-defined experimental sequences and some references. These latter were informed and the other dot corresponds to samples. The outgroup was used to root the tree. Bootstrap values to study robustness are indicated in red.
Figure S3: Flower phenotypes of violets belonging to *V. alba* (a), *V. sp* Hearn (b*), V. subs Rostratae* (c), *V. suavis* (d), and *V. odorata* (e).
Table S2 : Resemblance and divergence of morphological and anatomical characters of the six main species identified in the collection
Table S3: Doehlert matrix (coded values X1=balancing, X2=adsorption) and respective experimental conditions (real values)
Table S4 : Experimental number of peak (Yexp) and predicted one by the quadratic model (Ycalc)
Table S5 Experimental integration of peaks (Yexp), and predicted one by the quadratic model (Ycalc)
Table S6: Validation of the theoretical and practical optimums for the HS-SPME-GC-MS analysis of native volatiles
Figure S4: Kinetic study of volatile emission showing quantitative but not qualitative variabilities
Figure S5: Permutation test of GC-MS data
Figure S6 In silico identification of native volatiles biomarkers
Figure S7: Permutation test of UHPLC-HRMS data
Figure S8: PCA of kinetic study of secondary metabolite compounds from violet of Toulouse (V. alba) showing no qualitative variation due to a good clustering pattern
Figure S9: Consensus tree combining ITS sequences and LC-MS data built with parsimony method. The outgroup is used to root the tree. Bootstrap values to study robustness are indicated in red.
Figure S10:HCA analysis based on PCA plot confirming the five distinct clades correlated to species
Figure S11: Correlation network of NI and PI UHPLC data

Figure S12 In silico identification of secondary metabolite biomarkers

8. References

Aceto, S., Caputo, P., Cozzolino, S., Gaudio, L., & Moretti, A. 1999. Phylogeny and evolution of Orchis and allied genera based on ITS DNA variation: morphological gaps and molecular continuity. *Molecular Phylogenetics and Evolution*, 13(1): 67–76.

Ahmad, H., Khan, S. M., Ghafoor, S., & Ali, N. 2009. Ethnobotanical Study of Upper Siran. *Journal of Herbs Spices & Medicinal Plants*, 15: 86–97.

Allard, P. M., Péresse, T., Bisson, J., Gindro, K., Marcourt, L., et al. 2016. Integration of Molecular Networking and In-Silico MS/MS Fragmentation for Natural Products Dereplication. *Analytical Chemistry*, 88(6): 3317–3323.

Álvarez, I. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, 29(3): 417–434.

Anizon, J.-Y., Lamaire, B., & Surbled, M. 2003. Extraction assistée par micro-ondes. *Techniques de l'Ingénieur*, 1–10.

Ansari, M., & Emami, S. 2016. β-Ionone and its analogs as promising anticancer agents. *European Journal of Medicinal Chemistry*, 123: 141–154.

Arnold, M. L. 2006. *Evolution through genetic exchange*. Oxford ; New York: Oxford University Press.

Ayouni, K., Berboucha-rahmani, M., Kim, H. K., Atmani, D., Verpoorte, R., et al. 2016. angustifolia leaf and stem bark extracts. *Industrial Crops & Products*, 1–13.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., et al. 1995. The ITS Region of Nuclear Ribosomal DNA: A Valuable Source of Evidence on Angiosperm Phylogeny. *Annals of the Missouri Botanical Garden*, 82(2): 247.

Ballard, H. E., Sytsma, K. J., & Kowal, R. R. 1999. Shrinking the violets: phylogenetic relationships of infrageneric groups in Viola (Violaceae) based on internal transcribed spacer DNA sequences. *Systematic Botany*, 23(4): 439–458.

Barbeta, B. L., Marshall, A. T., Gillon, A. D., Craik, D. J., & Anderson, M. A. 2008. Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. *Proceedings of the National Academy of Sciences*, 105(4): 1221–1225.

Barry, D. G., Daly, N. L., Clark, R. J., Sando, L., & Craik, D. J. 2003. Linearization of a Naturally Occurring Circular Protein Maintains Structure but Eliminates Hemolytic Activity ·. *Biochemistry*, 42(22): 6688–6695.

Bartle, K. D., & Myers, P. 2002. History of gas chromatography. *Trends in Anlytical Chrmistry*, 21(02): 547–557.

Bedoukian, P. Z. 1982. Violet Fragrance Compounds. *Fragrance Chemistry: The Science of the Sense of Smell* (Academic Press Inc.): 285–315. London: Ernst T. Theimer.

Beille, L. 1909. Famille des Violacées. *Botanique pharmaceutique*, vol. II: 411–415. A. Maloine.

Beirnaert, C., Meysman, P., Vu, T. N., Hermans, N., Apers, S., et al. 2018. speaq 2.0: A complete workflow for high-throughput 1D NMR spectra processing and quantification. (D. Schneidman, Ed.)*PLOS Computational Biology*, 14(3): e1006018.

Benaissi, K. 2013. Le CO2 supercritique appliqué à l'extraction végétale. *Techniques de l'Ingénieur*, 1–16.

Benjamini, Y., & Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.

Bhatt, V. P., & Negi, G. C. S. 2006. Ethnomedicinal plant resources of Jaunsari tribe of Garhwal Himalaya, Uttaranchal. *Indian Journal of Traditional Knowledge*, 5(3): 331–335.

Blomberg, J., Schoenmakers, P. J., Beens, J., & Tijssen, R. 1997. Compehensive two-dimensional gas chromatography (GC$\times$ GC) and its applicability to the characterization of complex (petrochemical) mixtures. *Journal of High Resolution Chromatography*, 20(10): 539–544.

Bonnier, G., & Douin, R. 1990. *La Grande Flore en couleurs*, vol. 3. Belin.

Borges, F., Roleira, F., Milhazes, N., Santana, L., & Uriarte, E. 2005. Simple Coumarins and Analogues in Medicinal Chemistry : Occurrence , Synthesis and Biological Activity. *Current Medicinal Chemistry*, 12: 887–916.

Burgot, G., & Pellerin, F. 2003. Microextraction en phase solide (SPME). *Techniques de l'ingénieur*, 33(0): 0–20.

Burketova, L., Trda, L., Ott, P. G., & Valentova, O. 2015. Bio-based resistance inducers for sustainable plant protection against pathogens. *Biotechnology Advances*, 33(6): 994–1004.

Cabelli, D. E., & Bielski, B. H. J. 1986. STUDIES OF THE REACTIVITY OF TROLOX WITH Mn3+/Fe 3÷ COMPLEXES BY PULSE RADIOLYSIS. *Journal of Free Radicals in Biology & Medicine*, 2: 71–75.

Cai Ke, J. 2007. Violaceae. *Flora of China* (Science Press (Beijing) & Missouri Botanical Garden (St. Louis)), vol. 13: 72–111. Wu Zhengyi, Peter H. Raven & Hong Deyuan.

Calderón-Santiago, M., Fernández-Peralbo, M. A., Priego-Capote, F., & Luque de Castro, M. D. 2016. MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes. *Metabolomics*, 12(3): 1–12.

Cao, J., Yin, C., Qin, Y., & Cheng, Z. 2013. Chemical Constituents of Viola yedoensis and Their Antioxidant Activity. *Chinese Journal of Experimental Traditional Medical Formulae*, 19: 77–81.

Caude, M., & Jardy, A. 2016. *Chromatographie en phase liquide - Théorie et méthodes de séparation Chromatographie en phase liquide Théorie et méthodes de séparation*, 33(0).

Cawoy, H., Mariutto, M., Henry, G., Fisher, C., Vasilyeva, N., et al. 2014. Plant Defense Stimulation by Natural Isolates of *Bacillus* Depends on Efficient Surfactin Production. *Molecular Plant-Microbe Interactions*, 27(2): 87–100.

Cennamo, P., Guacchio, E. D., Jury, S. L., & Caputo, P. 2011. Molecular markers in Viola L. subsect. Viola : Application and taxonomic implications for the identification of dubious herbarium specimens. *Plant Biosystems - An International Journal Dealing with All Aspects of Plant Biology*, 145(2): 306–323.

Chauchard, B., Munzinger, J., Henry, M., & Cedex, N. 2013. L'étude caryologique de deux cultivars de Violettes odorantes remet en cause leur origine taxonomique. *Acta Botanica Gallica*, 150(March 2014): 85–93.

Chen, B., Colgrave, M. L., Daly, N. L., Rosengren, K. J., Gustafson, K. R., et al. 2005. Isolation and Characterization of Novel Cyclotides from *Viola hederaceae*: Solution structure and anti-HIV activity of vhl-1, a leak-specific expressed cyclotide. *Journal of Biological Chemistry*, 280(23): 22395–22405.

Chervin, J., Perio, P., Martins-Froment, N., Pharkeovilay, C., Reybier, K., et al. 2017. Dereplication of natural products from complex extracts by regression analysis and molecular networking: case study of redox-active compounds from Viola alba subsp. dehnhardtii. *Metabolomics*, 13(8). https://doi.org/10.1007/s11306-017-1227-6.

Clausen, J. 1927. Chromosome number and the relationship of species in the genus Viola. *Annals of Botany*, 41(164): 677–714.

Conesa, M. à., Mus, M., & Rosselló, J. A. 2008. Hybridization between insular endemic and widespread species of Viola in non-disturbed environments assessed by nuclear ribosomal and cpDNA sequences. *Plant Systematics and Evolution*, 273(3–4): 169–177.

Costa, R., Fanali, C., Pennazza, G., Tedone, L., Dugo, L., et al. 2015. Screening of volatile compounds composition of white truffle during storage by GCxGC-(FID/MS) and gas sensor array analyses. *LWT - Food Science and Technology*, 60(2): 905–913.

Cotton, J. 2015. *Application de l'analyse métabolomique à la détection ciblée et globale de contaminants organiques dans des matrices agroalimentaires et environnementales par spectrométrie de masse à ultra-haute résolution*. Université Pierre et Marie Curie Paris 6.

Cox, D. G., Oh, J., Keasling, A., Colson, K. L., & Hamann, M. T. 2014a. The Utility of Metabolomics in Natural Product and Biomarker Characterization. *Biochimica et Biophysica Acta*, 1840(12): 3460–3474.

Cox, D. G., Oh, J., Keasling, A., Colson, K. L., & Hamann, M. T. 2014b. The utility of metabolomics in natural product and biomarker characterization. *Biochimica et Biophysica Acta*, 1840(12): 3460–3474.

Craik, D. J. 2012. Host-Defense Activities of Cyclotides. *Toxins*, 4: 139–156.

Cu, J. Q., Perineau, F., & Gaset, A. 1992. Volatile components of violet leaves. *Phytochemistry*, 31(2): 571–573.

de Hoffmann, E., & Stroobant, V. 2005. *Spectrométrie de masse* (3e édition). Dunod.

De Lamarck, J.-B. P. A. de M., & De Candolle, A. P. 1805. *Flore française, ou description succinctes de toutes les plantes qui croissent naturellement en France*, vol. 4. Paris: Desray.

Dong, M. W., & Zhang, K. 2014. Ultra-high-pressure liquid chromatography (UHPLC) in method development. *TrAC Trends in Analytical Chemistry*, 63: 21–30.

Dudareva, N., Klempien, A., Muhlemann, J. K., & Kaplan, I. 2013. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist*, 198: 16–32.

Dunn, W. B., & Ellis, D. I. 2005. Metabolomics: Current analytical platforms and methodologies. *Trends in Analytical Chemistry*, 24(4): 285–294.

Erben, M. 1996. The significance of hybridization on the forming of species in the genus Viola. *Bocconea*, 5(1): 113–118.

Eriksson, L., Trygg, J., & Wold, S. 2008. CV-ANOVA for significance testing of PLS and OPLS® models. *Journal of Chemometrics*, 22(11–12): 594–600.

Favier, A. 2003. *Le stress oxydant*. L'actualité chimique.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4): 783–791.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. 1990. Electropsray ionization-principles and practice. *Mass Spectrometry Reviews*, 9: 32–70.

Fernandez, X., Breme, K., & Varlet, V. 2009. Couplage chromatographie en phase gazeuse / olfactométrie. *Technique de l'Ingénieur*, 33(Base documentaire : Gestion des odeurs et des nuisances olfactives): 26.

Fernandez, X., Cabrol-bass, D., Fernandez, X., & Cabrol-bass, D. 2007. Analyse des arômes. *Technique de l'Ingénieur*, 33(0): 1–22.

Ferrari, G., Donsı, F., & Pataro, G. 2010. Applications of Pulsed Electric Field Treatments for the Enhancement of Mass Transfer from Vegetable Tissue. *Food Eng Rev (2010)*, 2: 109–130.

Ferreira, S. 2004. Doehlert matrix: a chemometric tool for analytical chemistry - review. *Talanta*, 63(4): 1061–1067.

Finkelstein, E., Rosen, G. M., & Rauckman, E. J. 1980. Spin trapping of superoxide and hydroxyl radical: practical aspects. *Archives of Biochemistry and Biophysics*, 200(1): 1–16.

Flamini, G. 2007. Flavonoids and other compounds from the aerial parts of Viola etrusca. *Chemistry & Biodiversity*, 4(2): 139–44.

Flavonoids as antioxidants: determination of radical-scavenging efficiencies. n.d. .

Forman, H. J., & Davies, K. J. A. 2014. How Do Nutritional Antioxidants Really Work: Nucleophilic Tone and Para-Hormesis Versus Free Radical Scavenging in vivo. *Free Radical Biology and Medicine*, 66: 24–35.

Friaa, O., & Brault, D. 2006. Kinetics of the reaction between the antioxidant Trolox and the free radical DPPH • in semi-aqueous solution. *Organic & Biomolecular Chemistry*, 4: 2417–2423.

Gacche, R. N., & Jadhav, S. G. 2012. Antioxidant Activities and Cytotoxicity of Selected Coumarin Derivatives: Preliminary Results of a Structure-Activity Relationship Study Using Computational Tools. *Journal of Experimental and Clinical Medicine*, 4(3): 165–169.

Gerlach, S. L., Rathinakumar, R., Chakravarty, G., Göransson, U., Wimley, W. C., et al. 2010. Anticancer and chemosensitizing abilities of cycloviolacin O2 from Viola odorata and psyle cyclotides from Psychotria leptothyrsa. *Biopolymers*, 94(5): 617–625.

Goff, S. A. 2006. Plant Volatile Compounds: Sensory Cues for Health and Nutritional Value? *Science*, 311(5762): 815–819.

Goudable, J., & Favier, A. 1997. Radicaux libres oxyénés et antioxydants. *Nutur Clin Métabol*, 11: 115–120.

Gran, L., Sandberg, F., & Sletten, K. 2000. Oldenlandia affinis (R&S) DC A plant containing uteroactive peptides used in African traditional medicine. *Journal of Ethnopharmacology*, 70: 197–203.

Grapov, D., Wanichthanarak, K., & Fiehn, O. 2015a. MetaMapR: pathway independent metabolomic network analysis incorporating unknowns. *Bioinformatics*, 31(16): 2757–2760.

Grapov, D., Wanichthanarak, K., & Fiehn, O. 2015b. MetaMapR : pathway independent metabolomic network analysis incorporating unknowns. *Bioinformatics*, 31(16): 2757–2760.

Grasse, M.-C., Ellena, J.-C., & Peyron, L. 2009a. Survol historique de la culture des Violettes et de leur valorisation à Tourettes-sur-Loup. *Un jour, une plante : Les violettes en parfumerie*.

Grasse, M.-C., Ellena, J.-C., & Peyron, L. 2009b. Utilisation des feuilles de violette à Grasse. *Un jour, une plante : Les violettes en parfumerie*.

Gruber, C. W., Čemažar, M., Anderson, M. A., & Craik, D. J. 2007. Insecticidal plant cyclotides and related cystine knot toxins. *Toxicon*, 49(4): 561–575.

Gustafson, K. R., McKee, T. C., & Bokesch, H. R. 2004. Anti-HIV cyclotides. *Current Protein and Peptide Science*, 5: 331–340.

Haleng, J., Pincemail, J., Defraigne, J.-O., Charlier, C., & Chapelle, J.-P. 2007. Le stress oxydant. *Revue Médicale de Liège*, 62(10): 628–38.

Hammami, I., Kamoun, N., & Rebai, A. 2011. Biocontrol of Botrytis cinerea with essential oil and methanol extract of Viola odorata L . flowers. *Archive of Applied Science Research*, 3(5): 44–51.

Hanem Fathy, K. 2012. Prospect of Botanical Biopesticides in Insect Pest Management. *Pharmacologia*, 3(12): 641–656.

Harbourne, N., Marete, E., Jacquier, J. C., & O'Riordan, D. 2009. Effect of drying methods on the phenolic constituents of meadowsweet (Filipendula ulmaria) and willow (Salix alba). *LWT - Food Science and Technology*, 42(9): 1468–1473.

Harman, D. 1992. Free radical theory of aging. *Mutation Research/DNAging*, 275(3–6): 257–266.

Hasegawa, M., Kishino, H., & Yano, T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22(2): 160–174.

Hashempour, H., Koehbach, J., Daly, N. L., Ghassempour, A., & Gruber, C. W. 2013. Characterizing circular peptides in mixtures: sequence fragment assembly of cyclotides from a violet plant by MALDI-TOF/TOF mass spectrometry. *Amino Acids*, 44(2): 581–595.

Hekking, W. H. A. (Ed.). 1988. Violaceae ; P. 1. Rinorea and Rinoreocarpus. *Flora Neotropica*: 1–207. New York: New York Botanical Garden.

Herrmann, A., Burman, R., Mylne, J. S., Karlsson, G., Gullbo, J., et al. 2008. The alpine violet, Viola biflora, is a rich source of cyclotides with potent cytotoxicity. *Phytochemistry*, 69: 939–952.

Hubert, J., Berger, M., & Dayde, J. 2008. Effects of fermentation on the phytochemical composition and antioxidant properties of soy germ. *Food Chemistry 109*, 109: 709–721.

Ibrahim, R. M. 2017. Immunological, Cytogenetic and Hepatoprotective Effect of Viola odorata Methanolic Extract on Methotrexate Albino Male Mice. *Journal of Biotechnology Research Center*, 11(2): 47–55.

Ireland, David C., Colgrave, M. L., & Craik, D. J. 2006. A novel suite of cyclotides from *Viola odorata* : sequence variation and the implications for structure, function and stability. *Biochemical Journal*, 400(1): 1–12.

Ireland, David C., Colgrave, M. L., Nguyencong, P., Daly, N. L., & Craik, D. J. 2006. Discovery and Characterization of a Linear Cyclotide from Viola odorata: Implications for the Processing of Circular Proteins. *Journal of Molecular Biology*, 357: 1522–1535.

James W., B., Mark W., C., Maarten J. M., C., Michael F., F., Walter S., J., et al. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, 181(1): 1–20.

Jennings, C., West, J., Waine, C., Craik, D., & Anderson, M. 2001. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from Oldenlandia affinis. *Proceedings of the National Academy of Sciences*, 98(19): 10614–10619.

Jeong, Y. H., Oh, Y.-C., Cho, W.-K., Shin, H., Lee, K. Y., et al. 2016. Anti-inflammatory effects of Viola yedoensis and the application of cell extraction methods for investigating bioactive constituents in macrophages. *BMC Complementary and Alternative Medicine*, 16: 1–16.

Jukes, T. H., & Cantor, C. R. 1969. Evolution of protein molecules. *Mammalian Protein Metabolism* (Academic Press), vol. III: 21–132. Munro H. N.

Jung, M., Geiger, H., Zinsmeistert, H. D., & Botanik, F. 1995. *Tri- and tetrahydroxycoumarin derivatives from tetraphis pellucida**, 39(2): 379–381.

Kanchanapoom, T. 2007. Aromatic diglycosides from Cladogynos orientalis. *Phytochemistry*, 68(5): 692–696.

Kanti Das, T., Wati, M. R., & Fatima-Shad, K. 2014. Oxidative Stress Gated by Fenton and Haber Weiss Reactions and Its Association With Alzheimer's Disease. *Archives of Neuroscience*, 2(3): 1–8.

Katajamaa, M., Miettinen, J., & Ores, M. 2006. MZmine : toolbox for processing and visualization of mass spectrometry based molecular profile data ˇ ic. *Bioinformatics*, 22(5): 634–636.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16: 111–120.

Kind, T., & Fiehn, O. 2016. Phytochemistry Letters Strategies for dereplication of natural compounds using high-resolution tandem mass spectrometry. *Phytochemistry Letters*. https://doi.org/10.1016/j.phytol.2016.11.006.

Kostova, I. 2006. Synthetic and Natural Coumarins as Antioxidants. *Mini-Reviews in Medicinal Chemistry*, 6: 365–374.

Le Floch, F., Tena, M. T., R??os, A., & Valc??rcel, M. 1998. Supercritical fluid extraction of phenol compounds from olive leaves. *Talanta*, 46(5): 1123–1130.

Lee, S.-W., Chen, Z.-T., & Chen, C.-M. 1993. The 4α-Methylsterol Violasterol A from Viola Formosana Hayata. *Journal of the Chinese Chemical Society*, 40(3): 305–307.

Leopoldini, M., Marino, T., Russo, N., & Toscano, M. 2004. Antioxidant Properties of Phenolic Compounds: H-Atom versus Electron Transfer Mechanism. *The Journal of Physical Chemistry A*, 108(22): 4916–4922.

Levin, D. A., Francisco-Ortega, J., & Jansen, R. K. 1996. Hybridization and the extinction of rare plant species. *Conservation Biology*, 10(1): 10–16.

Liochev, S. I., & Fridovich, I. 2002. The Haber-Weiss cycle—70 years later: an alternative view. *Redox Report*, 7(1): 55–57.

Little, J. R., & McKinney, L. E. 2016. Violaceae Batsch. *Flora of North America*, vol. 6. Oxford University Press.

Liza, M. S., Abdul Rahman, R., Mandana, B., Jinap, S., Rahmat, A., et al. 2012. Supercritical fluid extraction of bioactive flavonoid from Strobilanthes crispus (Pecah kaca) and its comparison with solvent extraction. *International Food Research Journal*, 19(2): 503–508.

López-alarcón, C., & Denicola, A. 2013a. Evaluating the antioxidant capacity of natural products : A review on chemical and cellular-based assays. *Analytica Chimica Acta*, 763: 1–10.

López-alarcón, C., & Denicola, A. 2013b. Evaluating the antioxidant capacity of natural products : A review on chemical and cellular-based assays. *Analytica Chimica Acta*, 763: 1–10.

Luzzatto-Knaan, T., Melnik, A. V., & Dorrestein, P. C. 2015. Mass spectrometry tools and workflows for revealing microbial chemistry. *Analyst*, 140(15): 4949–4966.

Mahieu, N. G., & Patti, G. J. 2017. Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. *Analytical Chemistry*, 89(19): 10397–10406.

Malécot, V., Marcussen, T., Munziger, J., Yockteng, R., & Henry, M. 2007a. On the origin of the sweet-smelling Parma violet cultivars (Violaceae): wide intraspecific hybridization, sterility, and sexual reproduction. *American Journal of Botany*, 94(1): 29–41.

Malécot, V., Marcussen, T., Munziger, J., Yockteng, R., & Henry, M. 2007b. On the origin of the sweet-smelling Parma violet cultivars (Violaceae): wide intraspecific hybridization, sterility, and sexual reproduction. *American Journal of Botany*, 94(1): 29–41.

Malien-Aubert, C., Dangles, O., & Amiot, M. J. 2001. Color Stability of Commercial Anthocyanin-Based Extracts in Relation to the Phenolic Composition. Protective Effects by Intra- and Intermolecular Copigmentation. *Journal of Agricultural and Food Chemistry*, 49(1): 170–176.

Mandal, V., Mohan, Y., & Hamalatha, S. 2007. Microwave Assisted Extraction - An Innovative and Promising Extraction Tool for Medicinal Plant Research. *Pharmacognosy Reviews*, 1(1): 1–18.

Marcussen, T., Karlsson, T., Jonsell, B., & Wind, P. 2010. Violaceae. *Flora Nordica*, vol. 6: 12–52. Stockholm: The Bergius Foundation.

Marcussen, T., Yousefi, N., & Mehrvarz, S. S. 2012. Anatomical studies on selected species of Viola (Violaceae). *Nordic Journal of Botany*, 30(4): 461–469.

Martucci, M. E. P., De Vos, R. C. H., Carollo, C. A., & Gobbo-Neto, L. 2014. Metabolomics as a Potential Chemotaxonomical Tool: Application in the Genus Vernonia Schreb. (J. Chamani, Ed.)*PLoS ONE*, 9(4): e93149.

Marzouk, M. M., Elkhateeb, A., Ibrahim, L. F., Sameh, R., & Kawashty, S. A. 2012. Two Cytotoxic Coumarin Glycosides from the Aerial Parts of Diceratella elliptica (DC .) Jonsell Growing in Egypt. *Rec. Nat. Prod.*, 3: 237–241.

Masrani, K. V., Rama, H. S., & Bafna, S. L. 1974. Ultraviolet absorption spectra: Some substituted coumarins. *Journal of Applied Chemistry and Biotechnology*, 24(6): 331–341.

Mereda JR., P., Hodalova, I., & Kucera, J. 2011. Genetic and morphological variation in Viola suavis s.l. (Violaceae) in the western Balkan Peninsula : two endemic subspecies revealed. *Systematics and Biodiversity*, 9: 211–231.

Michener, C. D., & Sokal, R. R. 1957. A quantitative approach to a problem in classification. *Evolution*, 11(2): 130–162.

Milman, B. L., & Zhurkovich, I. K. 2016. Mass spectral libraries: A statistical review of the visible use. *TrAC Trends in Analytical Chemistry*, 80: 636–640.

Molnár, P., & Szabolcs, J. 1980. Occurrence of 15-cis-violaxanthin in Viola tricolor. *Phytochemistry*, 19(4): 623–627.

Moon, H.-I., Jung, J.-C., & Lee, J. 2007. Antiplasmodial activity of triterpenoid isolated from whole plants of Viola genus from South Korea. *Parasitology Research*, 100(3): 641–644.

Muhammad, N., Saeed, M., Awen, A., & Khan, H. 2012. Ethnomedicinal , phytochemical and pharmacological profile of genus Viola. *Phytopharmacology*, 3(1): 214–226.

Mullis, K., Faloona, F., Scharf, S., Saiki, R. K., Horn, G. T., et al. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor symposia on quantitative biology*, 51: 263–273. Cold Spring Harbor Laboratory Press.

Murray, A. P., Nacional, U., Rodriguez, S., & Nacional, U. 2015. Chemical Constituents and Biological Activities of Plants from the Genus Viola. *Chemistry & Biodiversity*, 12: 1777–1808.

Naruya, S., & Masatoshi, N. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4): 406–425.

Nejia, H. 2013. *Extraction et purification de substances naturelles : comparaison de l'extraction au CO2-supercritique et des techniques conventionnelles*. Université de Toulouse.

Nepveu, F. 2008. *APPRT- Région Midi-Pyrénées VALorisation d ' Agro-Produits : Identification de Marqueurs spécIfiques par leurs Propriétés redox Porteur : UPS*: 1–39. Document interne, Toulouse: Université Paul Sabatier.

Nguyen, H. T., Vo, B. T., Nguyen, L. H., Bernad, J., Alaeddine, M., et al. 2013. Extracts of Crinum latifolium inhibit the cell viability of mouse lymph oma cell line EL4 and induce activation of anti-tumour activity of macrophages in vitro. *Journal of Ethnopharmacology*, 149(1): 75–83.

Niederbacher, B., Winkler, J. B., & Schnitzler, J. P. 2015. Volatile organic compounds as non-invasive markers for plant phenotyping. *Journal of Experimental Botany*, 66(18): 5403–5416.

Nkhili, E., Tomao, V., El Hajji, H., El Boustani, E.-S., Chemat, F., et al. 2009. Microwave-assisted water extraction of green tea polyphenols. *Phytochemical Analysis*, 20(5): 408–415.

Ohashi, Y., & Ohshima, M. 1992. Stress-induced expression of genes for pathogenesis-related proteins in plants. *Plant and Cell Physiology*, 33(7): 819–826.

Overy, S. A. 2004. Application of metabolite profiling to the identification of traits in a population of tomato introgression lines. *Journal of Experimental Botany*, 56(410): 287–296.

Özel, M. Z., Göğüş, F., & Lewis, A. C. 2006. Comparison of direct thermal desorption with water distillation and superheated water extraction for the analysis of volatile components of Rosa damascena Mill. using GCxGC-TOF/MS. *Analytica Chimica Acta*, 566(2): 172–177.

Papet, Y., Brunet, B., & Mura, P. 2010. Headspace (HS) et micro-extraction en phase solide (SPME). Théorie et applications. *Annales de Toxicologie Analytique*, 22(2): 75–79.

Parshad, B., Duraisamy, A. J., Saini, S., Yadav, P., Vats, P., et al. 2016. Synthesis and SAR Study of Antioxidant Potential of Polyhydroxy Coumarin Derivatives. *Medicinal Chemistry*, 6(7): 506–514.

Parsley, N. C., Kirkpatrick, C. L., Crittenden, C. M., Rad, J. G., Hoskin, D. W., et al. 2018. PepSAVI-MS reveals anticancer and antifungal cycloviolacins in Viola odorata. *Phytochemistry*, 152: 61–70.

Perez de Souza, L., Naake, T., Tohge, T., & Fernie, A. R. 2017. From chromatogram to analyte to metabolite. How to pick horses for courses from the massive web resources for mass spectral plant metabolomics. *GigaScience*, 6(7): 1–20.

Pichersky, E. 2006. Biosynthesis of Plant Volatiles: Nature's Diversity and Ingenuity. *Science*, 311(5762): 808–811.

Pranting, M., Loov, C., Burman, R., Goransson, U., & Andersson, D. I. 2010. The cyclotide cycloviolacin O2 from Viola odorata has potent bactericidal activity against Gram-negative bacteria. *Journal of Antimicrobial Chemotherapy*, 65(9): 1964–1971.

Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., et al. 1987. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*, 238(4825): 336–341.

Procházkoá, D., Boušová, I., & Wilhelmová, N. 2011. Antioxidant and prooxidant properties of flavonoids. *Fitoterapia*, 82(4): 513–523.

Prost, I., Dhondt, S., Rothe, G., Vicente, J., Rodriguez, M. J., et al. 2005. Evaluation of the Antimicrobial Activities of Plant Oxylipins Supports Their Involvement in Defense against Pathogens. *Plant Physiology*, 139(4): 1902–1913.

Pruitt, K. D. 2004. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, 33(Database issue): D501–D504.

Qin, Y., Yin, C., & Cheng, Z. 2013. A New Tetrahydrofuran Lignan Diglycoside from Viola tianshanica Maxim. *Molecules*, 18(12): 13636–13644.

Quideau, S., Deffieux, D., Douat-Casassus, C., & Pouységu, L. 2011. Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. *Angewandte Chemie International Edition*, 50(3): 586–621.

Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J. 2006. OPLS discriminant analysis : combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics*, 20: 341–351.

Ribeiro, S. M., Porto, W. F., Silva, O. N., Santos, M. de O., Dias, S. C., et al. 2013. Plant Antifungal Peptides. *Handbook of Biologically Active Peptides*: 169–179. Elsevier.

Rodrigues, V. H., Cardoso, C., Paixão, J. A., & Costa, M. M. R. R. 2007. 2,2,6,6-Tetramethyl-4-oxopiperidinium nitrate. *Acta Crystallographica Section C Crystal Structure Communications*, 63(4): 243–245.

Rouadjia, I. 2009, October. *Infos Matériaux : Le "champ des odeurs."* Matériauthèque Ensad.

Ruzicka, L., Schinz, H., & Susz, B.-P. 1944. Veilchenriechstoffe.(14. Mitteilung). Zur Stereoisomerie von Hexen-(3)-ol-(1), Nonadien-(2, 6)-al-(1) und Nonadien-(2, 6)-ol-(1). *Helvetica Chimica Acta*, 27(1): 1561–1569.

Saint-lary, L., Debrauwer, M. L., Gougeon, M. R., Paris, M. J., Thomas, M. O., et al. 2015. *Evaluation de l'approche métabolomique pour l'authentification des extraits naturels utilisés dans le secteur arômes et parfums*. Université Nice Sophia-Antipolis.

Saldaña, G., Álvarez, I., Condón, S., & Raso, J. 2014. Microbiological Aspects Related to the Feasibility of PEF Technology for Food Pasteurization. *Critical Reviews in Food Science and Nutrition*, 54(11): 1415–1426.

Salem, M. A., Jüppner, J., Bajdzienko, K., & Giavalisco, P. 2016. Protocol: a fast, comprehensive and reproducible one-step extraction method for the rapid preparation of polar and semi-polar metabolites, lipids, proteins, starch and cell wall polymers from a single sample. *Plant Methods*, 12: 1–15.

Sanger, F., & Coulson, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94: 441–448.

Sangster, T., Major, H., Plumb, R., Wilson, A. J., & Wilson, I. D. 2006. A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis. *The Analyst*, 131(10): 1075.

Segovia, F. J., Luengo, E., Corral-Pérez, J. J., Raso, J., & Almajano, M. P. 2015. Improvements in the aqueous extraction of polyphenols from borage (Borago officinalis L.) leaves by pulsed electric fields: Pulsed electric fields (PEF) applications. *Industrial Crops and Products*, 65: 390–396.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., et al. 2003a. Cytoscape : A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, 13: 2498–2504.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., et al. 2003b. *Cytoscape : A Software Environment for Integrated Models of Biomolecular Interaction Networks*, (Karp 2001): 2498–2504.

Shen, X.-Y., & Xie, C. 2009. Study on Antimicrobial Activity of Viola tianshanica Maxim.Extracts. *Biotechnology*, 19: 71–73.

Sies, H. 1985. Oxidative Stress : Introductory Remarks. *Oxidative stress*.

Skowyra, M., Calvo, M. I., Gallego, M. G., Azman, N. A. M., & Almajano, M. P. 2014. Characterization of Phytochemicals in Petals of Different Colours From Viola × wittrockiana Gams. and Their Correlation With Antioxidant Activity. *Journal of Agricultural Science*, 6(9): 93–105.

Smedsgaard, J., & Nielsen, J. 2005. Metabolite profiling of fungi and yeast: from phenotype to metabolome by MS and informatics. *Journal of Experimental Botany*, 56(410): 273–286.

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., & Siuzdak, G. 2006. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Analytical Chemistry*, 78(3): 779–787.

Sun, Y., Du, L., Zhou, L., Miao, F., Yang, X., et al. 2011. Study on antibacterial active components from Viola yedoensis. *Zhongguo Zhong Yao Za Zhi.*, 36(19): 2666–2671.

Swofford, D. L. 2001. *PAUP*: Phylogenetic Analysis Using Parsimony (and other methods) 4.0.b5*. Sinauer Associates, Sunderland, Massachusetts.

Tholl, D., Boland, W., Hansel, A., Loreto, F., Röse, U. S. R., et al. 2006. Practical approaches to plant volatile analysis. *The Plant Journal*, 45(4): 540–560.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., et al. 2016. Hybridization and extinction. *Evolutionary Applications*, 9(7): 892–908.

Trabi, M. 2004. Tissue-Specific Expression of Head-to-Tail Cyclized Miniproteins in Violaceae and Structure Determination of the Root Cyclotide Viola hederacea root cyclotide1. *The Plant Cell*, 16(8): 2204–2216.

Tsuagawa, H. et al. 2015. MS-DIAL: Data Independent MS/MS Deconvolution for Comprehensive Metabolome Analysis. *Nat Methods*, 12(6): 523–526.

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., et al. 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature Methods*, 12: 523–526.

Tsugawa, H., Kind, T., Nakabayashi, R., Yukihira, D., Tanaka, W., et al. 2016a. Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. *Analytical Chemistry*. https://doi.org/10.1021/acs.analchem.6b00770.

Tsugawa, H., Kind, T., Nakabayashi, R., Yukihira, D., Tanaka, W., et al. 2016b. Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. *Analytical Chemistry*. https://doi.org/10.1021/acs.analchem.6b00770.

Tutin, T. G., Moore, D. M., Valentine, D. H., Walters, S. M., & Webb, D. A. 1968. Violaceae. *Flora Europaea*, vol. 2: 270–275. Cambridge University Press.

Uhde, G., & Ohloff, G. 1972. Parmon, eine phantomverbindung im veilchenblütenöl. *Helvetica Chimica Acta*, 55(7): 2621–2625.

van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J. 2006. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(1): 142.

Vercammen, J., Sandra*, P., Baltussen, E., Sandra, T., & David, F. 2000. Considerations on Static and Dynamic Sorptive and Adsorptive Sampling to Monitor Volatiles Emitted by Living Plants. *Journal of High Resolution Chromatography*, 23(9): 547–553.

Vergnes, S., Ladouce, N., Fournier, S., Ferhout, H., Attia, F., et al. 2014. Foliar treatments with Gaultheria procumbens essential oil induce defense responses and resistance against a fungal pathogen in Arabidopsis. *Frontiers in Plant Science*, 5(477): 1–8.

Villaverde, J. J., Sandín-España, P., Sevilla-Morán, B., López-Goti, C., & Alonso-Prados, J. L. 2016. Biopesticides from Natural Products: Current Development, Legislative Framework, and Future Trends. *BioResources*, 11(2). https://doi.org/10.15376/biores.11.2.Villaverde.

Vukics, V. 2009. *Antioxidant Flavonoid Glycosides in Viola Tricolor L.* SEMMELWEIS UNIVERSITY.

Vukics, V., Kery, A., Bonn, G. K., & Guttman, A. 2008. Major flavonoid components of heartsease (Viola tricolor L.) and their antioxidant activities. *Analytical and Bioanalytical Chemistry*, 390(7): 1917–1925.

Vukics, V., Ringer, T., Kery, A., Bonn, G. K., & Guttman, A. 2008. *Analysis of heartsease (Viola tricolor L .) flavonoid glycosides by micro-liquid chromatography coupled to multistage mass spectrometry*, 1206: 11–20.

Wang, C. K., Colgrave, M. L., Gustafson, K. R., Ireland, D. C., Goransson, U., et al. 2007. Anti-HIV cyclotides from the Chinese medicinal herb Viola yedoensis. *Journal of Natural Products*, 71: 47–52.

Wardencki, W., Michulec, M., & Curylo, J. 2004. A review of theoretical and practical aspects of solid-phase microextraction in food analysis. *International Journal of Food Science and Technology*, 39(7): 703–717.

Wätjen, W., Michels, G., Steffan, B., Niering, P., Chovolou, Y., et al. 2005. Low Concentrations of Flavonoids Are Protective in Rat H4IIE Cells Whereas High Concentrations Cause DNA Damage and Apoptosis. *The Journal of Nutrition*, 135(3): 525–531.

Watson, J. D., Gilman, M., Witkowski, J., & Zoller, M. 1994a. La réaction de polymérisation en chaîne. *ADN Recombinant*, 33(0): 79–98.

Watson, J. D., Gilman, M., Witkowski, J., & Zoller, M. 1994b. La réaction de polymérisation en chaîne. *ADN Recombinant*, 33(0): 79–98.

White, T. J., Bruns, T., Lee, S., & Taylor, J. 1990. Amplification and Direct Sequencing of Fungal Ribosomal Rna Genes for Phylogenetics. *PCR Protocols*, (January 2016): 315–322.

Willmer, P. 2011. The Timing and Patterning of Flowering. *Pollination and Floral Ecology*: 483–502. Princeton: Princeton University Press.

Witkowska-Banaszczak, E., Bylka, W., Matławska, I., Goślińska, O., & Muszyński, Z. 2005. Antimicrobial activity of Viola tricolor herb. *Fitoterapia*, 76(5): 458–61.

Wodehouse, R. P. 1935. *Pollen grains*. Mcgraw-Hill Book Company, Inc; New York; London.

Wolfender, J., Marti, G., & Queiroz, E. F. 2010. Advances in Techniques for Profiling Crude Extracts and for the Rapid Identification of Natural Products : Dereplication , Quality Control and Metabolomics. *Current Organic Chemistry*, 14: 1808–1832.

Wolfender, J.-L., Marti, G., Thomas, A., & Bertrand, S. 2015a. Current approaches and challenges for the metabolite profiling of complex natural extracts. *Journal of Chromatography A*, 1382: 136–164.

Wolfender, J.-L., Marti, G., Thomas, A., & Bertrand, S. 2015b. Current approaches and challenges for the metabolite profiling of complex natural extracts. *Journal of Chromatography A*, 1382: 136–164.

Wolfender, J.-L., Queiroz, E. F., & Hostettmann, K. 2006. The importance of hyphenated techniques in the discovery of new lead compounds from nature. *Expert Opinion on Drug Discovery*, 1(3): 237–260.

Worley, B., & Powers, R. 2013. Multivariate Analysis in Metabolomics. *Current Metabolomics*, 1(1): 92–107.

Xiao, J. F., Zhou, B., & Ressom, H. W. 2012. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *TrAC - Trends in Analytical Chemistry*, 32: 1–14.

Xie, Chen, Veitch, N. C., Houghton, P. J., & Simmonds, M. S. J. 2003. Flavone C-glycosides from Viola yedoensis MAKINO. *Chemical & Pharmaceutical Bulletin*, 51(10): 1204–7.

Xie, C., Veitch, N. C., Houghton, P. J., & Simmonds, M. S. J. 2003. Flavone C-Glycosides from Viola yedoensis Makino. *Chem. Pharm. Bull.*, 51(10): 1204–1207.

Yang, J., Guo, J., & Yuan, J. 2008. In vitro antioxidant properties of rutin. *LWT - Food Science and Technology*, 41(6): 1060–1066.

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., Bruns, N., Glukhov, E., Wodtke, A., Felicio, R. D., et al. 2013. Molecular Networking as a Dereplication Strategy. *Journal of Natural Products*, 76: 1686–1699.

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., Bruns, N., Glukhov, E., Wodtke, A., Felicio, R. De, et al. 2013. Molecular Networking as a Dereplication Strategy. *Journal of Natural Products*, 76: 1686–1699.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution*, 39(3): 306–314.

Yeshak, M. Y., Burman, R., Eriksson, C., & Göransson, U. 2012. Optimization of cyclotide extraction parameters. *Phytochemistry Letters*, 5: 776–781.

Yockteng, R., Jr, H. E. B., Mansion, G., Dajoz, I., & Nadot, S. 2003. *Relationships among pansies (Viola section Melanium) investigated using ITS and ISSR markers*, 241: 153–170.

Yoo, K.-O., & Jang, S.-K. 2010. Infrageneric relationships of Korean Viola based on eight chloroplast markers. *Journal of Systematics and Evolution*, 48(6): 474–481.

Zhang, B., Huang, Y., Zhang, Q., Liu, X., Li, F., et al. 2014. Fragrance discrimination of Chinese Cymbidium species and cultivars using an electronic nose. *Scientia Horticulturae*, 172: 271–277.

Zhou, H. Y., Hong, J. L., Shu, P., Ni, Y. J., & Qin, M. J. 2009. A new dicoumarin and anticoagulant activity from Viola yedoensis Makino. *Fitoterapia*, 80(5): 283–285.

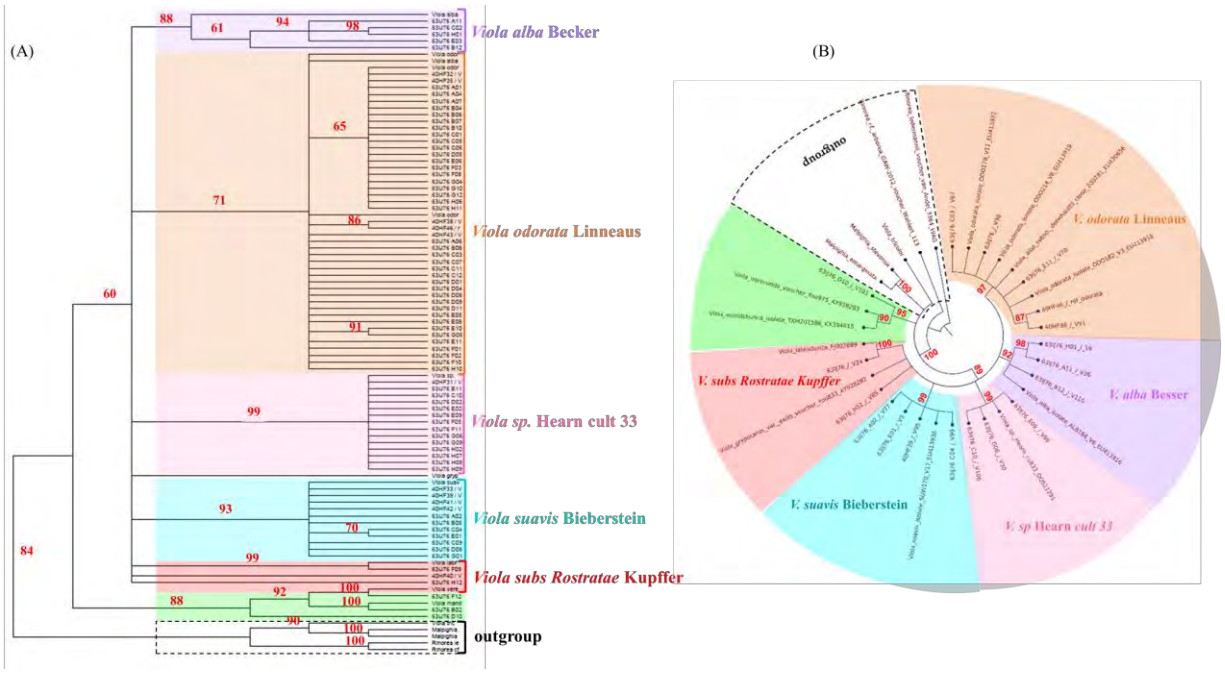


Fig. 1 Cladograms obtained with maximum parsimony (A) and maximum likelihood (B) method showing genetic relationship of selected experimental sequences and references based on ITS sequences. To improve the quality of the figure, only sequences of a few characteristic samples from each clade observed with distance model were selected for B. The outgroup is used to root the tree. Bootstrap values to study robustness are indicated in red (in colour).

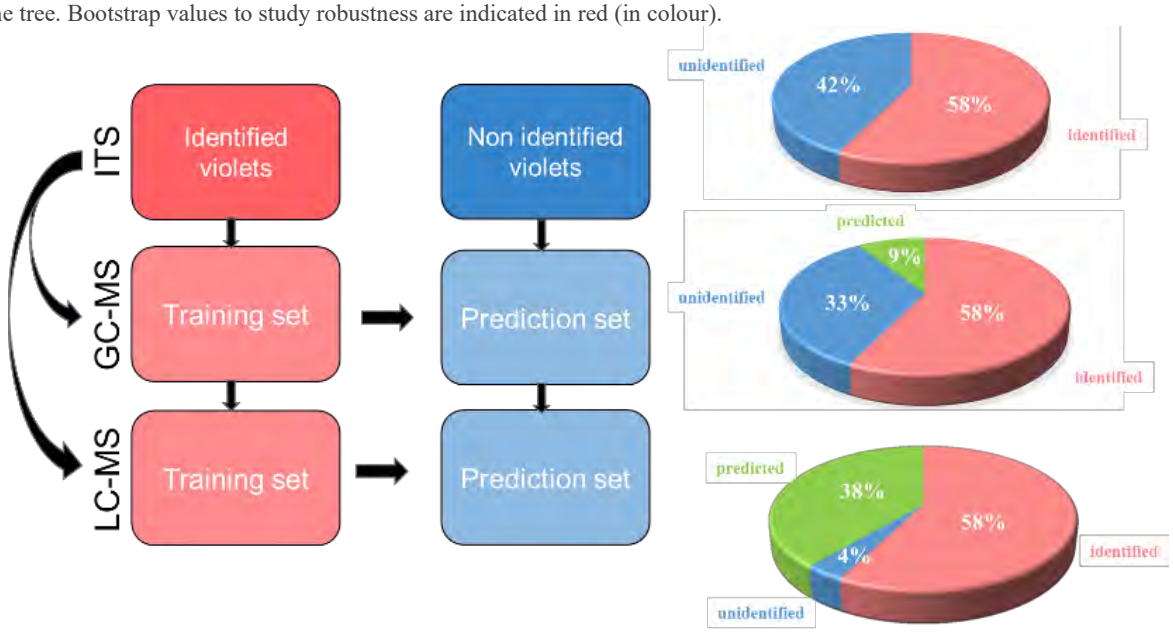


Fig. 2 Strategy of classification based on training and prediction sets. Pie charts show the distribution of results of classification (in colour).

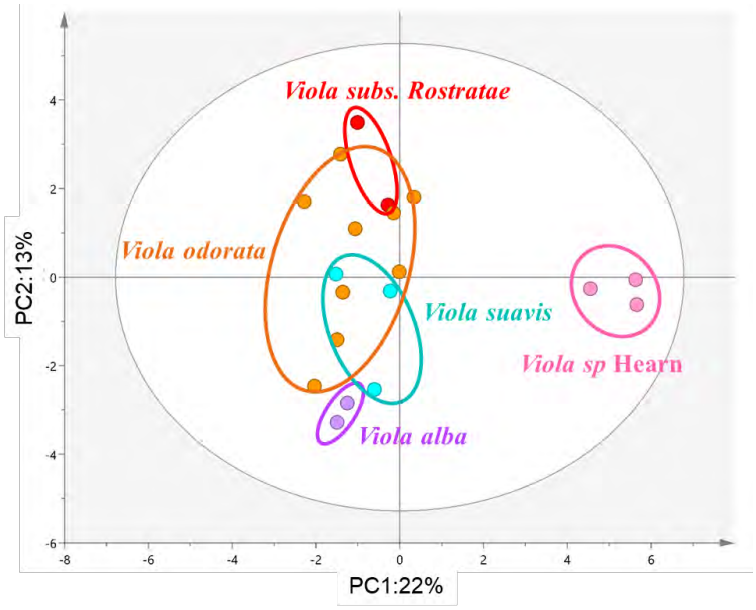


Fig. 3 PCA plot of EI data of GC-MS flower extracts with binary data (in colour).

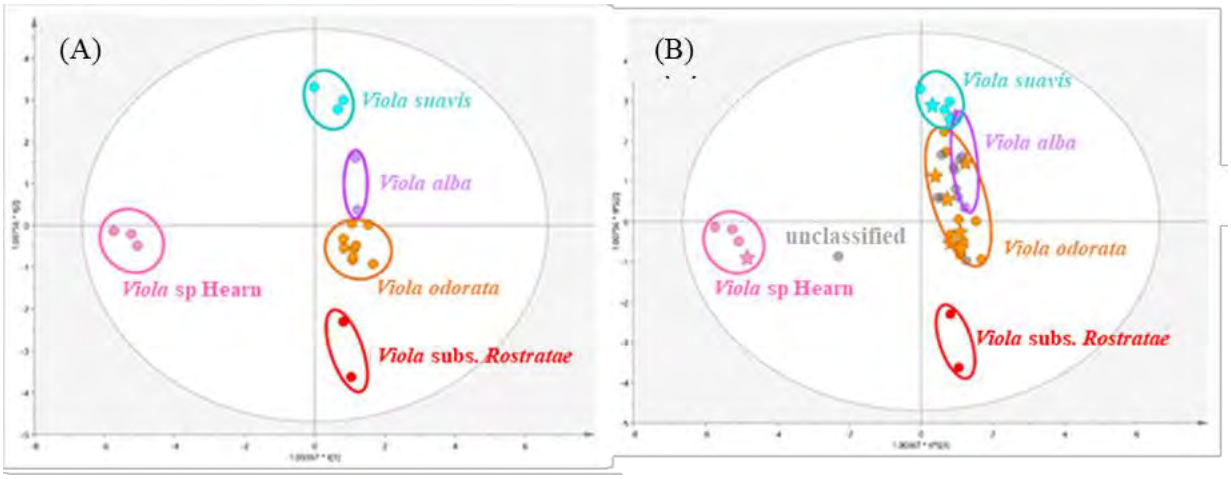


Fig. 4 OPLS-DA based model (A) and prediction set of unidentified violets (B). Well-projected violets are indicated by a star (in colour).

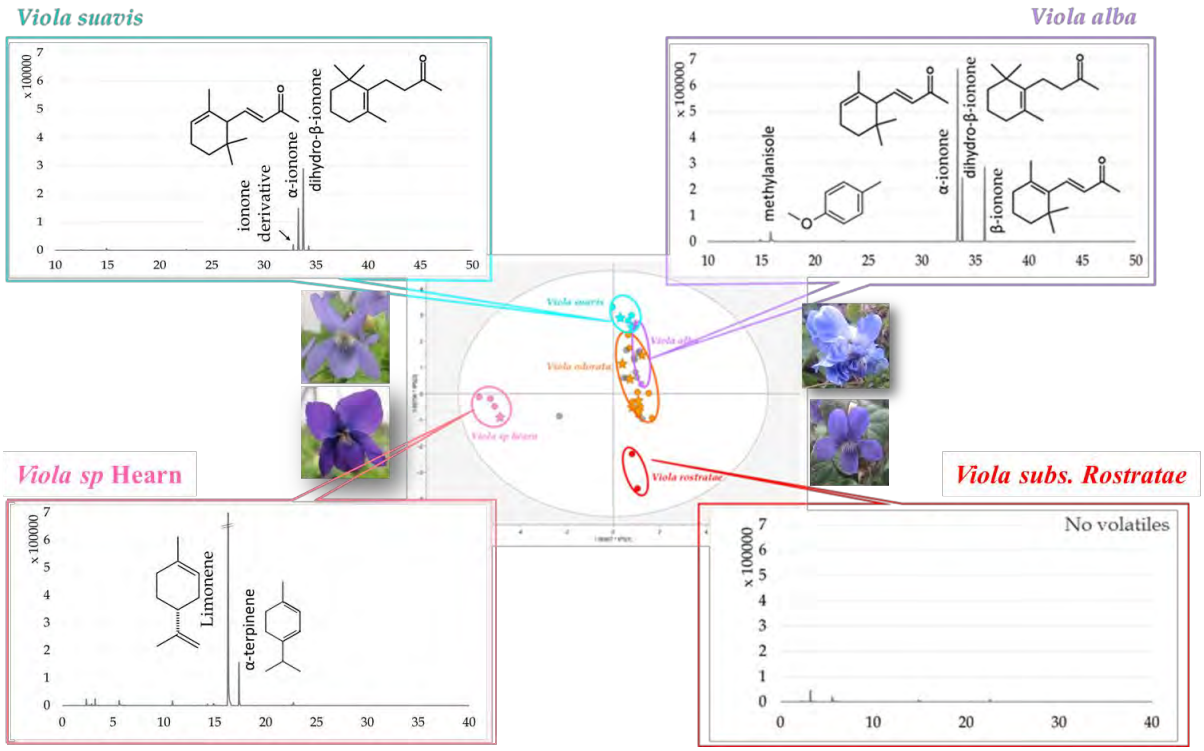


Fig. 5 Highlighting of specific volatile profiles with identification of characteristic volatile biomarkers. Stars indicate well-projected violets (in colour).

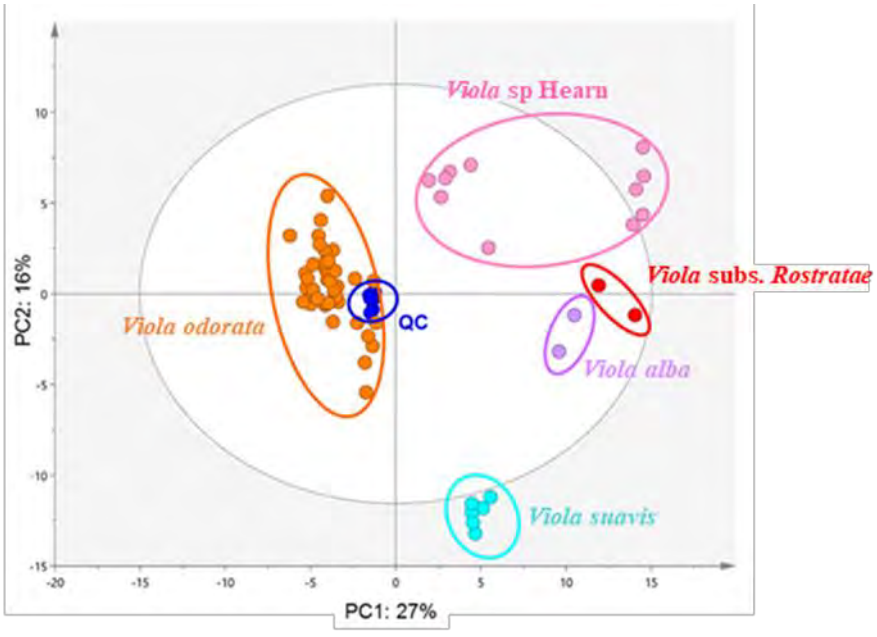


Fig. 6 PCA-X plot of ESI-NI/PI combined dataset of UHPLC-HRMS leaf extracts of violets (in colour).

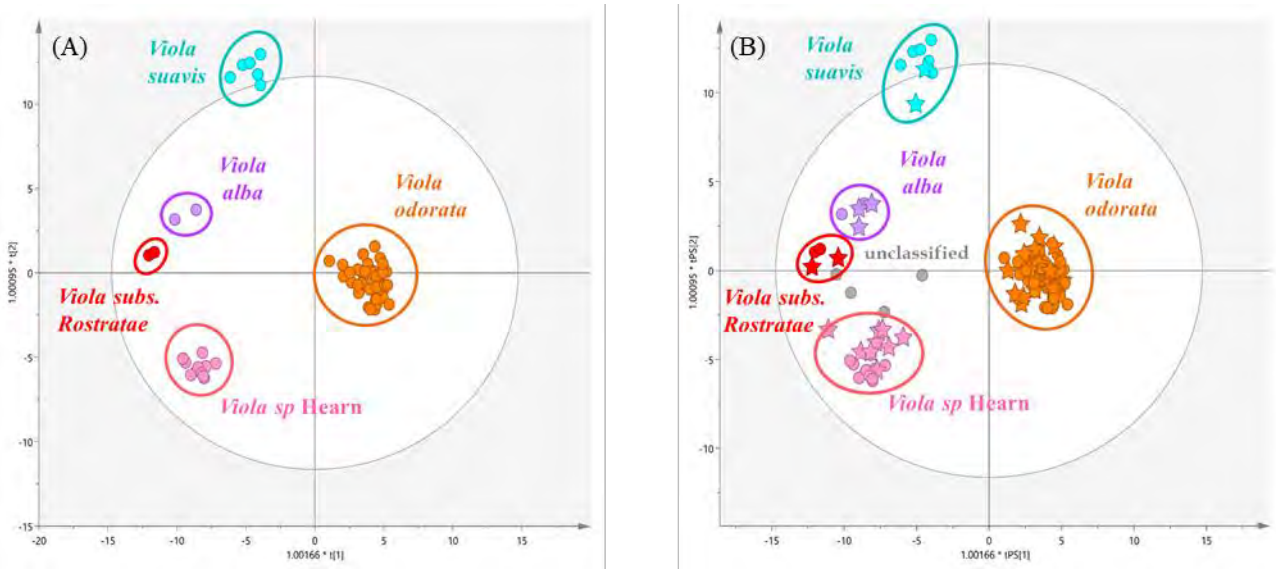


Fig. 7 OPLS-DA based model (A) and prediction set of unidentified violets (B) taking into account 95% of the data. Well projected violets are indicated by a star (in colour).

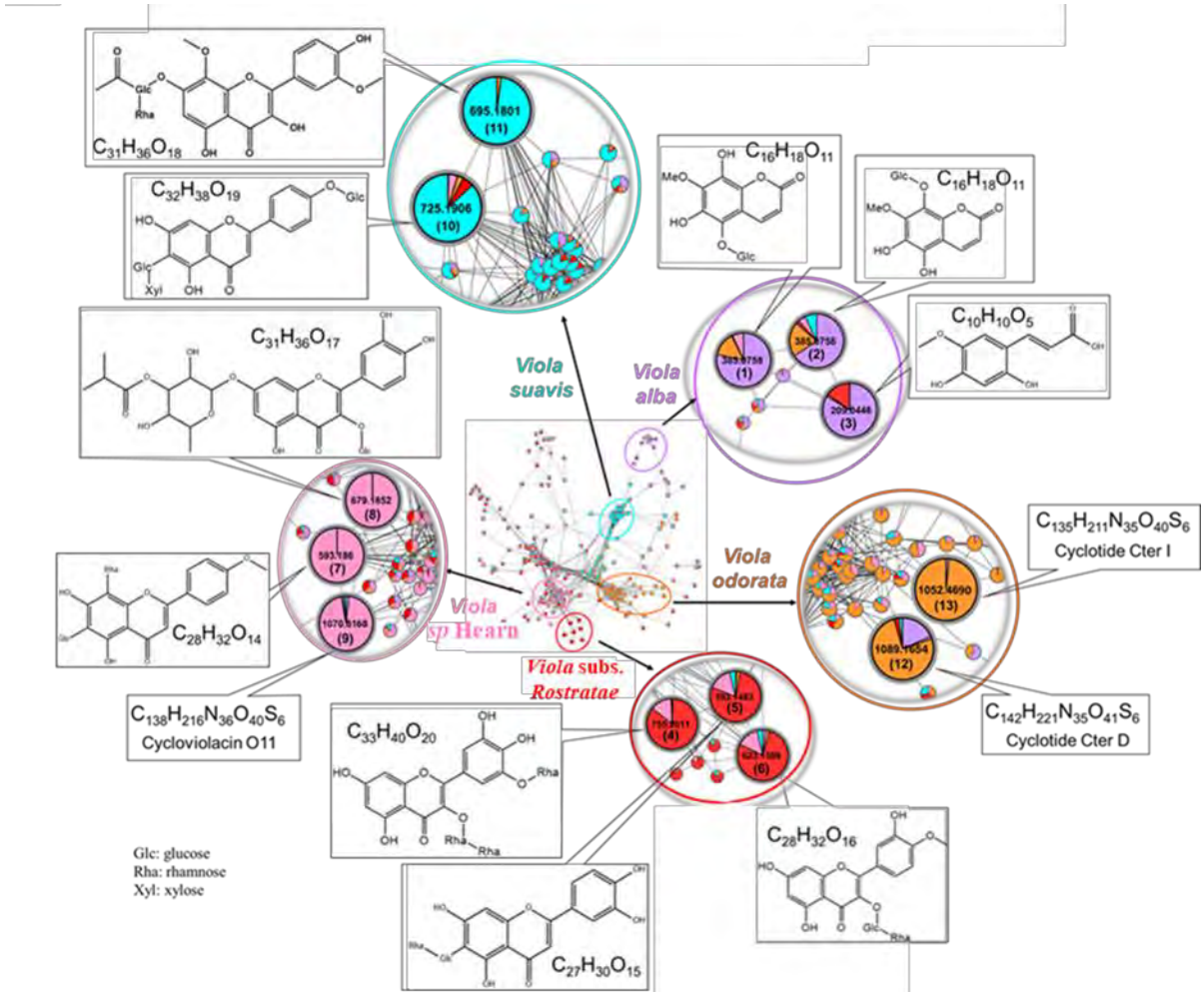


Fig. 8 Correlation network based on ESI-NI/PI combined dataset of UHPLC-HRMS performed on leaf extracts and identification of characteristic biomarkers. Color tag is based on species identification by genetic analysis. Node size was emphasized based on OPLS-DA coefficient value. Putative structural identification was based on HRMS and MS/MS spectra (in colour).

Table 1: Summary of all dereplicated compounds with MS-FINDER

N°	m/z	RT (min)	Mode	Molecular Formula	Error (ppm)	Putative id[a]	Chemical class	OPLS coefficient	Cluster Affiliation
1	385.0758	2.95	$[M-H]^-$	$C_{16}H_{18}O_{11}$	4.6744	7-Methoxy-5,6,8-dihydroxycoumarin-5-β-glucopyranoside[b]*	Coumarin	0.0486768	*alba*
2	385.0758	3.47	$[M-H]^-$	$C_{16}H_{18}O_{11}$	4. 6744	7-Methoxy-5,6,8-trihydroxycoumarin-8-β-glucopyranoside[b]*	Coumarin	0.040651	*alba*
3	209.0446	3.32	$[M-H]^-$	$C_{10}H_{10}O_5$	4.3053	3-(2,4,5-Trihydroxyphenyl)-2-propenoic acid; 5-Me ether*	Simple Aromatic	0.0474672	*alba*
4	755.2011	3.92	$[M-H]^-$	$C_{33}H_{40}O_{20}$	3.8625	3,3',4',5,5',7-Hexahydroxyflavone; 3-*O*-[α -L-Rhamnopyranosyl-(1,2)-α -L-rhamnopyranoside], 3'-*O*-α -L-rhamnopyranoside*	Flavonoid	0.0938658	*subs Rostratae*
5	593.1483	4.46	$[M-H]^-$	$C_{27}H_{30}O_{15}$	4.8891	6-C-β-D-Glucopyranosyl-3',4',5,7-tetrahydroxyflavone; 2''-O-a-L-Rhamnopyranosyl*	Flavonoid	0.0619192	*subs Rostratae*
6	623.1589	4.63	$[M-H]^-$	$C_{28}H_{32}O_{16}$	4.4932	Isorhamnetin 3-neohesperidoside*	Flavonoid	0.0909114	*subs Rostratae*
7	593.1860	5.42	$[M+H]^+$	$C_{28}H_{32}O_{14}$	0.8429	3,5,7-Trihydroxyflavone; 7-Me ether, 3-*O*-[α-L-rhamnopyranosyl-(1→6)-β -D-glucopyranoside]*	Flavonoid	0.0825076	*sp* Hearn
8	679.1852	5.71	$[M-H]^-$	$C_{31}H_{36}O_{17}$	3.8281	Quercetin 3,7-diglycosides; 3-*O*-β-D-Glucopyranoside, 7-*O*-[2-	Flavonoid	0.0587476	*sp* Hearn

						methylpropanoyl-(→3)-α-L-rhamnopyranoside]*			
9	1070.8168 803.3678	5.94	$[M+3H]^{3+}$ $[M+4H]^{4+}$	$C_{138}H_{216}N_{36}O_{40}S_6$	1.0273	Cycloviolacin O11*	Cyclotide	0.0574069	*sp* Hearn
10	725.1906	3.31	$[M-H]^-$	$C_{32}H_{38}O_{19}$	3.8610	6-C-β-D-Glucopyranosyl-4',5,7-trihydroxyflavone; 2"-*O*-α-L-Arabinopyranosyl, 4'-*O*-β-D-glucopyranoside*	Flavonoid	0.0430546	*suavis*
11	695.1801	3.52	$[M-H]^-$	$C_{31}H_{36}O_{18}$	4.0277	3,4',5,7-Tetrahydroxy-3',8-dimethoxyflavone; 7-*O*-[α-L-Rhamnopyranosyl-(1→2)-6-*O*-acetyl-β-D-glucopyranoside*	Flavonoid	0.0571469	*suavis*
12	1089.1653 817.1260	5.9	$[M+3H]^{3+}$ $[M+4H]^{4+}$	$C_{142}H_{221}N_{35}O_{41}S_6$	3.4889	Cter D**	Cyclotide	0.0403016	*odorata*
13	1052.4690	5.42	$[M+3H]^{3+}$	$C_{135}H_{211}N_{35}O_{40}S_6$	1.4252	Cter I**	Cyclotide	0.0533443	*odorata*
14	1047.1356 785.6131	5,17	$[M+3H]^{3+}$ $[M+4H]^{4+}$	$C_{133}H_{207}N_{37}O_{39}S_6$	4,8704	Cycloviolacin O9*	Cyclotide	-	*odorata*
15	1057.7927	5,31	$[M+3H]^{3+}$	-	-	-	Cyclotide	-	*odorata*

[a] Top ranked hit determined by *in silico* MS/MS fragmentation with MS-FINDER

[b] Compounds confirmed by NMR (Chervin and al. 2017)

* Hits found when interrogating the in-house database formed using the DNP composed of known compounds in *Viola* and derivatives

** Hits found when interrogating general natural product databases