

## TECHNICAL ADVANCE

# Computational metabolomics reveals overlooked chemodiversity of alkaloid scaffolds in *Piper fimbriulatum*

Tito Damiani<sup>1,†</sup> , Joshua Smith<sup>1,2,†</sup>, Téo Hebra<sup>1,†</sup> , Milana Perković<sup>1,3,†</sup>, Marijo Čičak<sup>1,3</sup>, Alžběta Kadlecová<sup>1</sup>, Vlastimil Rybka<sup>4</sup>, Martin Dráčinský<sup>1</sup> and Tomáš Pluskal<sup>1,\*</sup> 

<sup>1</sup>Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo náměstí 542/2, 160 00 Prague, Czech Republic,

<sup>2</sup>First Faculty of Medicine Charles University, Kateřinská 1660/32, 121 08 Prague, Czech Republic,

<sup>3</sup>University of Chemistry and Technology, Technická 5, 166 28 Prague, Czech Republic, and

<sup>4</sup>Prague Botanical Garden, Trojská 800/196, 171 00 Prague, Czech Republic

Received 11 December 2024; revised 15 February 2025; accepted 24 February 2025.

\*For correspondence (e-mail [tomas.pluskal@uochb.cas.cz](mailto:tomas.pluskal@uochb.cas.cz)).

†These authors contributed equally to this work.

## SUMMARY

Plant specialized metabolites play key roles in diverse physiological processes and ecological interactions. Identifying structurally novel metabolites, as well as discovering known compounds in new species, is often crucial for answering broader biological questions. The *Piper* genus (*Piperaceae* family) is known for its special phytochemistry and has been extensively studied over the past decades. Here, we investigated the alkaloid diversity of *Piper fimbriulatum*, a myrmecophytic plant native to Central America, using a metabolomics workflow that combines untargeted LC–MS/MS analysis with a range of recently developed computational tools. Specifically, we leverage open MS/MS spectral libraries and metabolomics data repositories for metabolite annotation, guiding isolation efforts toward structurally new compounds (i.e., dereplication). As a result, we identified several alkaloids belonging to five different classes and isolated one novel seco-benzylisoquinoline alkaloid featuring a linear quaternary amine moiety which we named fimbriulatamine. Notably, many of the identified compounds were never reported in *Piperaceae* plants. Our findings expand the known alkaloid diversity of this family and demonstrate the value of revisiting well-studied plant families using state-of-the-art computational metabolomics workflows to uncover previously overlooked chemodiversity. To contextualize our findings within a broader biological context, we employed a workflow for automated mining of literature reports of the identified alkaloid scaffolds and mapped the results onto the angiosperm tree of life. By doing so, we highlight the remarkable alkaloid diversity within the *Piper* genus and provide a framework for generating hypotheses on the biosynthetic evolution of these specialized metabolites. Many of the computational tools and data resources used in this study remain underutilized within the plant science community. This manuscript demonstrates their potential through a practical application and aims to promote broader accessibility to untargeted metabolomics approaches.

**Keywords:** *Piper fimbriulatum*, *Piperaceae*, alkaloids, mass spectrometry, computational metabolomics, Wikidata, angiosperms, technical advance.

## INTRODUCTION

*Piperaceae* ("pepper") is a family of pan-tropical flowering plants and includes two of the largest genera of angiosperms, *Piper* (>2400 species) and *Peperomia* (>1400 species) (Simmonds et al., 2021). Numerous *Piperaceae* plants are of significant economic importance due to their widespread use as spices (e.g., black pepper) and in traditional

medicine (e.g., Ayurveda) (Salehi et al., 2019). In particular, the *Piper* genus is known for its special phytochemistry, and its phytochemical investigation over the past four decades has led to the isolation of over 300 different amide alkaloids (often referred to as "piperamides") with a wide range of biological activities (Gómez-Calvario & Ríos, 2019; Gutierrez et al., 2013; Parmar et al., 1997; Salehi

et al., 2019). Despite such extensive exploration of the *Piper* phytochemistry, new piperamide structures are still regularly reported every year (Jung et al., 2024; Zhou et al., 2024). *Piper fimbriulatum* (Figure S1) is a myrmecophytic plant native to Central America that lives in symbiosis with *Pheidole bicornis* ants (Mayer et al., 2008). Because of this symbiosis, the phytochemical characterization of this plant has focused mostly on its volatile compounds (terpenoids in particular), believed to play a role in the plant–insect communication (Mayer et al., 2008; Mundina et al., 1998). During a plant metabolomics screening campaign (Jarmusch et al., 2022), we found evidence for *P. fimbriulatum* leaves to contain a high concentration of piperlongumine, a *Piper* alkaloid that has exhibited remarkable anticancer properties at a preclinical level (Conde et al., 2021). This finding, combined with the lack of reports of alkaloids in this plant, motivated a deeper investigation of the nonvolatile chemodiversity of *P. fimbriulatum*, where the presence of unique specialized metabolites might have been overlooked.

Over the past two decades, liquid chromatography coupled to high-resolution tandem mass spectrometry (LC–MS/MS) has become the method of choice for metabolite profiling of crude plant extracts (Tsugawa et al., 2021). Such popularity is owed to the high sensitivity and throughput of modern instruments as well as the ability of tandem mass spectrometry data (i.e., MS/MS) to provide sufficient structural information for the annotation of the detected metabolites (Jarmusch et al., 2021). In addition, recently developed computational tools allow for efficient exploration of the metabolomics data and greatly assist in the identification of new natural product analogs and scaffolds (Beniddir et al., 2021; Wolfender et al., 2019). Major advances in this direction were stimulated by the introduction of molecular networking (Watrous et al., 2012), which can capture structural relationships among (unknown) metabolites based on their fragmentation spectra (MS/MS or MS2) as well as the development of computational platforms to facilitate the (re)usability of open data (e.g., reference MS/MS libraries, metabolomics data sets) (Gomes et al., 2024; Wang et al., 2016). Finally, the recent establishment of open access resources for cataloging information available for known natural products (e.g., taxonomic occurrence, reported bioactivity), such as the LOTUS initiative (Rutz et al., 2022), allows for quick and reproducible mining of scientific literature, which can significantly aid data interpretation and dereplication (i.e., avoiding re-isolation of previously known compounds). Despite their potential, many of these computational tools remain underutilized within the plant science community, partly due to their recent introduction and, in our opinion, the perceived complexity for researchers outside the metabolomics field. The present manuscript showcases their application for the identification of specialized metabolites

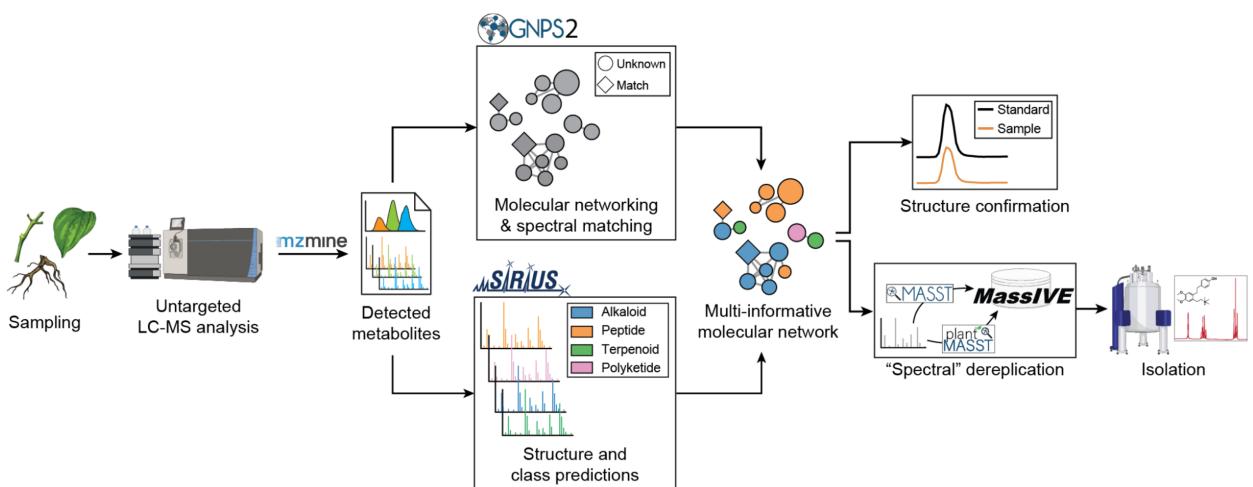
in non-model plant species, an often crucial task for addressing broader biological questions, and aims to promote broader accessibility to untargeted metabolomics approaches.

In this study, we used untargeted LC–MS/MS analysis to investigate the metabolite profiles of different organs of *P. fimbriulatum*. We primarily focused on alkaloids given the relevance of this compound class in the *Piper* genus. We used a range of computational tools to explore the (detected) chemodiversity of *P. fimbriulatum* and leveraged open MS/MS spectral libraries and metabolomics data repositories to direct our isolation efforts toward structurally new compounds (Figure 1). We identified several alkaloids belonging to five different alkaloid classes, many of which have never been reported in Piperaceae. Moreover, we isolated one novel seco-benzylisoquinoline alkaloid with a linear quaternary amine moiety, which is an unusual moiety for plant alkaloids. Finally, we mined the Wikidata framework and compared the occurrence of these alkaloid classes across angiosperm orders and families. Our findings uncovered a remarkable diversity of alkaloid scaffolds being produced by the *Piper* genus, which goes beyond the well-known piperamides. We believe that the workflows presented here can be applied to numerous biochemical and physiological studies requiring a deep and effective phytochemical characterization of specialized metabolism in non-model species.

## RESULTS AND DISCUSSION

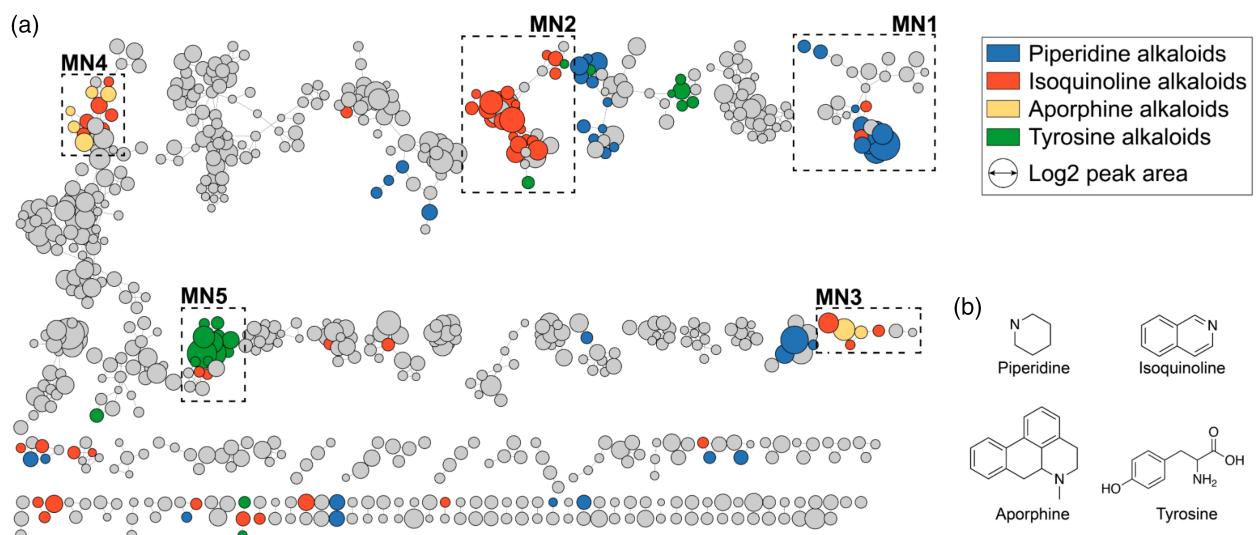
### LC–MS analysis of *P. fimbriulatum*

In the present study, we used untargeted metabolomics to explore the chemodiversity of *P. fimbriulatum* (Figure S1) with a particular focus on alkaloids. After confirming the taxonomic identity of a collected specimen of *P. fimbriulatum* using DNA barcoding (see **Experimental procedures** section), we sampled the plant leaves, stems, and root organs. From the collected samples, we prepared crude water–ethanol extracts and analyzed them using untargeted LC–MS/MS. The resulting raw data were processed using a pipeline of computational tools (Figure 1). We used the mzmine software for feature detection in order to convert the complex raw spectral data into a list of detected metabolites (i.e., LC–MS features) (Heuckerth et al., 2024). Next, we used feature-based molecular networking (FBMN) (Nothias et al., 2020) to group together unknown metabolites with similar chemical structures based on the similarity between their MS/MS spectra. Finally, we annotated (unknown) metabolites in the obtained molecular network using a combination of MS/MS spectral matching, *in silico* prediction of chemical structures and compound classes using CSI:FingerID (Dührkop et al., 2015) and CANOPUS (Dührkop et al., 2021), with manual inspection of the prediction results (see **Metabolite annotation** section). FBMN



**Figure 1.** Analytical and computational workflow used in the present study.

Different organs of the plant were sampled and analyzed by untargeted LC-MS/MS. Computational tools were used to aid the exploration of the detected chemical space, while publicly available spectral libraries and data repositories were used to focus isolation efforts toward structurally novel phytochemicals.



**Figure 2.** (a) A global molecular network of the chemical space detected in *Piper fimbriatum* (leaf, stem, and root organs). Nodes predicted as alkaloids by CANOPUS are colored based on the alkaloid class. Alkaloid-related molecular families are highlighted as MN1-5. (b) Chemical scaffolds corresponding to the alkaloid classes highlighted in the global molecular network.

resulted in the generation of 54 molecular networks (MNs) from 898 features (Figure 2), out of which only a small fraction (~6%) could be annotated via direct match against the GNPS MS/MS spectral library. In order to aid the exploration of the detected chemical space, we layered the chemical structure and compound class predictions over the global molecular network (Mutabdžija et al., 2024). Overall, 245 nodes in the network were predicted as alkaloids and clustered into five main molecular networks (Figure 2, hereafter referred to as **MN1**, **MN2**, **MN3**, **MN4**, and **MN5**), on which we focused our annotation efforts on.

#### *Piperidine alkaloids molecular network (MN1)*

MN1 contained 29 nodes, out of which 12 were predicted as “piperidine alkaloids” (Figure 2a), a class of lysine-derived alkaloids characterized by the presence of a piperidine ring (Figure 2b) in their chemical structures (Szóke et al., 2013). Spectral matching against the GNPS MS/MS library retrieved a highly confident hit (cosine similarity  $>0.96$ , Figure S2a) to piperlongumine (**1**), in accordance with our screening campaign results (see **Introduction** section). Moreover, several other spectra in MN1 were annotated as piperlongumine analogs (i.e., a

similar MS/MS spectrum, but different precursor  $m/z$ ) with high confidence (cosine similarity  $>0.93$ ). We confirmed the annotation of piperlongumine using a commercial standard (Figure S3a) and used this confirmation as a starting point to “propagate” the annotation throughout MN1 (Note S1; see also Metabolite annotation section). By doing so, we putatively annotated five piperlongumine analogs ((2), (3), (4), (5), (6); see Table S1) and one homodimer (7). Next, we synthesized pure standards (see Experimental procedures section) for compounds (2), (3), (5), (6) and confirmed our annotations by retention time and MS/MS spectral matching (Figure S3b–e). Regarding compound (7), [2 + 2] cycloaddition reactions are known to occur under visible and UV light (Hurtley et al., 2014; Nguyen & Al-Mourabit, 2016). Therefore, we generated a mixture of piperlongumine dimers by irradiation of the pure monomer with 365 nm UV light (see Experimental procedures section). LC–MS/MS analysis of the reaction mixture confirmed the presence of (7) in the native plant (Note S1; Figure S7a,b). Interestingly, we observed a nearly identical chromatographic profile between the reaction mixture and the leaf extract (Figure S7c). This suggests that the formation of these dimers in the plant might be UV-induced, rather than enzyme-catalyzed (Note S1). Finally, spectral matching retrieved a highly confident hit (cosine similarity  $>0.99$ , Figure S2b) to piperine (8), the main alkaloid in *Piper nigrum* fruits, which we also confirmed using a commercial standard (Figure S3f). Interestingly, piperine was detected exclusively in the root organ, unlike the other piperidine alkaloids (Figure 3c). Piperine and other piperamides were recently reported to localize within specific cells of the fruit perisperm and the root cortex in *P. nigrum* (Jäckel et al., 2022). Given that *P. fimbriatum* belongs to the same genus, it is reasonable to expect a similar phenomenon. However, our data allow us to assess only the interorgan distribution, as the analysis was performed on bulk samples (i.e., leaf, stem, and root). Investigating intertissue and cellular localization of these alkaloids would require spatial metabolomics techniques with tissue- or cellular-level resolution (e.g., single-cell metabolomics or laser microdissection), which we plan to pursue in future studies.

#### *Isoquinoline alkaloids molecular network (MN2)*

MN2 contained 34 nodes, out of which 12 were predicted as “isoquinoline alkaloids,” which are a large and

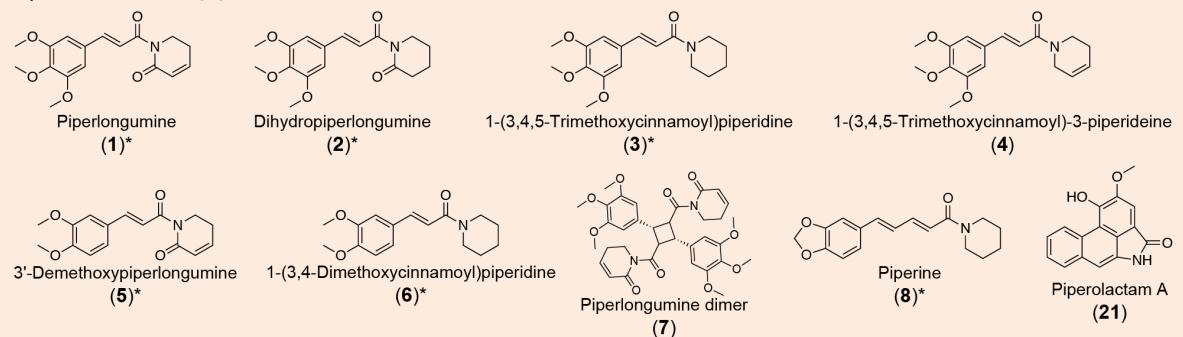
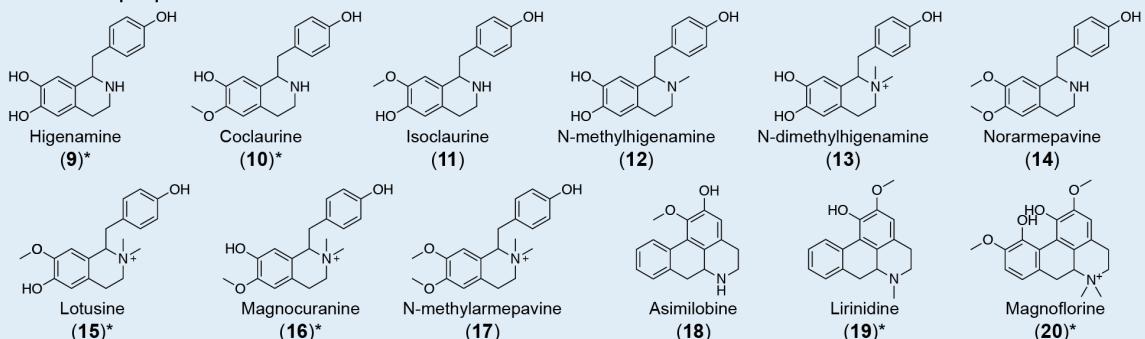
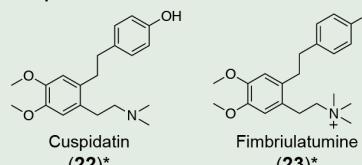
structurally diverse group of specialized metabolites (~2500 known structures) predominantly found in the plant kingdom (Yang et al., 2024). Isoquinoline alkaloids are further classified into benzylisoquinolines (BIAs), aporphines, and other subclasses based on their structural backbone, and are known to possess potent pharmacological properties (e.g., the narcotic morphine and the antimicrobial berberine) (Yang et al., 2024). Spectral matching against the GNPS MS/MS library retrieved highly confident hits (cosine similarity  $>0.95$ ) to higenamine (9) and coclaurine (10) (Figure S2c,d), two central intermediates in the biosynthesis of all plant BIAs (Tian et al., 2024). We confirmed both these annotations using commercial standards (Figure S8a,b) and, similar to what was done for MN1, we used these confirmations to “propagate” the annotation throughout MN2. Overall, we annotated 9 BIAs (Table S1), including *N*-methylarmepavine (17) and its intermediates, and confirmed 4 of these annotations using commercial standards (Figure S8a–d; Note S2). Overall, BIAs showed higher accumulation in the leaves, unlike piperamides, which were more often found consistently across all organs (Figure 3c).

#### *Aporphine and aristolactams alkaloids molecular networks (MN3 and MN4)*

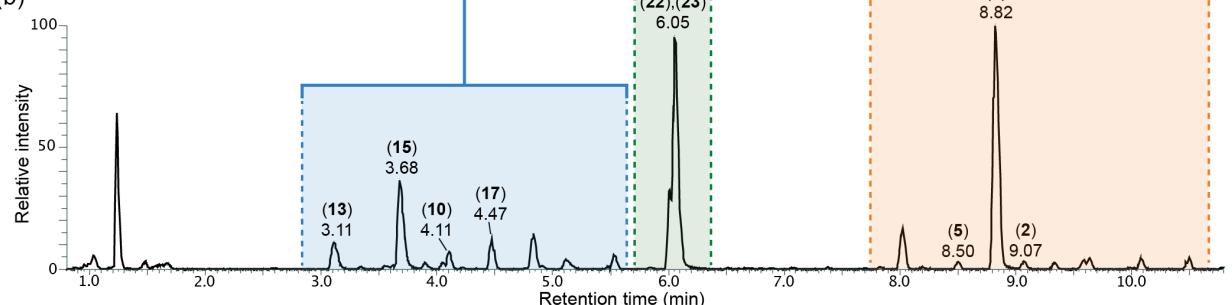
MN3 and MN4 contained various nodes predicted as either “isoquinoline” or “aporphine alkaloids” by CANOPUS. Aporphine alkaloids are a major class of BIAs, including more than 700 reported compounds (da Silva Mendes et al., 2023). Many aporphine alkaloids have shown promising therapeutic effects, particularly for the treatment of central nervous system diseases (Carbone et al., 2019) and metabolic syndromes (Wang et al., 2021). Spectral matching against the GNPS MS/MS library retrieved hits to asimilobine (18), magnoflorine (20), and piperolactam A (21) (Figure S2e–g). Piperolactams (a.k.a. aristolactams) are a relatively small group of phytochemicals with a variety of reported pharmacological activities, mainly occurring in plants from the Annonaceae, Aristolochiaceae, and Piperaceae families (Kumar et al., 2004). Similar to what was done for MN1 and MN2, we used commercial standards (when available) and manual inspection of the MS/MS spectra to confirm these annotations (Note S3). Overall, we confirmed the annotation for lirinidine (19) and magnoflorine (20) using commercial standards (Figure S8e,f) and putatively annotated asimilobine (18) and piperolactam A

**Figure 3.** (a) Chemical structures of the *Piper fimbriatum* alkaloids identified in the present study. Compounds confirmed by retention time match with commercial standard or NMR structural characterization are marked with asterisks. More information about the identified compounds are provided in Table S1. (b) Base peak LC–MS chromatogram of *P. fimbriatum* leaf acquired in positive ionization mode. (c) Heatmap representing the abundance (shown as log<sub>10</sub>-transformed LC–MS peak area) of the annotated alkaloids across the different plant organs (i.e., leaf, stem, root). It must be noted that the heatmap aims at highlighting the interorgan distribution of individual metabolites. Since absolute quantification for each compound was not performed, abundance comparisons between metabolites [e.g., compound (1) more abundant in the plant than compound (15)] cannot be made due to differences in ionization efficiency and other factors.

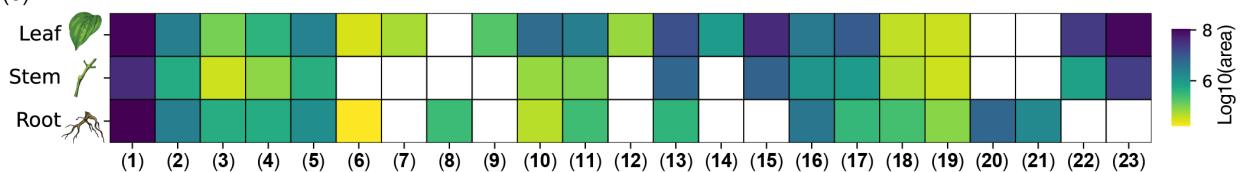
(a)

**Piperamides and piperolactam****BIAs and aporphine alkaloids****Cuspidatin and fimbriulatumine**

(b)



(c)



(21) based on MS/MS spectral matching against the GNPS MS/MS library and manual inspection of the results. Interestingly, all these alkaloids were almost exclusively detected in the root organ of the plant (Figure 3c). Such organ-specific production may result from the localized expression of biosynthetic enzymes, which could be leveraged in future studies to elucidate the biosynthetic pathway(s) for these alkaloids. In fact, organ-specific metabolite abundance can help narrow down candidate genes for experimental validation, focusing on those with similar expression patterns (see, e.g., Schnabel et al., 2021).

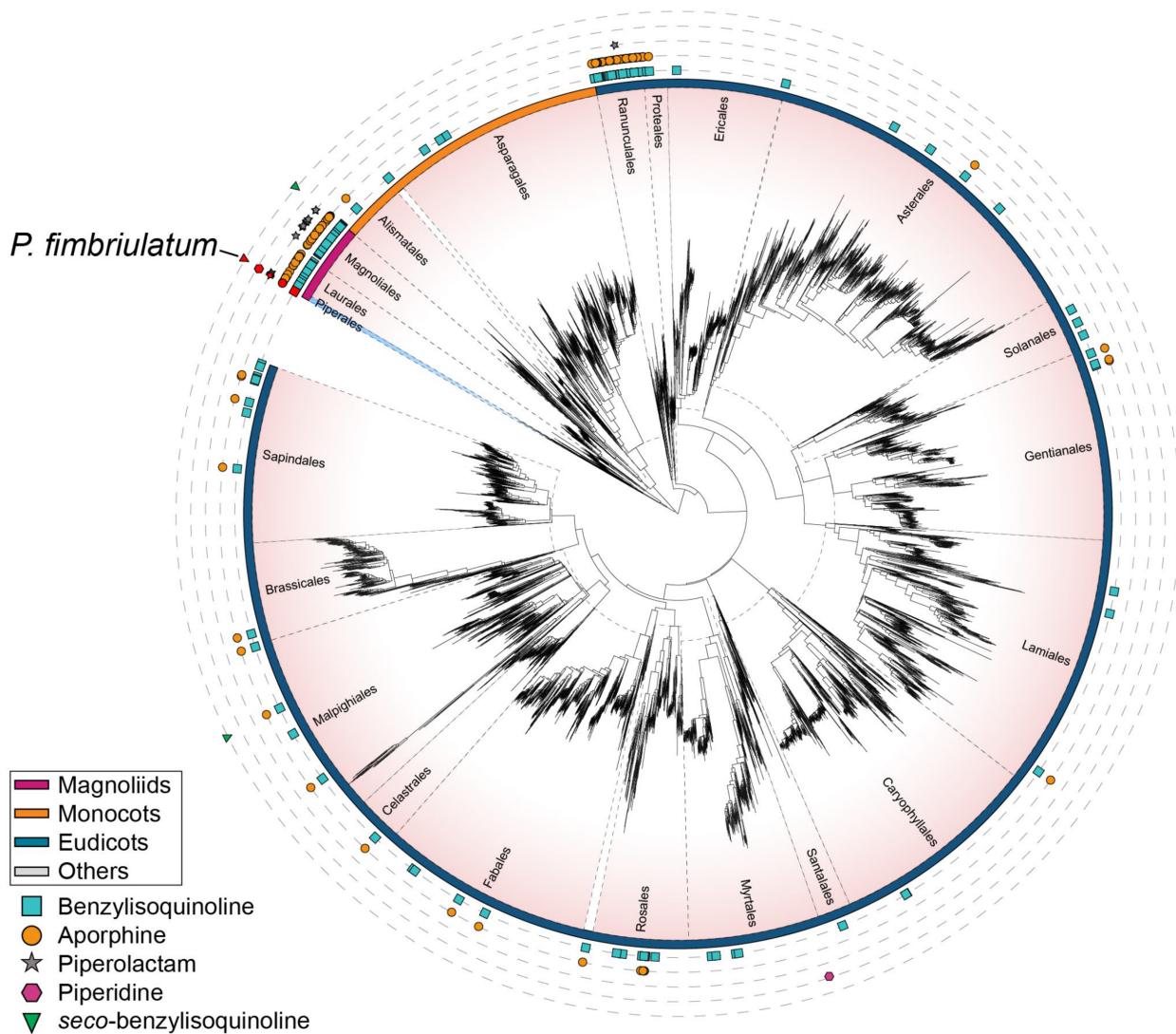
#### *Fimbriulatumine alkaloids molecular network (MN5)*

Finally, MN5 contained 18 nodes, out of which 13 were predicted as “tyrosine alkaloids” by CANOPUS. Interestingly, two nodes in the network (*m/z* 330.207 and 344.223, respectively) corresponded to dominant peaks in the LC-MS chromatogram of *P. fimbriulatum* leaves (see Figure 3b). No reliable match against the GNPS MS/MS spectral library was retrieved for any node within the network, thereby complicating the annotation process. For both peaks, CS:FinderID provided low-confidence (score <0.3) chemical structure predictions, and the manual inspection of the MS/MS spectra did not reveal fragment peaks or patterns common to other identified metabolites, which suggested these compounds could belong to a distinct alkaloid class. Before moving to the purification and structural characterization of these metabolites, we used an MS/MS-based dereplication strategy (see Figure 1 and Experimental procedures section) to reduce the chance of (re-)isolating already-known compounds. We started by using the recently developed plantMASST (Gomes et al., 2024) tool to verify whether a similar MS/MS spectrum was ever collected in publicly available plant metabolomics data sets. In particular, plantMASST searches MS/MS spectra across a curated database of LC-MS/MS data from plant extracts within the GNPS ecosystem which, at the time of writing, covers more than 2700 species from 1400 genera and 246 botanical families (Gomes et al., 2024). The plantMASST search revealed that both metabolites were exclusively found in *P. fimbriulatum* (Figures S12 and S13). Next, we searched the same MS/MS spectra against a wider range of public metabolomics data repositories using the Mass Spectrometry Search Tool (MASST) (Mongia et al., 2024; Wang et al., 2020). While plantMASST limits the search to curated plant LC-MS/MS data sets, MASST enables querying of the entire GNPS/MassLIVE (Wang et al., 2016) data repository (containing over 16 000 data sets and 8 billion of spectra at the time of writing). For both peaks, MASST returned no match outside data sets deposited by our own lab (MASST search 1, MASST search 2). Given the potential structural novelty, we proceeded with the targeted

isolation of these compounds from the original plant material (see Experimental procedures section). NMR analysis of the purified compounds confirmed the structure of two *seco*-benzyltetrahydroisoquinolines alkaloids (22) and (23) (Figures S21–S31), respectively, carrying a tertiary and quaternary linear amine moiety (Figure 3). While we found a previous report of (22), named cuspidatin, from the bark of *Antidesma cuspidatum* (Elya et al., 2014), a tropical plant from the Phyllanthaceae family, we did not find a previous report for (23) either in primary literature or in chemical structure databases (i.e., PubChem, Reaxys, SciFinder). Given the structural novelty and its unique occurrence in the *P. fimbriulatum* species (see Figure S13), we named the compound fimbriulatumine.

#### *Distribution of Piper alkaloids in angiosperms*

All identified alkaloids in this study are shown in Figure 3 (a) and summarized in Table S1. While *Piper* plants are mainly known for the production of piperamides (amide alkaloids), here we confirmed the occurrence of several non-amide alkaloids in *P. fimbriulatum*. Surprised by such diversity of alkaloid structures and scaffolds in a single *Piper* species, we contextualized our findings in an evolutionary framework using the recently published angiosperms tree of life (Zuntini et al., 2024), which covers 58% of the approximately 13 600 currently accepted genera and arguably constitutes the most comprehensive phylogenomic reconstruction of this clade’s evolution at the time of writing. Figure 4 shows the angiosperm phylogenetic tree mapped with literature reports, mined using Wikidata (see Experimental procedures section and Figure S14), of the alkaloid scaffolds we identified in *P. fimbriulatum* (i.e., benzylisoquinoline, aporphine, piperolactam, piperidine, *seco*-benzylisoquinoline). We found the widest distribution for the BIAs (203 genera from 48 families and 24 orders) and aporphine alkaloids (141 genera from 23 families and 16 orders) scaffolds. Within the Piperaceae family, we found only two reports of BIAs and aporphine alkaloids: one in *Piper argyrophyllum* dating back to 1996 (Singh et al., 1996) and, more recently, a second in *Piper sarmentosum* (Ware et al., 2024). Our work suggests a potential wider occurrence of these alkaloids within the *Piper* genus. Concerning the piperolactam scaffold, we found reports in 14 genera from 6 families and 3 orders and, interestingly, it always co-occurs with BIAs and/or aporphine alkaloids, except for two genera in the Saururaceae family (Wu et al., 2021; Zhuang et al., 2014). This supports the hypothesis that piperolactams are derived from 4,5-dioxoaporphine via benzilic acid rearrangement (Kumar et al., 2004), although evidence for this is still lacking. In contrast to the wide occurrence of the structural cores described above, we found the piperidine scaffold to be reported (almost) exclusively in Piperaceae plants, with only one report outside this family—that is, *Punica*



**Figure 4.** Angiosperm tree of life from Zuntini et al. (2024) mapped with literature reports, mined from Wikidata, of the alkaloid scaffolds considered in the present study (i.e., benzylisoquinoline, aporphine, piperolactam, piperidine, seco-benzylisoquinoline).

Each leaf in the tree corresponds to a representative species for each genus as described in the original publication. Different plant orders are separated by dashed black lines, and the Piperales order is highlighted in light blue. Reports for each scaffold are represented with different colored shapes (see legend in the figure). Reports for the *Piper* genus are highlighted in red. Colored arcs around the tree indicate the four main clades of angiosperms as described in the original publication: Magnoliids, Monocots, and Eudicots. The tree in the figure only shows the plant orders for which at least one scaffold was reported. A full version of the tree is provided in Figure S15. More details about the construction of the tree are provided in the [Experimental procedures](#) section. The figure was created using iTOL (Letunic & Bork, 2024). The original tree can be accessed at <https://itol.embl.de/tree/14723112167224931731658296>.

*granatum* (Chaturvedi et al., 2013) (Lythraceae family, Myrtales order). Finally, we found reports for the seco-benzylisoquinoline scaffold only in *Polyalthia insignis* (Lee et al., 1997) (Annonaceae family, Magnoliales order) and *Antidesma cuspidatum* (Phyllanthaceae family, Malpighiales order), while we report it for the first time in the *Piper* genus in this study. Notably, *Piper* is the only genus across the phylogenetic tree where all of the five considered scaffolds were reported, highlighting its remarkable alkaloid diversity.

We believe that the workflow shown in Figure S14 provides a powerful approach for efficiently mining knowledge bases (e.g., Wikidata) and tracking the occurrence of NPs across the plant kingdom, which, in our opinion, remain largely underestimated. This can be crucial for studies addressing broader biological questions, such as uncovering the evolutionary history of these specialized metabolic pathways, by enabling the generation of new hypotheses. For instance, the presence of the benzylisoquinoline and aporphine scaffolds in the early diverging

Piperales order may support their monophyletic evolution in angiosperms (Liscombe et al., 2005). In contrast, the occurrence of the piperidine scaffold exclusively in the *Piper* and *Punica* genera might indicate convergent evolution of piperidine alkaloid biosynthesis, considering the large phylogenetic distance between these two genera. While proving this claim would require a much deeper investigation (see, e.g., Huang et al., 2016), we postulate that leveraging knowledge bases combined with insightful visualization strategies can significantly support this exploration.

## CONCLUSIONS

In the present study, we investigated the alkaloid diversity of *P. fimbriulatum* using untargeted LC–MS/MS metabolomics. We used a range of computational tools to assist in the exploration of the detected chemical space and leveraged open MS/MS spectral libraries and data repositories to direct our isolation efforts towards structurally novel compounds. Overall, we documented the natural occurrence of 23 alkaloids belonging to five different classes, including a novel seco-benzylisoquinoline alkaloid carrying a linear quaternary amine nitrogen, which we named fimbriulatamine. Next, we contextualized these findings within the angiosperm evolutionary framework by mining the Wikidata open knowledge base and mapping the results onto the angiosperm tree of life. This analysis highlighted a remarkable diversity of alkaloid scaffolds within the *Piper* genus.

While expanding the known alkaloid diversity of the Piperaceae family and confirming it as a great source of specialized metabolites, our findings demonstrate the value of revisiting well-studied plant families using state-of-the-art computational metabolomics tools to uncover overlooked chemodiversity. In fact, many of the identified alkaloids were never reported in Piperaceae plants, despite their extensive phytochemical investigation over the past four decades. Another highlight of our work is the effectiveness of the “spectral dereplication” in the proposed LC–MS/MS data analysis workflow, which leverages the ever-growing MS/MS spectral libraries and metabolomics data repositories to efficiently annotate both known and structurally novel metabolites. This is demonstrated by the fact that, apart from the new fimbriulatamine, all other alkaloid structures were confirmed simply by purchasing the corresponding analytical standard, thus saving significant time and resources that would have otherwise been spent on the (re)isolation of known compounds.

We believe that the analytical and computational workflows presented here can be instrumental in improving our capabilities for the phytochemical characterization of non-model plant species. This is crucial for studies addressing broader biological questions, such as physiological processes and ecological interactions, which often

rely on identifying novel metabolites mediating these processes. Moreover, a more effective identification and reporting of structurally novel metabolites, as well as known compounds in new species, will accelerate efforts to map their distribution across the plant kingdom, which, in our opinion, remains largely underestimated. This could ultimately lead to the generation of new evolutionary hypotheses about the biosynthetic origin of specialized metabolic pathways.

Despite their potential, many of the computational tools employed in this study have not yet become routine resources for many researchers outside the metabolomics field. For this reason, we ensured the full reproducibility of the computations and results by sharing all generated data and scripts in open repositories. We hope this will help nonexperts familiarize themselves with the computational tools used in this manuscript and facilitate their broader adoption within the plant science community.

## EXPERIMENTAL PROCEDURES

### Chemicals

Commercial analytical standards Piperlongumine (>97% purity), piperidine (99% purity), and higenamine (95% purity) were purchased from Merck-Sigma-Aldrich (Prague, Czech Republic). Higenamine, coclaurine, lotusine, magnocurarine, lirinidine, and magnoflorine were all obtained in a natural product compound library from Targetmol (Wellesley, USA). Chemical synthesis compounds: 3,4,5-Trimethoxycinnamic acid (>98% purity) were purchased from BLDpharm (Turnov, Czech Republic). Dichloromethane (99.8% purity) and Tetrahydrofuran (99.5% purity) extra dry over molecular sieve from Fisher Scientific (Germany). Pivaloyl-chloride (99% Purity), Triethylamine (>99.5% purity) 2-Piperidinone were from Merck-Sigma-Aldrich (Prague, Czech Republic). Solvents and additives for LC–MS analysis (i.e., acetonitrile, water, formic acid, all Optima LCMS Grade) were purchased from Fisher Scientific (Germany). Concerning the molecular biology reagents, ultra-pure agarose was purchased from Invitrogen-Fisher (Prague, Czech Republic). NaCl p.a. was from Penta Chemicals (Prague, Czech Republic), molecular biology-grade NaHCO<sub>3</sub>, Na<sub>2</sub>SO<sub>4</sub>, MgSO<sub>4</sub>, and CTAB extraction buffer were from SERVA (Heidelberg, Germany). DMSO molecular biology grade was from Merck-Sigma-Aldrich (Prague, Czech Republic)..

### Sample collection and taxonomic confirmation

*P. fimbriulatum* samples were collected from the Prague Botanical Garden in January 2021 (specimen no. 2016.00648). Leaf, stem, and root organs were collected in triplicate from a single plant. After collection, samples were immediately stored in dry ice and transferred to –80°C until analysis. The taxonomic identity of *P. fimbriulatum* was confirmed through DNA barcoding on the internal transcribed spacer (ITS) region, which proved to be an efficient marker for *Piper* species differentiation (Jaramillo et al., 2008). Genomic DNA was isolated from 50 mg of fresh leaf tissue using the cetyltrimethylammonium bromide (CTAB) protocol as described in Aboul-Maaty and Oraby (2019). The ITS region was amplified by polymerase chain reaction (PCR) using the Phusion High-Fidelity DNA Polymerase (M0530L, New England Biolabs) and the following primers (as reported in Jaramillo et al., 2008):

- ITS-A (forward, 5' → 3'): GGAAGGAGAAGTCGTAACAAAGG
- ITS-B (reverse, 5' → 3'): CTTTCCTCCGCTTATTGATATG

PCR reactions were conducted in a 3 × 32-well ProFlex thermocycler (Thermo Fisher) and in 20 µL volume: 12.75 µL of nuclease-free water, 4 µL of 5X HF buffer, 1 µL of 10 µM forward primer (ITS-A), 1 µL of 10 µM reverse primer (ITS-B), 0.5 µL of 10 mM dNTPs mix, 0.5 µL of template DNA, and 0.25 µL of Phusion DNA polymerase. After an initial denaturation step at 98 °C for 30 sec, the following thermocycling conditions were applied for 35 cycles: 98°C for 10 sec (denaturation), 61°C for 15 sec (annealing), 7°C for 30 sec (extension). The reaction was concluded with a final extension step at 72°C for 10 min. PCR products were separated by gel electrophoresis (80 V for 45 min) on a 1% (w/v) agarose gel and purified using NucleoSpin silica gel cartridges (Macherey-Nagel) prior to sequencing. Sanger sequencing was performed and a consensus sequence was generated by aligning the sequencing results from both primers. The resulting consensus sequence was searched against the NCBI database using the BLAST alignment tool (Camacho et al., 2009). The top match (100% identity, 99% query coverage, E-value of 0) was retrieved against an ITS sequence deposited for *P. fimbriulatum* (GenBank EF056251.1; NCBI:txid425151). Raw sequencing results and the resulting consensus sequence are provided in the Supplementary Information.

### Metabolite extraction

Frozen plant material was homogenized under liquid nitrogen using a mortar and pestle. In this study, 50 mg ( $\pm$ 5 mg) of material was weighed in a prefrozen 2 mL plastic tube with a round bottom. Samples were further homogenized by adding a prefrozen stainless steel bead to each tube and high-speed shaking (30 sec, 25 rpm) using TissueLyser II (QIAGEN). For secondary metabolites extraction, each sample was added to 1 mL of ethanol/H<sub>2</sub>O (75:25) solution, incubated for 5 min at 40°C, and shaken again for 60 sec at 25 rpm. We opted for this extraction mixture as, based on preliminary testing, it offered the best compromise in terms of metabolite coverage and (relative) extraction yield. Following the extraction step, samples were centrifuged for 10 min at 14 000 rpm; 750 µL of supernatant was transferred into new 1.5 mL tubes and dried using a SpeedVac vacuum concentrator (Thermo Fisher). Finally, samples were resuspended in 750 µL of H<sub>2</sub>O/CH<sub>3</sub>CN (50:50) solution and transferred to glass vials for LC-MS analysis. Extraction blanks were prepared and analyzed in order to identify and remove signals arising from contaminants introduced during the extraction process (see Feature detection section).

It is important to note that we have not performed absolute quantification of the metabolites reported in this manuscript (see Results and discussion section). Therefore, the relative abundances of detected metabolites may be influenced by their ionization efficiency, the chosen extraction mixture as well as the resuspension in H<sub>2</sub>O/CH<sub>3</sub>CN (50:50) prior to LC-MS analysis. For this reason, abundance comparisons between metabolites (see, e.g., Figure 3) should be made with this in mind.

### LC-MS analysis

LC-MS/MS analyses were performed using a Vanquish UHPLC system (Thermo Fisher Scientific) coupled to an Orbitrap ID-X mass spectrometer equipped with a heated electrospray ionization (HESI) source. Reverse-phase (RP) separation of analytes was performed on an Acquity BEH C18 column, 150 mm × 2.1 mm, 1.7 µm particle size (Waters). The temperature in the autosampler and column oven was set to 10 and 40°C, respectively. The mobile

phase consisted of water (A) and CH<sub>3</sub>CN (B), both containing 0.1% formic acid. A constant flow rate of 350 µL min<sup>-1</sup> was used, and the chromatographic gradient was as follows: initial hold of 0.5 min at 5% of B, linear gradient of 15 min from 5 to 100% of B, isocratic step of 1.8 min at 100% of B for column wash, and isocratic step of 2 min at 5% of B for column reconditioning. The injection volume was set to 1 µL. HESI source parameters were set to 50 arbitrary units (AU) of sheath gas, 12 AU of auxiliary gas, 1 AU of sweep gas, vaporizer temperature of 350°C, spray voltage of 3000 V, and ion transfer tube temperature of 325°C. MS data were acquired in positive ionization and data-dependent acquisition mode (DDA), with MS2 spectra collected with a cycle time of 0.6 sec (i.e., maximum time between MS1 scans). MS1 data were collected in profile mode, 100–1000 m/z scan range, 60 000 resolution, 1 microscan, 45% RF lens, normalized AGC target of 50% (i.e.,  $\approx$ 2E5), and maximum ion injection time of 118 milliseconds. MS2 data were collected in profile mode, 15 000 resolution, 1 microscan, normalized AGC target of 100% (i.e.,  $\approx$ 5E4), and maximum ion injection time of 80 msec. Selected precursor ions were fragmented with a quadrupole isolation window of 0.8 m/z, 1E5 minimum intensity threshold, dynamic exclusion of 2 sec (5 ppm tolerance, isotope exclusion set to "ON"), fixed normalized HCD collision energy of 35%, apex detection set to "ON" with expected peak width (FWHM) of 4 sec and desired apex window of 50%. Raw LC-MS/MS data were visualized using Freestyle v1.8 (Thermo Fisher Scientific).

### LC-MS data processing

**Feature detection and feature-based molecular networking**  
Raw LC-MS data were converted from Thermo Fisher proprietary format (.RAW) to open format (.mzML) using the MSConvert software (ProteoWizard) (Chambers et al., 2012). Untargeted feature detection was performed using the mzmine software (Schmid et al., 2023) (v4.2.0) as described in Heuckeroth et al. (2024). Briefly, the mass detection noise level was set to 5.0E4 and 2.0E3 for MS1 and MS2 levels, respectively, and all signals below these intensities were discarded. Extracted ion chromatograms (XIC) were built for each m/z with a 0.002 Da or (5 ppm) tolerance and retained if exhibiting at least seven consecutive data points above 1.5E5 intensity, and a minimum absolute height or 5.0E5. After smoothing (LOESS method, 4-scans window), XICs were resolved (Local Minimum Resolver module) using the following parameters: chromatographic threshold = 90%, minimum search range = 0.05, minimum absolute height = 5.0E5, minimum ratio of peak top/edge = 1.70, peak duration range = 0.0–2.0 min, minimum scans = 5. MS2 spectra were paired to the resolved peaks with a 0.002 Da or (5 ppm) "MS1 to MS2 precursor" tolerance and using feature edges as retention time (RT) limits. Redundant <sup>13</sup>C isotope features were filtered using 0.001 Da (or 3.5 ppm) and 0.01 min tolerances; a monoisotopic shape of the isotopic pattern was required and the most intense isotope was kept as representative. Feature alignment (Join aligner module) was performed using 0.002 Da (or 5 ppm) and 0.08 min sample-to-sample tolerances (equal weight was attributed to m/z and RT). Duplicated features (i.e., potential artifacts from the alignment) were removed using 0.0005 Da (or 2 ppm) and 0.02 min tolerances. Features detected in the extraction blank samples were removed from the aligned feature table unless they exhibited a 3x fold change compared to the blank. Correlation grouping and ion identity networking were carried out using default parameters (see provided batch file). Finally, features that were not detected in at least three samples (since each plant organ was collected and analyzed in triplicate) and/or did not contain at least two isotopes in the isotopic

pattern were filtered out. The resulting feature table was exported for further analysis by FBMN and the SIRIUS software. For FBMN, MS/MS spectra were merged across samples (weighted average) using the following parameters: intensity merge mode = sum, expected mass deviation: 0.02 Da (or 5 ppm), cosine threshold = 0, signal count threshold = 0, isolation window offset = 0, isolation window width of 0.8 Da. For SIRIUS, MS/MS spectra were not merged and exported with a *m/z* tolerance of 0.002 Da (or 5 ppm). The above-described feature detection steps and parameters are also provided as a mzmine configuration batch file (see **Supplementary Information**).

FBMN was performed through the Global Natural Products Social Molecular Networking (GNPS) platform (Wang et al., 2016) using the following parameters: 0.01 Da parent mass tolerance, 0.01 Da MS/MS fragment ion tolerance, a minimum cosine similarity of 0.7, and a minimum of 6 matched peaks. The FBMN results were visualized using Cytoscape (v3.10.2) (Shannon et al., 2003).

#### *Metabolite annotation*

For the (putative) annotation of unknown metabolites in the molecular network, we used a combination of MS/MS spectral library matching (Bittremieux et al., 2022), *in silico* prediction of chemical structures and compound classes (Dührkop et al., 2015, 2021) and manual interpretation of MS/MS spectra. During the FBMN workflow in GNPS, all spectra are automatically searched against the public reference MS/MS libraries hosted on the platform. Both exact and analog matches with cosine similarity above 0.7 and  $\geq 6$  matched fragment peaks were retrieved. "Analog matches" are retrieved by using a spectral similarity score that takes into account both fragment peaks and neutral losses in the similarity calculation. To increase the annotation coverage, we computationally annotated molecular formulas, chemical structures, and compound classes for all MS/MS spectra in the network using the SIRIUS software (v5.8.5) (Dührkop et al., 2019). Molecular formulas were computed with the formula module by matching the experimental and predicted isotopic patterns (Böcker et al., 2009) and from fragmentation tree analysis (Böcker & Dührkop, 2016) of MS/MS. Default parameters were used, except for instrument type = orbitrap, mass accuracy for MS1 = 5 ppm, mass accuracy for MS2 = 7.5 ppm. Formula predictions were refined with the ZODIAC module (Ludwig et al., 2020) using default parameters. *In silico* structure annotation was done with CSI:FingerID (Dührkop et al., 2015) using structures from biological databases. Systematic compound class annotations were obtained with CANOPUS (Dührkop et al., 2021) using default parameters and the NPClassifier (Kim et al., 2021) ontology.

For the MS/MS-based dereplication, we used MASST (Mongia et al., 2024; Wang et al., 2020) and plantMASST (Gomes et al., 2024). MASST (Mass Spectrometry Search Tool) is an MS/MS similarity search tool for querying MS/MS spectra of unknown molecules against public data repositories, such as the GNPS/MassIVE repository (over 16 000 metabolomics data sets and 8 billion spectra at the time of writing). Spectral matches are returned along with data set(s) metadata, which can provide biological context information about the queried MS/MS spectrum, even when the corresponding chemical structure is unknown. PlantMASST is a domain-specific version of MASST, which searches the query MS/MS spectrum within a curated reference database of LC–MS/MS data acquired from plant extracts. Spectral matches are returned along with metadata information about the plant species, genus, etc. At the time of writing, the plantMASST reference database contains data from over 19 000 plant extracts

covering 246 botanical families, 1400 genera, and 2700 species (Gomes et al., 2024). Both MASST and plantMASST searches were performed with the following default parameters: precursor mass tolerance = 0.05 Da, fragment mass tolerance = 0.05 Da, cosine threshold = 0.7, minimum matched peaks = 3, analog search = OFF.

Manual interpretation of MS/MS spectra was done using a range of dedicated software tools, such as mzmine, SIRIUS, Metabolomics Spectrum Resolver (Bittremieux et al., 2020), and the GNPS Dashboard (Petras et al., 2022). In particular, to propagate the annotation from confirmed metabolites to neighboring nodes in the molecular network, we took advantage of the recently developed ModiFinder (Shahneh et al., 2024) tool, which compares the MS/MS spectra of a known structure and its unknown analog to locate the most likely modification site. All the putative annotations were manually inspected and, whenever possible, confirmed with either commercially available or *in-house* synthesized analytical standards (see **Chemical synthesis of piperamides** section).

#### **Chemical synthesis of piperamides**

Compounds (2), (3), (5), and (6) were synthesized as described in Wu et al. (2014) and Rao et al. (2012) (Figure S16a,b). For compounds (2) and (5), in an argon atmosphere, trimethoxycinnamic acid (1.0 eq) was dissolved in dry THF (5 mL), cooled to 0°C, and triethylamine (1.5 eq) was added. Pivaloyl chloride (1.1 eq) was added dropwise to the reaction mixture and stirred for 1 h at 0°C. Afterward, 2-piperidinone (1.1 eq) was dissolved in dry THF (2 mL), cooled to –78°C, and 1.6 M *n*-BuLi in hexane (1.1 eq) was added dropwise and stirred for 1 h at –78°C. A previously prepared mixture of anhydride was added to the 2-piperidinone reaction mixture dropwise at –78°C and stirred for 1 h. The reaction mixture was quenched with saturated NH<sub>4</sub>Cl (5 mL), diluted with H<sub>2</sub>O (30 mL) and extracted with EtOAc (3 × 25 mL). Combined organic layers were washed with saturated NaCl (30 mL), dried over anhydrous MgSO<sub>4</sub>, and evaporated to dryness. The crude product was purified by RP-flash chromatography (50 g column, H<sub>2</sub>O/MeOH, 0–100%), yielding pure (2) and (5). For compounds (3) and (6), trimethoxycinnamic acid (1.00 eq) was dissolved in dry DCM (5 mL), cooled to 0°C, and then triethylamine (2 eq) was added. Pivaloyl chloride (1.1 eq) was added dropwise to the reaction mixture and stirred for 1 h at 0°C. Afterward, piperidine (1.5 eq) was added dropwise and stirred overnight. Triethylamine (1.100 eq) was added to a stirred solution of piperidine (1.5 eq) in dry DCM (2 mL) at 0°C and stirred for 1 h. The reaction mixture was added dropwise to the anhydride reaction mixture at 0°C and stirred overnight. The reaction mixture was diluted with DCM, washed with saturated NaHCO<sub>3</sub> (20 mL) and saturated NaCl (20 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated, and purified by RP-flash chromatography (50 g column, H<sub>2</sub>O/MeOH, 0–100%), yielding pure (3) and (6). All pure compounds were dissolved in DMSO-*d*<sub>6</sub> for the subsequent NMR measurements. The NMR spectra were measured on a 400-MHz Bruker Avance III HD spectrometer. Chemical shifts are reported in ppm ( $\delta$ ) relative to the solvent (DMSO-*d*<sub>6</sub>) signal:  $\delta$  = 2.54 and 40.45 ppm for <sup>1</sup>H and <sup>13</sup>C, respectively (Figures S17–S20).

For the synthesis of piperlongumine dimer, piperlongumine (1) standard was dissolved in 1.67 mL of methanol (2 mM), transferred to a glass vial, and evaporated in the dark under a gentle nitrogen flow. This resulted in a thin, homogeneous film of the compound. The film was then irradiated with a UV lamp at 365 nm. At 0, 1, 5, 10, and 30 min, irradiation was stopped, the film was dissolved in MeOH as above, a time point was taken (10  $\mu$ L), and then again evaporated into a film. After irradiation,

samples were immediately stored in the dark at  $-20^{\circ}\text{C}$  until analysis. Aliquots were diluted to 1  $\mu\text{M}$  concentration and analyzed by LC–MS.

### Isolation of cuspidatin and fimbriulatumine

Lyophilized, ground leaves ( $\approx 2$  g) of *P. fimbriulatum* were sequentially extracted by maceration four times with 150 mL of methanol for 1.5 h (each round), at room temperature and protected from direct light sources. The resulting solutions were filtered and combined in a round-bottom flask and evaporated under reduced pressure at  $40^{\circ}\text{C}$ , yielding  $\approx 500$  mg of a dark-green residue. The obtained extract was redissolved in 20 mL of methanol and transferred to a separatory funnel and washed three times with 10 mL of *n*-octane. The methanol fractions resulting from each wash were pooled and evaporated again, yielding  $\approx 350$  mg of residue. The so-obtained residue was dissolved in methanol (concentration  $\approx 100$  mg  $\text{mL}^{-1}$ ), centrifuged and fractionated using a preparative HPLC system (1290 Infinity II Preparative Binary Pump, Agilent Technologies) equipped with a UV–Vis detector (1260 Infinity II Variable Wavelength Detector, Agilent Technologies) and fraction collector (1290 Infinity II Preparative Open-Bed Fraction Collector, Agilent Technologies). The sample was separated by reversed-phase chromatography using an XBridge C18 column, 10 mm  $\times$  250 mm, 5  $\mu\text{m}$  particle size (Waters). The temperature in the column oven was set to  $40^{\circ}\text{C}$ . The mobile phase consisted of water (A) and MeOH (B), both containing 0.1% formic acid. A constant flow rate of 17  $\text{mL min}^{-1}$  was used and the chromatographic gradient was as follows: initial hold of 3 min at 10% of B, linear gradient of 62 min from 10 to 70% of B, isocratic step of 5 min at 100% of B for column wash, isocratic step of 5 min at 10% of B for column reconditioning. The injection volume was set to 200  $\mu\text{L}$  and the sample was injected multiple times until collection of sufficient material for structural confirmation via NMR. This separation yielded  $\approx 1$  mg of compounds (22) and (23). All compounds were dissolved in  $\text{CD}_3\text{CN}$  for subsequent structural confirmation by NMR. The NMR spectra were measured on a 500- or a 600-MHz Bruker Avance III HD spectrometer ( $^1\text{H}$  at 500.0 or 600.1 MHz and  $^{13}\text{C}$  at 125.7 or 150.9 MHz). The spectra were referenced to solvent signals ( $\delta = 1.94$  and 1.32 ppm for  $^1\text{H}$  and  $^{13}\text{C}$ , respectively). The signal assignment was performed using a combination of 1D ( $^1\text{H}$  and  $^{13}\text{C}$ ) experiments and 2D correlation experiments (H,H-COSY, H,C-HSQC and H,C-HMBC). In some cases, the sample quantity was insufficient for the detection of some or all  $^{13}\text{C}$  signals in the 1D experiments. The structural analysis and signal assignment were then based on the 2D correlation experiments. Through literature search, we found a previous report (Elya et al., 2014) of compound (22), with NMR data deposited in the SpectraBase repository (SpectraBase ID: HqNww9TxQ60).

### Mapping alkaloid scaffolds to the phylogeny of angiosperms

To investigate the distribution of specific alkaloid scaffolds across angiosperms, we designed SPARQL queries to retrieve natural products containing one of the five alkaloid scaffolds reported in *P. fimbriulatum* in the present study (i.e., benzylisoquinoline, aporphine, pipеролактам, piperidine, seco-benzylisoquinoline, Figure S14). The queries targeted reported natural product structures containing the scaffold (sub)structure (defined using SMILES string), along with the corresponding plant species they were isolated from, as documented in Wikidata based on literature reports. The original SPARQL queries are available in the Supplementary Information and in the GitHub repository (see Data availability section). Query results were manually inspected and refined as

follows. First, structures that were retrieved by the initial query but were clearly unrelated to the target scaffold were filtered out. This was achieved by defining chemical substructures common to the “unwanted” structure retrieved by the initial query (Figure S14b) and using the HasSubstructMatch module from the RDKit Python library (v2023.03.2). The substructures used for this filtering are shown in Figure S14(c,d), and both unfiltered and filtered data sets are included in the Supplementary Information. In addition, to facilitate review and reproducibility of this manuscript, we created a Jupyter Notebook (clean\_wikidata.ipynb, available in the GitHub repository) that interactively visualizes the substructures removed for each scaffold. After structural filtering, we manually inspected the cleaned data sets to identify false or unreliable reports due to errors in Wikidata and/or primary literature. In particular, the query results for the piperidine scaffold included erroneous reports of piperine in *Capsicum* species (attributed to an error in Wikidata; Weaver et al., 1984) and entries for *Centaurea aegyptica* (Dahmy et al., 1985) and *Aglaia perviridis* (Zhang et al., 2010), both of which were excluded due to a lack of NMR data in the original publication. Next, we mapped the so-cleaned literature reports onto the recently published phylogenomics tree of angiosperms (Zuntini et al., 2024). In particular, the tree file named global\_tree\_brlen\_pruned\_renamed.tree was used (see Zuntini & Carruthers, 2024). The latter contains one representative species per genus; therefore, the mapping was done at the genus level—that is, the genus names were extracted from the literature reports and matched with the genera present in the angiosperm tree. This was done using custom Python scripts (03\_clean\_wikidata.py and 05\_create\_itol\_annotation.py) available in the GitHub repository. Phylogenetic tree visualization and figure creation were done using the Interactive Tree Of Life (iTOL) online tool (Letunic & Bork, 2024). The phylogenetic tree displayed in Figure 4 is a smaller version of the original tree created by removing plant orders for which no report for any alkaloid scaffold was retrieved (see Python scripts 05\_create\_small\_tree.py in the GitHub repository). The tree can be accessed at <https://itol.embl.de/tree/14723112167224931731658296>. The full tree is shown in Figure S15 and can be accessed at <https://itol.embl.de/tree/14723112167277531728383616>.

### ACKNOWLEDGMENTS

We thank Klára Lorencová, Petr Vacík, and Ludvík Bortl from the Prague Botanical Garden for their support with the sample collection. TP was supported by the Czech Science Foundation (GA CR) grant 21-11563M and by the European Union’s Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No. 891397. TD was supported by the European Regional Development Fund; P JAC; Project “IOCB MSCA PF Mobility” (No. CZ.02.01.01/00/22\_010/0002733). TH was supported by the European Union’s Horizon Europe research and innovation program under the Marie Skłodowska-Curie grant agreement No. 101130799. MP received support from the Visegrad Fund for presenting this research at conferences. Open access publishing facilitated by Ustav organické chemie a biochemie Akademie ved České republiky, as part of the Wiley - CzechELib agreement.

### CONFLICT OF INTEREST

The authors have not declared a conflict of interest.

### DATA AVAILABILITY STATEMENT

Raw LC–MS files (.mzML), mzmine configuration batch file, and processing results are available through Zenodo

(<https://zenodo.org/records/14336729>). NMR data are also available through Zenodo at the same link. SPARQL queries and corresponding output (both raw and filtered), phylogenetic trees, and iTOL annotation files are also available through Zenodo at the same link. All code and software is available through GitHub under the following link <https://github.com/pluskal-lab/PiperFIM>. The FBMN results are available at the following link: <https://gnps2.org/status?task=43853471d153488b95144abd3af36a9d>. Annotated phylogenetic trees can be accessed at <https://itol.embl.de/tree/14723112167277531728383616> (Figure 4), and <https://itol.embl.de/tree/14723112167224931731658296> (Figure S15).

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

**Table S1.** List of *Piper fimbriulatum* alkaloids identified in the present study.

**Figure S1.** Photo of the *Piper fimbriulatum* plant sampled in Prague Botanical Garden.

**Figure S2.** Spectral matches against the GNPS MS/MS spectral library.

**Figure S3.** Confirmation of piperamide annotations via retention time matching with analytical standards.

**Figure S4.** Interpretation of piperamides MS/MS spectra.

**Figure S5.** Interpretation of piperamides MS/MS spectra.

**Figure S6.** MS/MS mirror plot of piplartin dimer (7) and piperlongumine (1).

**Figure S7.** Results of the direct LC-MS analysis of the piperlongumine dimer reaction mixture (panel A and B).

**Figure S8.** Confirmation of BIAs and aporphine alkaloids via retention time matching with analytical standards.

**Figure S9.** Proposed MS/MS fragmentation pathway for higenamine [in accordance with Zhao et al. (2022)].

**Figure S10.** Interpretation of BIAs MS/MS spectra.

**Figure S11.** Interpretation of BIAs MS/MS spectra.

**Figure S12.** plantMASST search output of cuspidatin (22).

**Figure S13.** plantMASST search output of fimbriulatumine (23).

**Figure S14.** Computational workflow used to map literature reports for the alkaloid scaffolds found in *Piper fimbriulatum* in the present study onto the Angiosperm tree of life (Zuntini et al., 2024).

**Figure S15.** Angiosperm tree of life from Zuntini et al. (2024) mapped with literature reports, mined from Wikidata, of the alkaloid scaffolds considered in the present study (i.e., benzylisoquinoline, aporphine, piperolactam, piperidine, seco-benzylisoquinoline).

**Figure S16.** Scheme of the chemical synthesis of piperamides.

**Figure S17.**  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra of dihydropiperlongumine (2).

**Figure S18.**  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra of 1-(3,4,5-Trimethoxycinnamoyl)piperidine (3).

**Figure S19.**  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra of 3'-Demethoxypiperlongumine (5).

**Figure S20.**  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra of 1-(3,4-Dimethoxycinnamoyl)piperidine (6).

**Figure S21.**  $^1\text{H}$  NMR spectrum of cuspidatin (22).

**Figure S22.**  $^{13}\text{C}$  NMR spectrum of cuspidatin (22).

**Figure S23.** COSY spectrum of cuspidatin (22).

**Figure S24.** HSQC spectrum of cuspidatin (22).

**Figure S25.** HMBC spectrum of cuspidatin (22).

**Figure S26.**  $^1\text{H}$  NMR spectrum of fimbriulatumine (23).

**Figure S27.**  $^{13}\text{C}$  NMR spectrum of fimbriulatumine (23).

**Figure S28.** COSY spectrum of fimbriulatumine (23).

**Figure S29.** HSQC spectrum of fimbriulatumine (23).

**Figure S30.** HMBC spectrum of fimbriulatumine (23).

**Figure S31.** Structural elucidation of fimbriulatumine (23).

**Note S1.** Piperamides annotation.

**Note S2.** Benzylisoquinoline alkaloids annotation.

**Note S3.** Aporphine alkaloids and piperolactams annotation.

## REFERENCES

- Aboul-Maaty, N.A.-F. & Oraby, H.A.-S. (2019) Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. *Bulletin of the National Research Centre*, **43**(1), 25.
- Benidir, M.A., Kang, K.B., Genta-Jouve, G., Huber, F., Rogers, S. & van der Hooft, J.J.J. (2021) Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches. *Natural Product Reports*, **38**(11), 1967–1993.
- Bittremieux, W., Chen, C., Dorrestein, P.C., Schymanski, E.L., Schulze, T., Neumann, S. et al. (2020) Universal MS/MS visualization and retrieval with the metabolomics spectrum resolver web service. *bioRxiv*. Available from: <https://doi.org/10.1101/2020.05.09.086066>
- Bittremieux, W., Wang, M. & Dorrestein, P.C. (2022) The critical role that spectral libraries play in capturing the metabolomics community knowledge. *Metabolomics: Official Journal of the Metabolomic Society*, **18**(12), 94.
- Böcker, S. & Dürkop, K. (2016) Fragmentation trees reloaded. *Journal of Cheminformatics*, **8**(1), 5.
- Böcker, S., Letzel, M.C., Lipták, Z. & Pervukhin, A. (2009) SIRIUS: decomposing isotope patterns for metabolite identification. *Bioinformatics*, **25**(2), 218–224.
- Camacho, C., Couloris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**(1), 421.
- Carbone, F., Djamshidian, A., Seppi, K. & Poewe, W. (2019) Apomorphine for Parkinson's disease: efficacy and safety of current and new formulations. *CNS Drugs*, **33**(9), 905–918.
- Chambers, M.C., Maclean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S. et al. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, **30**(10), 918–920.
- Chaturvedi, A.K., Luqman, S., Dubey, V., Thakur, J.P., Saikia, D., Chanotiya, C.S. et al. (2013) Inhibition of Cathepsin D protease activity by *Punica granatum* fruit Peel extracts, isolates, and semisynthetic analogs. *Medicinal Chemistry Research: An International Journal for Rapid Communications on Design and Mechanisms of Action of Biologically Active Agents*, **22**(8), 3953–3958.
- Conde, J., Pumroy, R.A., Baker, C., Rodrigues, T., Guerreiro, A., Sousa, B.B. et al. (2021) Allosteric antagonist modulation of TRPV2 by piperlongumine impairs glioblastoma progression. *ACS Central Science*, **7**(5), 868–881.
- da Silva Mendes, J.W., Cunha, W.E.M., Filho, R.B., de Carvalho, N.K.G. & da Costa, J.G.M. (2023)  $^{13}\text{C}$  NMR spectroscopic data of aporphine alkaloids. *The Alkaloids. Chemistry and Biology*, **89**, 39–171.
- Dahmy, S., Bohlmann, F., Sarg, T., Ateya, A. & Farrag, N. (1985) New guianolides from *Centaurea aegyptica*. *Planta Medica*, **51**(2), 176–177.
- Dürkop, K., Fleischhauer, M., Ludwig, M., Aksenen, A.A., Melnik, A.V., Meusel, M. et al. (2019) SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nature Methods*, **16**(4), 299–302.
- Dürkop, K., Nothias, L.-F., Fleischhauer, M., Reher, R., Ludwig, M., Hoffmann, M.A. et al. (2021) Systematic classification of unknown

- metabolites using high-resolution fragmentation mass spectra. *Nature Biotechnology*, **39**(4), 462–471.
- Dührkop, K., Shen, H., Meusel, M., Rousu, J. & Böcker, S. (2015) Searching molecular structure databases with tandem mass spectra using CSI:FingerID. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(41), 12580–12585.
- Elya, B., Forestrania, R.C., Ropi, M., Kosela, S., Awang, K., Omar, H. et al. (2014) The new alkaloids from *Antidesma cuspidatum* M.A. Available from: <https://acgpubs.org/doc/2018080915074048-RNP-1310-448.pdf> [Accessed 26th November 2024]
- Gomes, P.W.P., Mannochio-Russo, H., Schmid, R., Zuffa, S., Damiani, T., Quiros-Guerrero, L.-M. et al. (2024) plantMASST—community-driven chemotaxonomic digitization of plants. *bioRxiv.org: The Preprint Server for Biology*, May. Available from: <https://doi.org/10.1101/2024.05.13.593988>
- Gómez-Calvario, V. & Rios, M.Y. (2019) <sup>1</sup>H and <sup>13</sup>C NMR data, occurrence, biosynthesis, and biological activity of piper amides. *Magnetic Resonance in Chemistry: MRC*, **57**(12), 994–1070.
- Gutiérrez, M.P., Rosa, A.M., Gonzalez, N. & Hoyo-Vadillo, C. (2013) Alkaloids from piper: a review of its phytochemistry and pharmacology. *Mini Reviews in Medicinal Chemistry*, **13**(2), 163–193.
- Heuckeroth, S., Damiani, T., Smirnov, A., Mokshyna, O., Brungs, C., Korf, A. et al. (2024) Reproducible mass spectrometry data processing and compound annotation in MZmine 3. *Nature Protocols*, **19**(9), 2597–2641.
- Huang, R., O'Donnell, A.J., Barboline, J.J. & Barkman, T.J. (2016) Convergent evolution of caffeine in plants by co-option of exapted ancestral enzymes. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(38), 10613–10618.
- Hurtley, A.E., Lu, Z. & Yoon, T.P. (2014) [2 + 2] cycloaddition of 1,3-dienes by visible light photocatalysis. *Angewandte Chemie*, **53**(34), 8991–8994.
- Jäckel, L., Schnabel, A., Stellmach, H., Klauß, U., Matschi, S., Hause, G. et al. (2022) The terminal enzymatic step in piperine biosynthesis is Co-localized with the product piperine in specialized cells of black pepper (*Piper nigrum* L.). *The Plant Journal: For Cell and Molecular Biology*, **111**(3), 731–747.
- Jaramillo, M.A., Callejas, R., Davidson, C., Smith, J.F., Stevens, A.C. & Tepe, E.J. (2008) A phylogeny of the tropical genus piper using ITS and the chloroplast intron psbJpetA. *Systematic Botany*, **33**(4), 647–660.
- Jarmusch, A.K., Aron, A.T., Petras, D., Phelan, V.V., Bittremieux, W., Acharya, D.D. et al. (2022) A universal language for finding mass spectrometry data patterns. *bioRxiv*. Available from: <https://doi.org/10.1101/2022.08.06.503000>
- Jarmusch, S.A., van der Hooft, J.J.J., Dorrestein, P.C. & Jarmusch, A.K. (2021) Advancements in capturing and mining mass spectrometry data are transforming natural products research. *Natural Product Reports*, **38**(11), 2066–2082.
- Jung, Y., Choi, S.Y., Choi, J.W., Cho, C.Y., Lee, D.K., Park, J. et al. (2024) Identification of novel dimers and chemical profiling of acid amide alkaloids in *Piper nigrum*. *Food Chemistry*, **450**, 139199.
- Kim, H.W., Wang, M., Leber, C.A., Nothias, L.-F., Reher, R., Kang, K.B. et al. (2021) NPClassifier: a deep neural network-based structural classification tool for natural products. *Journal of Natural Products*, **84**(11), 2795–2807.
- Kumar, V., Poonam, P., Prasad, A.K. & Parmar, V.S. (2004) Naturally occurring aristolactams, aristolochic acids and dioxaoporphines and their biological activities. *ChemInform*, **35**(11), 565–583. Available from: <https://doi.org/10.1002/chin.200411269>
- Lee, K.-H., Chuah, C.-H. & Goh, S.-H. (1997) Seco-benzyltetrahydroisoquinolines from *Polyalthia insignis* (annonaceae). *Tetrahedron Letters*, **38**(7), 1253–1256.
- Letunic, I. & Bork, P. (2024) Interactive tree of life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, **52**(W1), W78–W82.
- Liscombe, D.K., Macleod, B.P., Loukanina, N., Nandi, O.I. & Facchini, P.J. (2005) Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms. *Phytochemistry*, **66**(11), 1374–1393.
- Ludwig, M., Nothias, L.-F., Dührkop, K., Koester, I., Fleischauer, M., Hoffmann, M.A. et al. (2020) Database-independent molecular formula annotation using Gibbs sampling through ZODIAC. *Nature Machine Intelligence*, **2**(10), 629–641. Available from: <https://doi.org/10.1038/s42256-020-00234-6>
- Mayer, V., Schaber, D. & Hadacek, F. (2008) Volatiles of myrmecophytic piper plants signal stem tissue damage to inhabiting pheidole ant-partners. *The Journal of Ecology*, **96**(5), 962–970.
- Mongia, M., Yasaka, T.M., Liu, Y., Guler, M., Lu, L., Bhagwat, A. et al. (2024) Fast mass spectrometry search and clustering of untargeted metabolomics data. *Nature Biotechnology*, **42**(11), 1672–1677. Available from: <https://doi.org/10.1038/s41587-023-01985-4>
- Mundina, M., Vila, R., Tomi, F., Gupta, M.P., Adzet, T., Casanova, J. et al. (1998) Leaf essential oils of three Panamanian piper species. *Phytochemistry*, **47**(7), 1277–1282.
- Mutabdija, L., Myoli, A., de Jonge, N.F., Damiani, T., Schmid, R., van der Hooft, J.J.J. et al. (2024) Studying plant specialized metabolites using computational metabolomics strategies. *Methods in Molecular Biology*, **2788**, 97–136.
- Nguyen, T.B. & Al-Mourabit, A. (2016) Remarkably high homoselectivity in [2 + 2] photodimerization of trans-cinnamic acids in multicomponent systems. *Photochemical & Photobiological Sciences: Official Journal of the European Photochemistry Association and the European Society for Photobiology*, **15**(9), 1115–1119.
- Nothias, L.-F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A. et al. (2020) Feature-based molecular networking in the GNPS analysis environment. *Nature Methods*, **17**(9), 905–908.
- Parmar, V.S., Jain, S.C., Bisht, K.S., Jain, R., Taneja, P., Jha, A. et al. (1997) Phytochemistry of the genus piper. *Phytochemistry*, **46**(4), 597–673.
- Petras, D., Phelan, V.V., Acharya, D., Allen, A.E., Aron, A.T., Bandeira, N. et al. (2022) GNPS dashboard: collaborative exploration of mass spectrometry data in the web browser. *Nature Methods*, **19**(2), 134–136.
- Rao, V.R., Puppala Muthenna, G.S., Akileshwari, C., Babu, K.H., Suresh, G., Babu, K.S. et al. (2012) Synthesis and biological evaluation of new piplartine analogues as potent aldose reductase inhibitors (ARIs). *European Journal of Medicinal Chemistry*, **57**(November), 344–361.
- Rutz, A., Sorokin, M., Galonek, J., Mietchen, D., Willighagen, E., Gaudry, A. et al. (2022) The LOTUS initiative for open knowledge management in natural products research. *eLife*, **11**, e70780. Available from: <https://doi.org/10.7554/eLife.70780>
- Salehi, B., Zakaria, Z.A., Gyawali, R., Ibrahim, S.A., Rajkovic, J., Shinwari, Z.K. et al. (2019) Piper species: a comprehensive review on their phytochemistry, biological activities and applications. *Molecules*, **24**(7), 1364.
- Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., Dyrlund, T.S. et al. (2023) Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nature Biotechnology*, **41**(4), 447–449.
- Schnabel, A., Athmer, B., Manke, K., Schumacher, F., Cotinguiba, F. & Vogt, T. (2021) Identification and characterization of Piperine synthase from black pepper, *Piper nigrum* L. *Communications Biology*, **4**(1), 445.
- Shahneh, M.R.Z., Strobel, M., Vitale, G.A., Geibel, C., El Abied, Y., Garg, N. et al. (2024) ModiFinder: tandem mass spectral alignment enables structural modification site localization. *Journal of the American Society for Mass Spectrometry*, **35**(11), 2564–2578. Available from: <https://doi.org/10.1021/jasms.4c00061>
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D. et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, **13**(11), 2498–2504.
- Simmonds, S.E., Smith, J.F., Davidson, C. & Buerki, S. (2021) Phylogenetics and comparative plastome genomics of two of the largest genera of angiosperms, piper and peperomia (Piperaceae). *Molecular Phylogenetics and Evolution*, **163**(107229), 107229.
- Singh, S.K., Ashok, P.K., Olsen, C.E., Jha, A., Jain, S.C., Parmar, V.S. et al. (1996) Neolignans and alkaloids from piper Argyrophyllum. *Phytochemistry*, **43**(6), 1355–1360.
- Szóke, É., Lemberkovics, É. & Kursinszki, L. (2013) Alkaloids derived from lysine: piperidine alkaloids. In: *Natural products: phytochemistry, botany and metabolism of alkaloids, phenolics and terpenes*. Berlin Heidelberg: Springer, pp. 303–341.
- Tian, Y., Kong, L., Li, Q., Wang, Y., Wang, Y., An, Z. et al. (2024) Structural diversity, evolutionary origin, and metabolic engineering of plant specialized benzylisoquinoline alkaloids. *Natural Product Reports*, **41**(11), 1787–1810. Available from: <https://doi.org/10.1039/d4np00029c>
- Tsugawa, H., Rai, A., Saito, K. & Nakabayashi, R. (2021) Metabolomics and complementary techniques to investigate the plant phytochemical cosmos. *Natural Product Reports*, **38**(10), 1729–1759.

- Wang, F.-X., Zhu, N., Zhou, F. & Lin, D.-X.** (2021) Natural aporphine alkaloids with potential to impact metabolic syndrome. *Molecules*, **26**(20), 6117.
- Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y. et al.** (2016) Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nature Biotechnology*, **34**(8), 828–837.
- Wang, M., Jarmusch, A.K., Vargas, F., Aksenov, A.A., Gauglitz, J.M., Welldon, K. et al.** (2020) Mass spectrometry searches using MASST. *Nature Biotechnology*, **38**(1), 23–26. Available from: <https://doi.org/10.1038/s41587-019-0375-9>
- Ware, I., Franke, K., Frolov, A., Bureiko, K., Kysil, E., Yahayu, M. et al.** (2024) Comparative metabolite analysis of *Piper sarmentosum* organs approached by LC-MS-based metabolic profiling. *Natural Products and Bioprospecting*, **14**(1), 30.
- Watrous, J., Roach, P., Alexandrov, T., Heath, B.S., Yang, J.Y., Kersten, R.D. et al.** (2012) Mass spectral molecular networking of living microbial colonies. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(26), 10150.
- Weaver, K.M., Luker, R.G. & Neale, M.E.** (1984) Rapid quality control procedure for the determination of Scoville heat units and the detection of chillies in black pepper, via high-performance liquid chromatography. *Journal of Chromatography A*, **301**, 288–291.
- Wolfender, J.-L., Litaudon, M., Touboul, D. & Queiroz, E.F.** (2019) Innovative omics-based approaches for prioritisation and targeted isolation of natural products—new strategies for drug discovery. *Natural Product Reports*, **36**(6), 855–868.
- Wu, Y., Min, X., Zhuang, C., Li, J., Yu, Z., Dong, G. et al.** (2014) Design, synthesis and biological activity of piperlongumine derivatives as selective anticancer agents. *European Journal of Medicinal Chemistry*, **82**, 545–551. Available from: <https://doi.org/10.1016/j.ejmech.2014.05.070>
- Wu, Z., Deng, X., Hu, Q., Xiao, X., Jiang, J., Ma, X. et al.** (2021) *Houttuynia cordata* Thunb: an ethnopharmacological review. *Frontiers in Pharmacology*, **12**, 714694.
- Yang, X., Miao, X., Dai, L., Guo, X., Jenis, J., Zhang, J. et al.** (2024) Isolation, biological activity, and synthesis of Isoquinoline alkaloids. *Natural Product Reports*, **41**(11), 1652–1722. Available from: <https://doi.org/10.1039/d4np00023d>
- Zhang, L., Zhang, J.-H., Yang, S.-M., Tan, C.-H., Luo, H.-F. & Zhu, D.-Y.** (2010) Chemical constituents from the leaves of *Aglaias perviridis*. *Journal of Asian Natural Products Research*, **12**(3), 215–219.
- Zhou, K., Han, L., Li, W., Liu, S., Chen, T., Chen, J. et al.** (2024) Pipersarmenoids, new amide alkaloids from *Piper sarmentosum*. *Fitoterapia*, **177**, 106090. Available from: <https://doi.org/10.1016/j.fitote.2024.106090>
- Zhuang, T., Liang, J.-Y., Sun, J.-B., Wu, Y., Huang, L.-R. & Qu, W.** (2014) Secondary metabolites from *Saururus Chinensis* and their chemotaxonomic significance. *Biochemical Systematics and Ecology*, **56**, 95–98.
- Zuntini, A.R. & Carruthers, T.** (2024) Phylogenomics and the rise of the angiosperms. *Zenodo*. Available from: <https://doi.org/10.5281/ZENODO.10778206>
- Zuntini, A.R., Carruthers, T., Maurin, O., Bailey, P.C., Leempoel, K., Brewer, G.E. et al.** (2024) Phylogenomics and the rise of the angiosperms. *Nature*, **629**(8013), 843–850. Available from: <https://doi.org/10.1038/s41586-024-07324-0>