Accelerated characterization of active natural extracts with single-step microfractionation coupled to mass spectrometry and bioassays

Authors

Joshua David Smith[1&], Erik Bouchal[1&], Tito Damiani[1], Artur Jasanský[1], Alžběta Kadlecová[1], Martin Dračínský[1], Zdeněk Knejzlík[1], Tomáš Pluskal[1*], Téo Hebra[1*]

[1]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czechia
& These authors contributed equally to this work.
* To whom correspondence should be addressed

Abstract

Microfractionation is a prominent alternative for the discovery and characterization of bioactive natural products in complex extracts (e.g., plant, fungal, or microbial); however, the method faces limitations as it requires arduous scale-up for complete characterization. Furthermore, the integration of microfractionation data with metabolomic data can be a complex and tedious process. To circumvent current limitations, we developed a microfractionation method by creating an integrated, reproducible, milligram-scale, 96-deep-well plate microfractionation framework combining both biological and computational metabolomic readouts. Fractionation is coupled to a single quadrupole mass analyzer, which enables integration with high-resolution metabolomic outputs using easily adaptable Python scripts. To illustrate the feasibility, we applied our microfractionation characterization method to extracts of *Macleaya microcarpa* and *Macleaya cordata*, both moderately active against *Mycobacterium smegmatis*. From milligram-scale fractions, we directly linked bioactivity to isoquinoline alkaloids, in particular to chelerythrine. We confirmed these findings at fraction resolution with NMR structural elucidation in combination with computational metabolomic structure prediction. Pure compounds were structurally characterized from as little as 9 milligrams of crude extract and 80 micrograms of fractionated material. Our high-throughput, milligram-scale workflow reduces the need for scale-up and is broadly applicable to natural extracts and biological systems, thus allowing for rapid characterization of natural extracts.

Introduction

Natural products have been sought after for their complex molecular scaffolds and their pharmacological relevance [1,2]. The discovery of novel natural products often begins with crude extract biological screening, usually considering a specific target (e.g., receptor) or function (e.g., antioxidant). After screening, active crude extracts are fractionated and again tested for bioactivity—a process repeated until a pure active compound is isolated. This method of discovery is tedious and laborious [3] and often requires hundreds of grams of material [4,5]. Furthermore, bioactivity-guided methods are susceptible to both redundant or pan-assay interference compound (PAINS) isolation. PAINS produce false positive signals during the biological screening, resulting in the isolation of non-relevant compounds [6]. To reduce these problems, state-of-the-art computational metabolomic tools using data from ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) analysis have been developed to aid exploration of crude extracts [7–9].

Tandem mass spectrometry (MS/MS) fragments molecules into small ions representative of the original structure. Different *in silico* tools take MS/MS fragmentation spectra as input and provide output representative of the original molecule. For example, MS/MS fragmentation can be used to predict compound structure or molecular classes [7,10] or to create molecular networks of structurally related molecules [8,9]. MS/MS fragmentation can also be used for library matching to standards [11]. Often, these tools are used in combination during the discovery process by creating molecular networks with annotations of *in silico* structure predictions and library matches [12]. These tools only facilitate discovery through tentative annotation, thus requiring complete isolation and characterization of targets. Furthermore, since characterization often occurs at the crude extract scale, it lacks single-compound resolution when combined with biological activity data.

To obtain higher resolution of specific compound bioactivity, crude extracts can be microfractionated at milligram scale into many smaller, more pure fractions and then tested for biological activity [13–15]. While microfractionation methods can better prioritize targets for isolation, they are limited by requiring scale-up to gram amounts for compound isolation and extract characterization [15–18]. Furthermore, linking microfractionation data with metabolomic outputs often requires tedious method transfer between two different instruments [17,19]. To overcome these limitations, we sought to develop a microfractionation method that allows for complete characterization of an extract without the need for arduous scale-up and tedious method transfer. We accomplished this by creating a single-step microfractionation coupled with a single quadrupole mass analyzer, which allows for the seamless integration of MS/MS data. This enables mass matching and retention time alignment with MS/MS data without the need for highly tailored method transfer. By further combining the microfractionation with reproducible chromatography and multiple injections, only a few tens of milligrams of starting plant material are needed for NMR structural characterization of abundant bioactive compounds from complex mixtures. Ultimately, we created a simple single-step microfractionation workflow that reduces the laborious process of characterizing complex extracts (Figure 1).

Experimental Section

Chemicals

HPLC-grade methanol and acetonitrile were purchased from VWR International (Randor, Pennsylvania, USA), UPLC-grade methanol, water and acetonitrile were purchased from Thermo Fisher Scientific (Waltham, Massachusetts, USA), DMSO from Sigma-Aldrich (Darmstadt, Germany), Acetonitrile D3 99.8%D from Eurisotop (Saint-Aubin, Francie), Milli-Q water was obtained from a Milli-Q IQ 7000 water purification system (MilliporeSigma, Burlington, MA, USA), 99.0+% formic acid was purchased from Thermo Fisher Scientific (Waltham, Massachusetts, USA), Middlebrook 7H9 Broth Base (MilliporeSigma, Burlington, MA, USA), Bovine serum albumin from SERVA Electrophoresis GmbH (Heidelberg, Germany), Tyloxapol from Sigma-Aldrich (Darmstadt, Germany), Glycerol from PENTA (Prague, Czech republic), Resazurin sodium salt from Sigma-Aldrich (Darmstadt, Germany).

Extract library building and sample preparation for fractionation.

Plants were collected at the Botanical Garden of the Faculty of Tropical AgriSciences of the Czech University of Life Sciences in Prague. Plant tissues were immediately transferred to dry ice and then stored in – 80 °C. For UPLC-MS/MS analysis, plant material weighing at 50 ± 5 mg was placed into a 2.0 mL round-bottom Eppendorf tube (Hamburg, Germany). Samples were frozen with liquid nitrogen along with a metal bead for homogenization, which occurred in a TissueLyser for 30 seconds at 25 Hz. A mixture of ethanol:water 75:25 at 1000 µL was added to each sample followed by a 3-minute incubation at 40 °C on ThermoMixer C (Eppendorf, Hamburg, Germany). Samples were further agitated with the TissueLyser for 60 seconds at 25 Hz. Samples were centrifuged for 5 minutes at 14000 rpm with an Eppendorf centrifuge 5427 R (Hamburg, Germany) and supernatant was transferred to a clean 1.5 mL Eppendorf tube (Hamburg, Germany), evaporated under nitrogen gas flow, and then diluted in acetonitrile:water (50:50) mixture. Samples were sonicated for 15 seconds, vortexed for 30 seconds and then centrifuged for 5 minutes and 14000 rpm with an Eppendorf centrifuge 5427 R (Hamburg, Germany) before UPLC-MS/MS analysis.

After UPLC-MS/MS analysis, plant samples were dried under vacuum using a centrifugal evaporator at room temperature (Labconco, Kansas City, MO, USA). Ten milligrams of each lyophilized plant sample was extracted with 1 mL of methanol in a 2 mL microcentrifuge tube (Eppendorf, Hamburg, Germany) and with shaking at 900 rpm for 10 minutes at 40 °C using a ThermoMixer C (Eppendorf, Hamburg, Germany), followed by further agitation with a stainless steel ball using a TissueLyser II (Qiagen, Hilden, Germany) set at 30 Hz for 1 minute. Extracts were spun down with an Eppendorf centrifuge 5427 R at 14000 rpm and the supernatants were transferred to a clean microcentrifuge tube, evaporated under nitrogen flow, and resuspended to

a concentration of 10 mg/mL in dimethyl sulfoxide (DMSO) for subsequent biological testing. All samples from both extractions were stored at – 80 °C.

For fractionation and bioactivity screening of active plants, lyophilized samples of *Macleaya microcarpa and Macleaya cordata* were weighed out to 52.4 mg and 51.5 mg, respectively. Samples were extracted three times using 1 mL of methanol by vortexing at 2000 rpm, shaking at 900 rpm and 35 °C for 30 minutes. Supernatants from each extraction were pooled and evaporated under nitrogen flow. Extracts, 9.0 mg for *M. microcarpa* and 3.7 mg for *M. cordata*, were dissolved in methanol to 10 mg/mL for fractionation.

Bioactivity screening

A starter culture of *Mycobacterium smegmatis* strain mc^2155 was grown from a single colony in 7H9 medium supplemented with 0.5% glycerol and 0.1% tyloxapol overnight at 37 °C and 200 rpm. Overnight culture was used to inoculate fresh 7H9 medium to a final OD_{600} of 0.1. The inoculum was incubated overnight at 37 °C and 200 rpm, resulting in a bacterial stock in exponential phase with an OD_{600} of ~0.5. Into each well of a Nunclon Delta 96-Well MicroWell Plate (Thermo Fisher, Waltham, MA, USA), 100 µL of 7H9/ADS medium (1 part of albumin-dextrose-saline in 9 parts of 7H9 medium), 10 µL of bacterial stock diluted to an OD_{600} of 0.1, and 1 µL of plant extracts (10 mg/mL in DMSO) or 1 µL of fractions (50 µL of DMSO added into each fraction) were added. Kanamycin was used as a positive control with a working concentration of 0.45 mg/ml. Plates were covered with a lid and placed in a closed box with a water beaker and were incubated at 37 °C for 24 h. After the incubation period, 10 µL of 1 mM resazurin in phosphate-buffered saline was added to each well. After 8 h of incubation at 37 °C, quantitative fluorescence measurements were taken to determine cell metabolic activity. A Spark Multimode Microplate Reader (Tecan, Männedorf, Switzerland) with excitation wavelength set to 540 nm, emission to 590 nm, and bandwidth of 20 nm was used for plant extracts. An Infinite M1000 PRO (Tecan, Männedorf, Switzerland) set to excitation wavelength of 530 nm and an emission wavelength of 590 nm with a bandwidth of 5 nm was used for fractionation experiments.

HPLC Fractionation to pure compound with NMR confirmation

For high-performance liquid chromatography (HPLC) fractionation, the mobile phase was delivered by an Agilent 1290 Infinity II preparative binary pump (Santa Clara, CA, USA). A flow rate of 1 mL/min with a 38-minute gradient with solvents A (Milli-Q water + 0.1% formic acid) and solvent B (acetonitrile + 0.1% formic acid) was applied, starting at isocratic 10% B for 2 min, a gradient from 10% to 100% B over 28 min, isocratic at 100% B for 4 min, followed by a drop to 10% B for 4 minutes of equilibration. Extracts were separated on a XBridge BEH C18, 130 Å, 3.5 µm, 3x150 mm (Waters, Milford, MA, USA), with heating to 45 °C using MonoSLEEVE (Analytical Sales and Services, Flanders, NJ, USA). The analytes were detected by UV at 254 nm, 280 nm and a low-resolution single quadrupole MS with electrospray ionization (Agilent,

Santa Clara, CA, USA). Fractions were collected in a time-based manner every 22.5 seconds after the second minute by a 1290 Infinity II Preparative Open-Bed Fraction Collector (Agilent). MS analysis was achieved by active splitting of the primary flow post column with a ratio of 40:1 and a dilution ratio of 20:1. Split flow was pushed by a makeup flow of 0.500 mL/min using a 1260 quaternary pump (Agilent) with the same gradient and solvents as the main flow: 38-minute gradient, solvents A (Milli-Q water + 0.1% formic acid) and solvent B (acetonitrile + 0.1% formic acid), isocratic 10% B for 2 min, a gradient from 10% to 100% B over 28 min, isocratic at 100% B for 4 min, followed by an immediate drop to 10% B for 4 minutes. There was an additional passive splitting at the MS inlet to an ELSD; however, the ELSD was not used in this study. Analytes were ionized in positive mode, with a capillary voltage of 3000 V, a nebulizer gas pressure of 35 psi, a drying gas flow rate 12.0 L/s, a drying gas temperature 350 °C, a fragmentor voltage of 70 V, a scan speed of 743 mass units per second and a mass range of 100-800 *m/z*.

Crude extracts were injected at 10 mg/mL and 50 µL along with a plug solvent of 300 µL (10% ACN in 90% Milli-Q water). For *M. microcarpa*, 16 sequential injections were completed and for *M. cordata*, 7 injections. Fractions were collected into 96-deepwell plates (2.2 mL), yielding 4 plates for *M. microcarpa* and 2 for *M. cordata*. The mobile phase was evaporated on a SpeedVac SPD121P-230 (Thermo Fisher, Waltham, MA, USA) overnight at 40°C. Dried fractions were pooled together into a single plate with MeOH. 50 µL of DMSO was added to each well of the pooled fractions for bioactivity screening.

After bioassay screening, fractions of interest were checked on NMR or UPLC-MS/MS before NMR analysis by 1000x times dilution in UPLC-grade methanol and injected at 1 µL. Otherwise, DMSO was evaporated under nitrogen and then dissolved in 500 µL of acetonitrile-d_3 followed by 1H NMR analysis on a 400 MHz Bruker Avance III HD (Billerica, MA, USA). For some fractions, after NMR screening, 20µl of acetonitrile-d_3 was further screened on UPLC-MS/MS. Pooled fractions were ultimately evaporated and resuspended in 500 µL of acetonitrile-d_3 and analyzed on a 600 MHz Bruker Avance III HD spectrometer (Billerica, MA, USA) for ^{1}H, ^{13}C, COSY, HSQC, and HMBC experiments.

UPLC-MS/MS analysis and computational metabolomics

Crude extracts in 75:25 ethanol:water mixture were injected at 1 µL on an Orbitrap ID-X mass spectrometer (Thermo Fisher, Waltham, MA, USA) with a Vanquish Duo UHPLC system (Thermo Fisher). A flow rate of 0.350 mL/min was used with an ACQUITY UPLC BEH C18 column, 130 Å, 1.7 µm, 2.1 mm × 150 mm (Waters, Milford, MA, USA). The gradient used solvent A (water + 0.1% formic acid), solvent B (acetonitrile + 0.1% formic acid) and started with 0.5 min at 5% B, followed by a 15 min gradient to 100% B, isocratic elution from 15.5 min to 17.3 min and finally B was set to 5% again from 17.3 min to 19.3 min.

The electrospray ion source in positive mode was set to 3000 V, static gas mode, sheath gas at 50 arb, aux gas at 10 arb, sweep gas at 1 arb. The temperature of the ion transfer tube was set to 325 °C, and the vaporizer temperature to 350 °C. During MS^1 data collection, a scan

range of 100-1000 *m/z*, resolution of 60 000, 45% RF lens, and maximum injection time of 118 ms were set. For MS^2, data-dependent acquisition was used. The intensity threshold was set to 10 000. Fragmentation was done in fixed energy mode at 35% and an isolation window of 0.8 *m/z*. Cycle time was 0.6 s. Dynamic exclusion was used with a 5 ppm tolerance for 2 s after initial triggering by fragmented ions. Isotopes were also excluded. Apex detection used an expected peak width set to 4 s and a desired apex window of 50%. MS^2 used a resolution of 15 000, and the maximum injection time was 80 ms.

All raw data files in .raw format were converted to .mzML using MSconvert. Data were processed in MZmine 4.7.8 and SIRIUS 6.2.2, see supporting information for full processing details. All data were submitted to Zenodo and scripts for post-processing can be found at the microfractionation GitHub repository.

Results and discussion

Plant screening led to three moderately active extracts for microfractionation

We collected a small library of 143 different medicinally related plant species from the Botanical Garden of the Faculty of Tropical AgriSciences of the Czech University of Life Sciences in Prague. Lyophilized plant material was extracted with two different methods: one for UPLC-MS/MS analysis and one for biological testing and screening. Plants used can be found in Table S1. Methanolic extraction for bioassay was used on a subset of 76 species, while UPLC-MS/MS samples were extracted from all 143 species with a 75:25 ethanol:water mixture. Despite differences in extraction methods, we expect only minor divergences between the extracted metabolomes due to the similarity in polarity and our previous experience. Herein, crude methanolic extracts were evaporated and resuspended in DMSO at 10 mg/mL and were screened against *Mycobacterium smegmatis*.

Mycobacterium smegmatis is a non-pathogenic model for identifying potential antibacterials against slow-growing pathogenic *Mycobacterium spp*., such as *Mycobacterium tuberculosis* [20]. *M. tuberculosis*-based antibacterials are highly sought after due to the complex, hard-to-penetrate cell wall and unique way of obtaining resistance [21]. We incubated crude plant extracts overnight with *M. smegmatis* and indirectly measured cell death through the use of the established redox indicator resazurin [22]. While most plant extracts did not indicate any activity against the *M. smegmatis*, we identified three moderately active extracts (Figure S1): *Macleaya cordata*, *Macleaya microcarpa*, and *Ruta corsica*.

Based on UPLC-MS/MS data from crude plant extracts, in combination with SIRIUS [7] and CANOPUS [10] we attribute bioactivity to isoquinoline-like molecules in *Macleaya spp.* (Figure S2) and coumarins in *Ruta corsica* (Figure S3). Both coumarins [23–25] and isoquinolines [26–28] have been reported to possess antibacterial properties. Due to an excess of plant material and the proof-of-concept nature of the project, we proceeded with microfractionation using only *M. cordata* and *M. microcarpa*. We completed an exhaustive methanolic extraction with 51.5 mg of

lyophilized *M. cordata* and 52.4 mg of lyophilized *M. microcarpa*, which yielded 3.7 mg and 9.0 mg of extract, respectively. Crude extracts were then suspended in methanol at 10 mg/mL for fractionation.

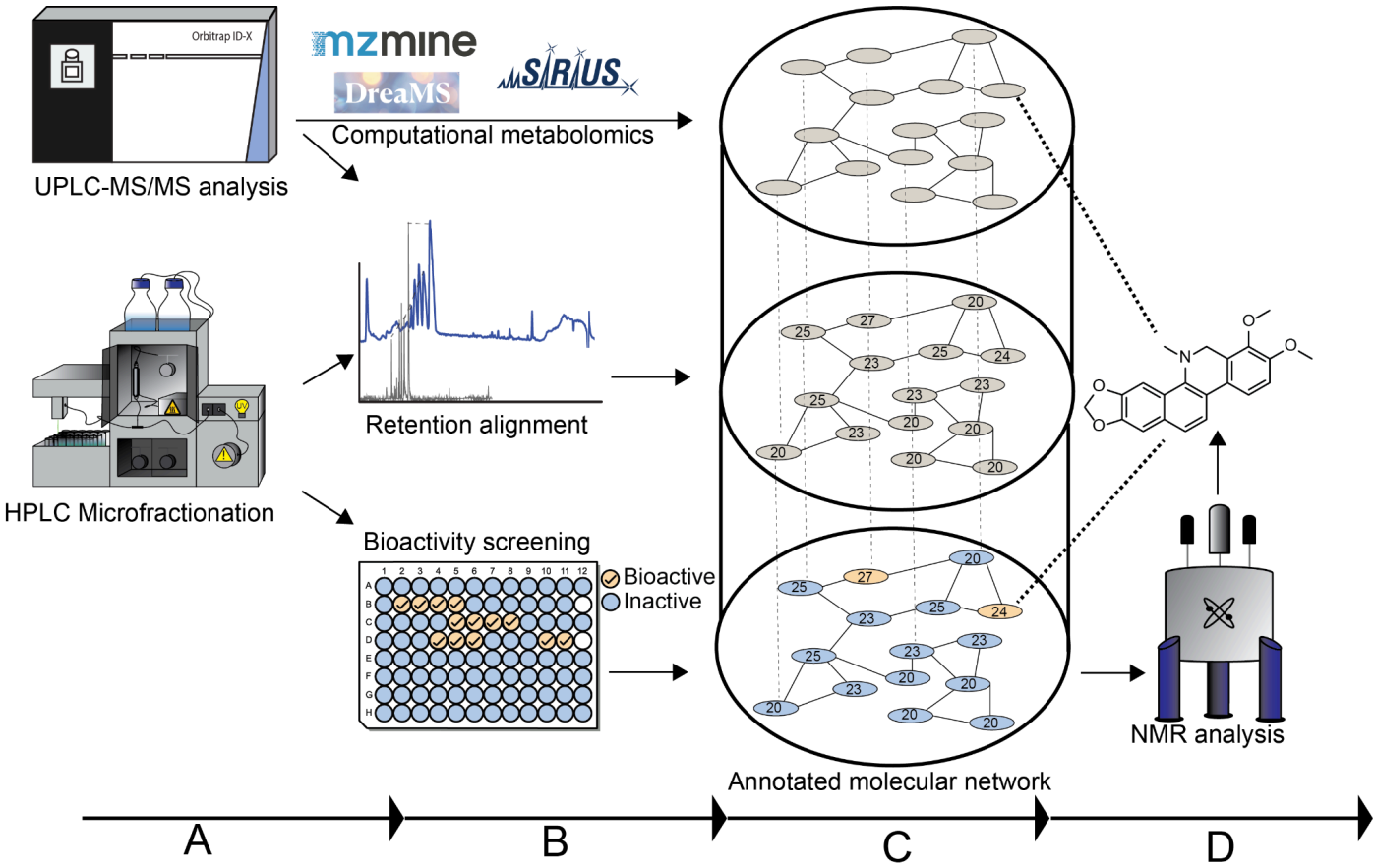

Figure 1. A) UPLC-MS/MS analysis and HPLC microfractionation of selected species can occur in parallel B) Retention times from UPLC-MS/MS are aligned to the HPLC-MS chromatogram. Fractions from crude extracts can also be screened for bioactivity. C) A molecular network is created using DreaMS, an embedding based similarity, from UPLC-MS/MS data. The network is further annotated with fraction number and fraction bioactivity. D) Selected fractions can be analyzed by NMR with MS/MS predictions to aid structure elucidation.

Automatized microfraction into a deep 96-well plate allowed for simultaneous bioscreening and NMR characterization with aid from *in silico* outputs

Our HPLC chromatographic gradient was based on a previous microfractionation publication [17]; however, we increased the number of fractions from 24 to 96. We performed multiple injections at 50 µL, which allowed for adequate peak shape and separation and to obtain sufficient amounts for both biological screening and NMR testing (Figure 2 and S4). We repeated injections into the same plate until the entire crude extract was fractionated or an additional plate was needed. Our method provided reproducible chromatography for both samples based on overlapping retention time and intensity for major features (Figure S5 and S6). Corresponding fractions in different plates were evaporated, pooled together into a single plate, and resuspended in DMSO for bioactivity screening.

By overlaying bioactivity with the microfractionation chromatogram, we found similar activity between most wells in both species of *Macleaya*(Figure 2 and S4), suggesting that no single molecule has potent activity, but rather there may be some synergistic effect of the bioactive class of molecules. To determine an estimated purity of each well, we used mzmine [29] to create extracted ion chromatograms (XICs) with a retention time restriction corresponding to each well. The resulting XICs provided an approximate number and amount of molecules in each well based on the peak area (Figure 2). Furthermore, we have uploaded scripts and mzmine batch parameters to GitHub for others to adapt this plot for their microfractionation pipelines.

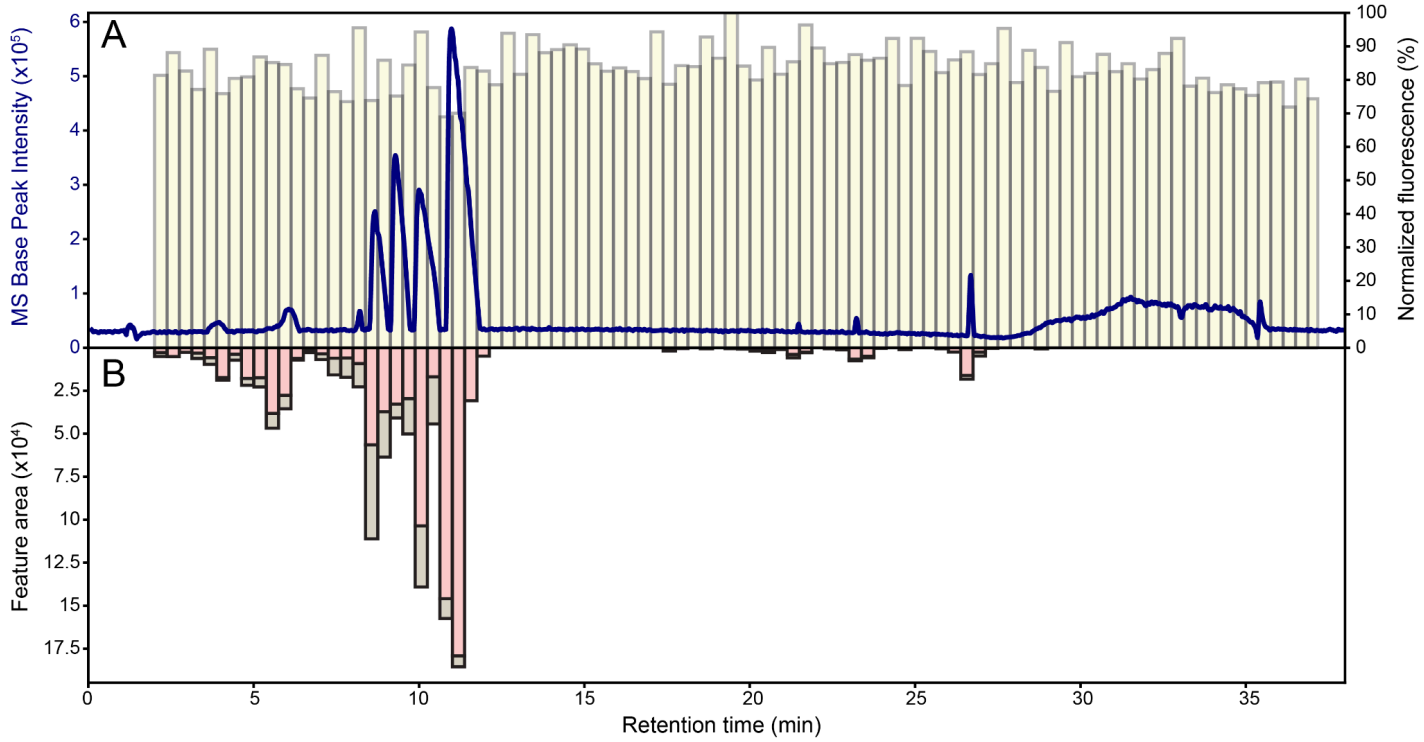

Figure 2. A) HPLC-MS trace from a fractionation injection of *M. microcarpa* with bioactivity shaded in the background. Fraction 25 produced both the most intense MS signal and the greatest response in the bioassay, about 30% more active than the least active fraction. B) The inverse plot is the total feature area of a metabolite found in a given fraction. Bars highlighted in red indicate the signal from the most abundant. Fraction 25 indicated that the most bioactive compound should be related to a single signal.

M. microcarpa fraction 25 was the most active, but also the most abundant and pure based on the *m/z* signal of 348. We leveraged this information to align the *m/z* of 348 and other features found in each well from microfractionation with the UPLC-MS/MS data and other *in silico* outputs, including SIRIUS structure and compound class prediction [7], spectral library matching, and molecular networking [8,29]. By using a combination of SIRIUS prediction and spectral library annotation, we attributed the activity of fraction 25 and the *m/z* of 348 to chelerythrine (**1**) (Figure 3). Furthermore, chelerythrine (**1**) has been reported to have antibacterial properties [30].

By combining fractions 24 and 25, we obtained 80 µg of residue with the expected *m/z* of 348. Ultimately, we obtained both 1D and 2D NMR confirmation of the molecule dihydrochelerythrine (**2**) (Figure S7-S10), despite expecting chelerythrine (**1**) (Figure 3). We attribute the identification of dihydrochelerythrine (**2**) to the reduction of the alkylpyridinium moiety in chelerythrine (**1**) [31] sometime after fractionation (Figure S11 and S12). We also confirmed that *M. cordata* contains dihydrochelerythrine (**2**) with retention time and MS/MS matching (Figure S13).

We further pooled less active fractions 15, 16, 17, and 18 that contained mostly an *m/z* of 322 from *M. cordata*, which yielded 190 µg of residue. 1D and 2D NMR analysis of the residue, with the aid of SIRIUS and MS/MS library matching (Figure 3), led to the structure of dehydrocheilanthifoline (**3**) (Figure S14-S17). Dehydrocheilanthifoline (**3**), also known as isogroenlandicine, has been previously identified in *M. microcarpa* [32] and more recently in *Corydalis bracteata* [33]. To the best of our knowledge, this compound has not been thoroughly investigated for antibacterial activity.

Lastly, we pooled fractions 9, 10, 11, and 12 from *M. microcarpa* with an *m/z* of 350 (Figure S18) and expected a structure similar to sanguinarine based on SIRIUS predictions (Figure 3). 1H NMR analysis resulted in a conflicting structure and mass of dihydrosanguinarine (**4**) and *m/z* of 334, respectively (Figure S19). Further investigation of UPLC-MS/MS analysis of both individual fractions 11 and 12 before NMR (Figure S20) and the analyzed pooled sample after NMR (Figure S21) revealed that the feature with an *m/z* of 350 likely degraded into dihydrosanguinarine (**4**). Dihydrosanguinarine (**4**) has been previously isolated from both *M. microcarpa* [34,35] and *M. cordata* [36].

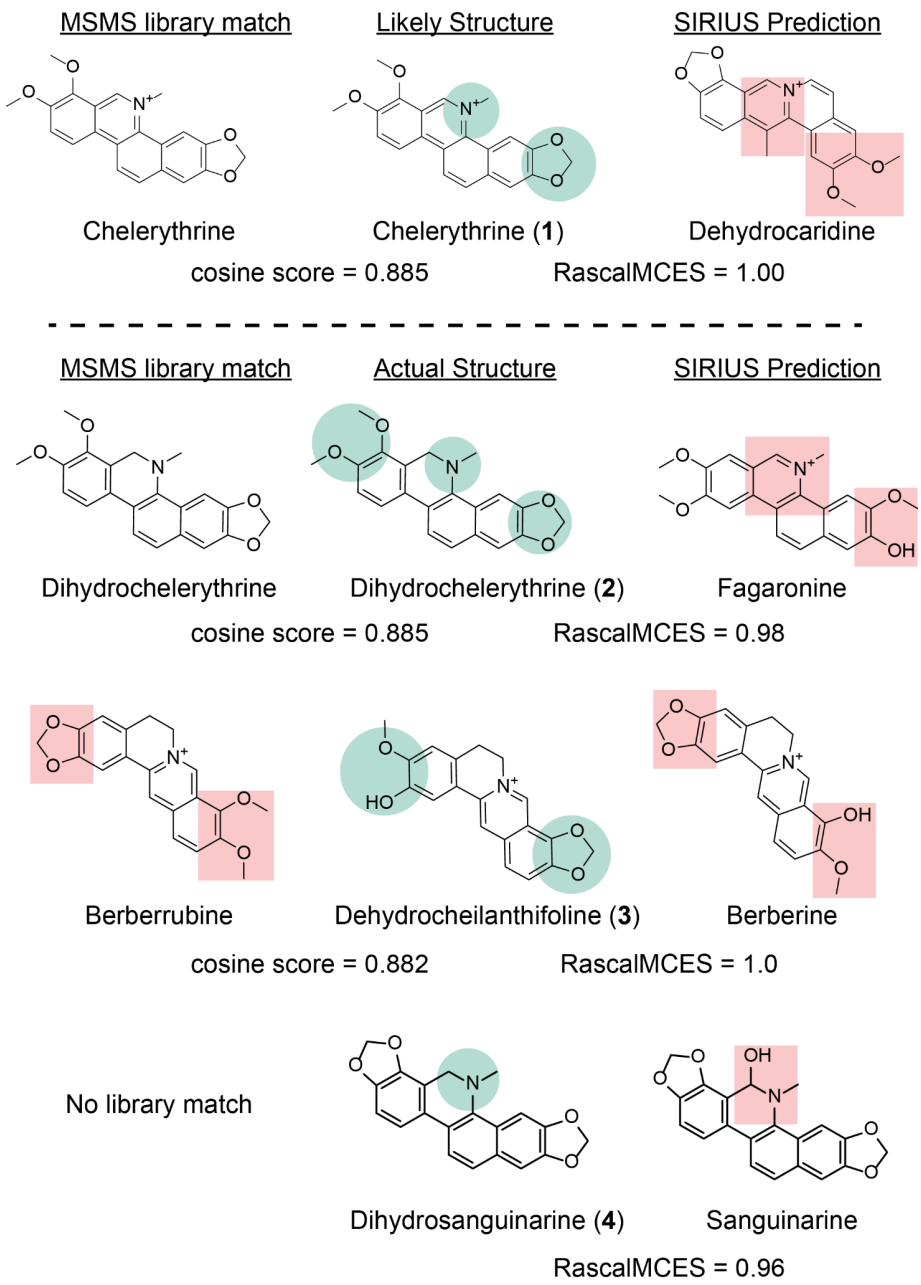

Figure 3. Comparison of obtained NMR structures to the MS/MS spectral library matches and SIRIUS-predicted structures. Green circles indicate regions with correct structures, while red squares highlight inaccuracies in structures from either SIRIUS or library matching. Using Johnson similarity with RascalMCES [37], we compared the structural similarity between the top SIRIUS-predicted and actual structures for dihydrochelerythrine (**2**). For dehydrocheilanthifoline (**3**), we opted for the SIRIUS prediction that better aligned with the MS/MS match. Lastly, we used the SIRIUS prediction related to *m/z* 350 to resolve the structure of dihydrosanguinarine (**4**). The high MCES score suggests a high structural similarity between the predicted structure and the actual molecule. Furthermore, MS/MS spectral matching provided the correct structure for dihydrochelerythrine (**2**).

Linked molecular networking further confirms the homogeneity of isoquinoline alkaloids as the active class of compounds

Lastly, we leveraged DreaMS embeddings [8] to create molecular networks in mzmine [38] using the entire crude extract library of 143 different plant species. The network was created with a focus on compounds directly connected to dihydrochelerythrine (**2**) (Figure 4)**.** The network is almost exclusive to *Macleaya spp.*, further suggesting that there may be a shared bioactive moiety. The activity of chelerythrine (**1**), fraction 25, may be linked to the alkylpyridinium moiety and the overproduction of reactive oxygen species [6,31]; however, due to the connection to many other features, it is unlikely that it was the sole driving force. Based on the homogeneity of the network and lack of major activity, further exploration is not warranted.

A DreaMS molecular network focused on dehydrocheilanthifoline (**3**) (Figure S22) also revealed direct connections to structurally related compounds from *Macleaya spp.*, but was much smaller in comparison. While other plants were found in connection, no significant activity or novelty would warrant further exploration. DreaMS, to the best of our knowledge, has been used only one other time for molecular networking [39]. Our network of structurally related compounds (Figure S22) illustrates DreaMS' ability to capture structural information using embeddings trained on MS/MS spectra. Ultimately, when all outputs are combined, the annotated networks (Figure 4 and Figure S22) strongly suggests that isoquinoline-like alkaloids are responsible for the activity caused by a specific shared moiety.

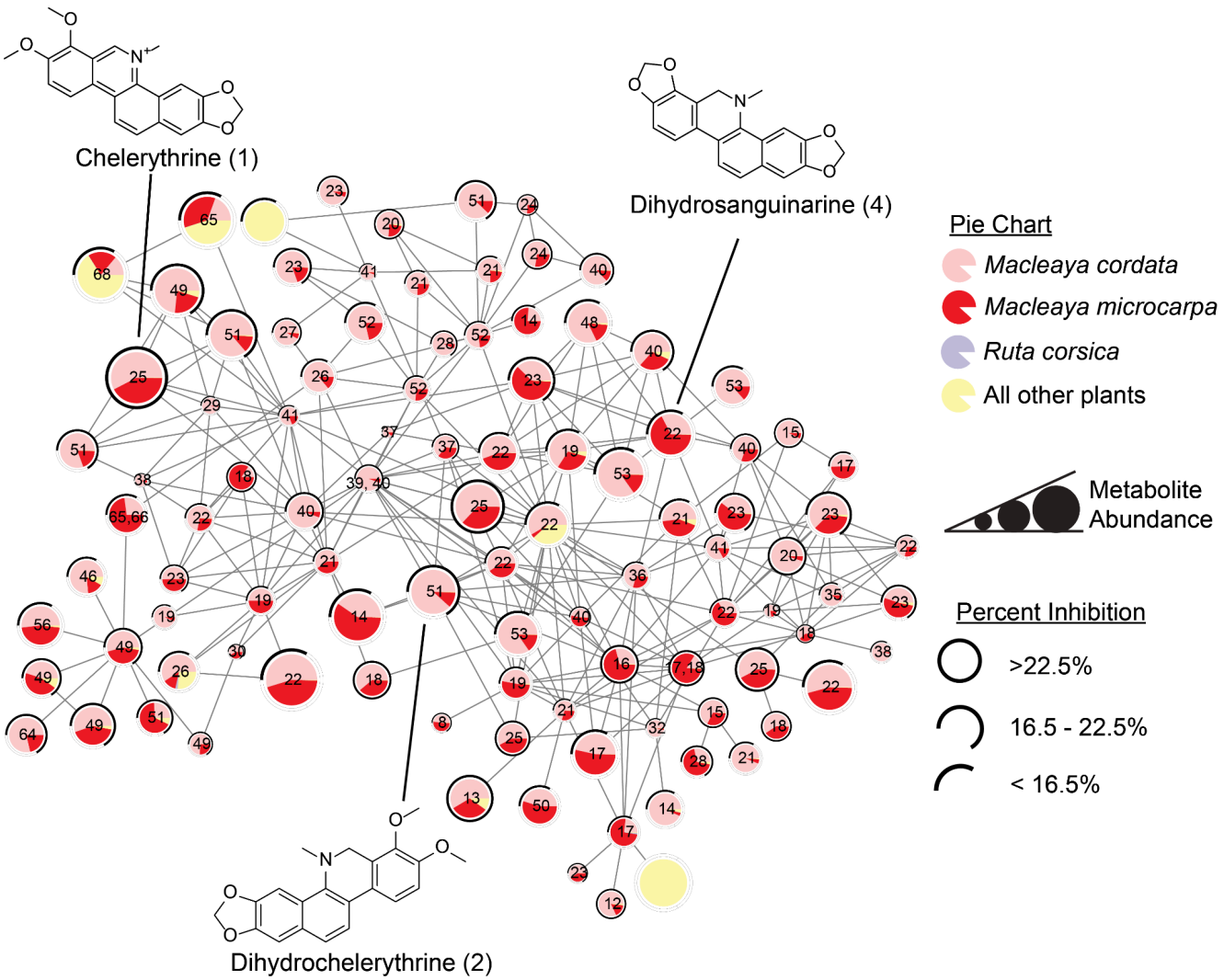

Figure 4. Molecular network built with connections starting from dihydrochelerythrine (**2**) and extending up to three closest neighbors. Node size is proportional to the log-sum intensity of the total peak area of a given feature found in all samples. The pie chart represents the peak area in either *M. cordata*, *M. microcarpa*, *R. corsica*, or a summation of the peak area for all other species. Nodes are also labeled with their corresponding fraction in *M. microcarpa*. Almost all features were exclusively found in bioactive plants.

Conclusions

Our novel method can be applied or easily adapted to any complex extract with a system capable of fractionating into a 96-deep-well plate and connecting to a mass spectrometer. Furthermore, we have made all scripts available for the complete adaptation of the workflow. In general, the method highlights how well computational tools and dereplication workflows combine to aid in the characterization of compounds and crude extracts. Despite our isolation of known compounds, our method demonstrates the reliability of annotations from MS/MS library matching and computational predictions. User expertise should also guide the selection of, and trust in, the annotations from the predictions or matches. Future studies using this method can consider employing small liquid-liquid prefractionation or other multi-solvent extraction for library creation to reduce the complexity of the crude extract. This suggested step, on a milligram scale, would be easily achievable and add little overall work. Ultimately, we were able to obtain the structures of three compounds from as little as 51.5 milligrams of crude material and 3.1 milligrams of subsequent extract. This method provides a template for quick isolation of target compounds or unknowns in a robust and high-throughput manner and allows for rapid characterization of crude extracts. We anticipate that our method will encourage the community to adopt microscale purification of natural products, saving time, effort, and material.

Data availability

All LC-MS/MS data has been uploaded to MassIVE:MSV000099695 . UPLC-MS/MS files, HPLC fractionation data and NMR data have been uploaded to Zenodo: 10.5281/zenodo.17495351. NMR data was also made available to the NMR-NP repository.

Supporting Information

Table S1: Contains the names of all plant species used in this manuscript, along with the UPLC-MS/MS file name and bioactivity.

Supplementary Information: Contains information on specific processing parameters used for MSconvert, mzmine, and SIRIUS processing. In addition, it contains all additional figures.

Acknowledgements

We thank Anežka Daníčková from the Botanical Garden of the Faculty of Tropical AgriSciences of the Czech University of Life Sciences in Prague for the continued support and collaboration. T.P. was supported by the Czech Science Foundation (GA CR) grant 21-11563M and by the European Union's Horizon Europe program (ERC, TerpenCode, 101170268). T.H. was supported by the European Union's Horizon Europe program under the Marie Skłodowska-Curie

grant agreement No. 101130799. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. Computational resources were provided by the e-INFRA CZ project (ID:90254), supported by the Ministry of Education, Youth and Sports of the Czech Republic.

References

(1) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83* (3), 770–803.
(2) Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; Supuran, C. T. Natural Products in Drug Discovery: Advances and Opportunities. *Nat. Rev. Drug Discov.* **2021**, *20* (3), 200–216.
(3) Muchiri, R. N.; van Breemen, R. B. Affinity Selection-Mass Spectrometry for the Discovery of Pharmacologically Active Compounds from Combinatorial Libraries and Natural Products. *J. Mass Spectrom.* **2021**, *56* (5), e4647.
(4) Pajouhesh, H.; Delwig, A.; Beckley, J. T.; Klas, S.; Monteleone, D.; Zhou, X.; Luu, G.; Du Bois, J.; Hunter, J. C.; Mulcahy, J. V. Discovery of Selective Inhibitors of NaV1.7 Templated on Saxitoxin as Therapeutics for Pain. *ACS Med. Chem. Lett.* **2022**, *13* (11), 1763–1768.
(5) Cicek, S. S.; Khom, S.; Taferner, B.; Hering, S.; Stuppner, H. Bioactivity-Guided Isolation of GABA(A) Receptor Modulating Constituents from the Rhizomes of Actaea Racemosa. *J. Nat. Prod.* **2010**, *73* (12), 2024–2028.
(6) Baell, J. B. Feeling Nature's PAINS: Natural Products, Natural Product Drugs, and Pan Assay INterference compoundS (PAINS). *J. Nat. Prod.* **2016**, *79* (3), 616–628.
(7) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16* (4), 299–302.
(8) Bushuiev, R.; Bushuiev, A.; Samusevich, R.; Brungs, C.; Sivic, J.; Pluskal, T. Self-Supervised Learning of Molecular Representations from Millions of Tandem Mass Spectra Using DreaMS. *Nat. Biotechnol.* **2025**, 1–11.
(9) Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; Martin H, C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; van der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. Feature-Based Molecular Networking in the GNPS Analysis Environment. *Nat. Methods* **2020**, *17* (9), 905–908.
(10) Dührkop, K.; Nothias, L.-F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M. A.; Petras, D.; Gerwick, W. H.; Rousu, J.; Dorrestein, P. C.; Böcker, S. Systematic Classification of Unknown Metabolites Using High-Resolution Fragmentation Mass Spectra. *Nat. Biotechnol.* **2020**, *39* (4), 462–471.
(11) Brungs, C.; Schmid, R.; Heuckeroth, S.; Mazumdar, A.; Drexler, M.; Šácha, P.; Dorrestein,

P. C.; Petras, D.; Nothias, L.-F.; Veverka, V.; Nencka, R.; Kameník, Z.; Pluskal, T. Efficient Generation of Open Multi-Stage Fragmentation Mass Spectral Libraries. **2025**.
(12) Damiani, T.; Smith, J.; Hebra, T.; Perković, M.; Čičak, M.; Kadlecová, A.; Rybka, V.; Dračínský, M.; Pluskal, T. Remarkable Diversity of Alkaloid Scaffolds in*Piper Fimbriulatum*. *bioRxiv*, 2024, 2024.12.10.627739. https://doi.org/10.1101/2024.12.10.627739.
(13) Mohotti, S.; Rajendran, S.; Muhammad, T.; Strömstedt, A. A.; Adhikari, A.; Burman, R.; de Silva, E. D.; Göransson, U.; Hettiarachchi, C. M.; Gunasekera, S. Screening for Bioactive Secondary Metabolites in Sri Lankan Medicinal Plants by Microfractionation and Targeted Isolation of Antimicrobial Flavonoids from Derris Scandens. *J. Ethnopharmacol.* **2020**, *246* (112158), 112158.
(14) Giera, M.; Heus, F.; Janssen, L.; Kool, J.; Lingeman, H.; Irth, H. Microfractionation Revisited: A 1536 Well High Resolution Screening Assay. *Anal. Chem.* **2009**, *81* (13), 5460–5466.
(15) Challal, S.; Buenafe, O. E. M.; Queiroz, E. F.; Maljevic, S.; Marcourt, L.; Bock, M.; Kloeti, W.; Dayrit, F. M.; Harvey, A. L.; Lerche, H.; Esguerra, C. V.; de Witte, P. A. M.; Wolfender, J.-L.; Crawford, A. D. Zebrafish Bioassay-Guided Microfractionation Identifies Anticonvulsant Steroid Glycosides from the Philippine Medicinal Plant Solanum Torvum. *ACS Chem. Neurosci.* **2014**, *5* (10), 993–1004.
(16) Zhao, Y.; Gericke, O.; Li, T.; Kjaerulff, L.; Kongstad, K. T.; Heskes, A. M.; Møller, B. L.; Jørgensen, F. S.; Venter, H.; Coriani, S.; Semple, S. J.; Staerk, D. Polypharmacology-Labeled Molecular Networking: An Analytical Technology Workflow for Accelerated Identification of Multiple Bioactive Constituents in Complex Extracts. *Anal. Chem.* **2023**, *95* (9), 4381–4389.
(17) Hell, T.; Rutz, A.; Dürr, L.; Dobrzyński, M.; Reinhardt, J. K.; Lehner, T.; Keller, M.; John, A.; Gupta, M.; Pertz, O.; Hamburger, M.; Wolfender, J.-L.; Garo, E. Combining Activity Profiling with Advanced Annotation to Accelerate the Discovery of Natural Products Targeting Oncogenic Signaling in Melanoma. *J. Nat. Prod.* **2022**, *85* (6), 1540–1554.
(18) Chang, S.; Li, Y.; Huang, X.; He, N.; Wang, M.; Wang, J.; Luo, M.; Li, Y.; Xie, Y. Bioactivity-Based Molecular Networking-Guided Isolation of Epicolidines A-C from the Endophytic Fungus Epicoccum Sp. 1-042. *J. Nat. Prod.* **2024**, *87* (6), 1582–1590.
(19) Guillarme, D.; Nguyen, D. T. T.; Rudaz, S.; Veuthey, J.-L. Method Transfer for Fast Liquid Chromatography in Pharmaceutical Analysis: Application to Short Columns Packed with Small Particle. Part II: Gradient Experiments. *Eur. J. Pharm. Biopharm.* **2008**, *68* (2), 430–440.
(20) Sparks, I. L.; Derbyshire, K. M.; Jacobs, W. R., Jr; Morita, Y. S. Mycobacterium Smegmatis: The Vanguard of Mycobacterial Research. *J. Bacteriol.* **2023**, *205* (1), e0033722.
(21) Sachan, R. S. K.; Mistry, V.; Dholaria, M.; Rana, A.; Devgon, I.; Ali, I.; Iqbal, J.; Eldin, S. M.; Mohammad Said Al-Tawaha, A. R.; Bawazeer, S.; Dutta, J.; Karnwal, A. Overcoming Mycobacterium Tuberculosis Drug Resistance: Novel Medications and Repositioning Strategies. *ACS Omega* **2023**, *8* (36), 32244–32257.
(22) Riss, T.; Moravec, R.; Niles, A.; Duellman, S.; Benink, H.; Worzella, T.; Minor, L. Cell Viability Assays. *Assay guidance manual [Internet]* **2016**. https://doi.org/10.1007/978-1-4939-6960-9.
(23) Borges, F.; Roleira, F.; Milhazes, N.; Santana, L.; Uriarte, E. Simple Coumarins and Analogues in Medicinal Chemistry: Occurrence, Synthesis and Biological Activity. *Curr. Med. Chem.* **2005**, *12* (8), 887–916.
(24) Kayser, O.; Kolodziej, H. Antibacterial Activity of Simple Coumarins: Structural Requirements for Biological Activity. *Z. Naturforsch. C* **1999**, *54* (3-4), 169–174.
(25) Sridevi, D.; Sudhakar, K. U.; Ananthathatmula, R.; Nankar, R. P.; Doble, M. Mutation at G103 of MtbFtsZ Altered Their Sensitivity to Coumarins. *Front. Microbiol.* **2017**, *8*, 578.
(26) Jin, S.; Sato, N. Benzoquinone, the Substance Essential for Antibacterial Activity in

Aqueous Extracts from Succulent Young Shoots of the Pear Pyrus Spp. *Phytochemistry* **2003**, *62* (1), 101–107.
(27) Villar, A.; Mares, M.; Rios, J. L.; Canton, E.; Gobernado, M. Antimicrobial Activity of Benzylisoquinoline Alkaloids. *Pharmazie* **1987**, *42* (4), 248–250.
(28) Sai, C.-M.; Li, D.-H.; Li, S.-G.; Han, T.; Guo, Y.-Z.; Pei, Y.-H.; Bai, J.; Jing, Y.-K.; Li, Z.-L.; Hua, H.-M. Racemic Alkaloids from Macleaya Cordata: Structural Elucidation, Chiral Resolution, and Cytotoxic, Antibacterial Activities. *RSC Adv.* **2016**, *6* (47), 41173–41180.
(29) Heuckeroth, S.; Damiani, T.; Smirnov, A.; Mokshyna, O.; Brungs, C.; Korf, A.; Smith, J. D.; Stincone, P.; Dreolin, N.; Nothias, L.-F.; Hyötyläinen, T.; Orešič, M.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X.; van der Hooft, J. J. J.; Schmid, R.; Pluskal, T. Reproducible Mass Spectrometry Data Processing and Compound Annotation in MZmine 3. *Nat. Protoc.* **2024**, 1–45.
(30) He, N.; Wang, P.; Wang, P.; Ma, C.; Kang, W. Antibacterial Mechanism of Chelerythrine Isolated from Root of Toddalia Asiatica (Linn) Lam. *BMC Complement. Altern. Med.* **2018**, *18* (1), 261.
(31) Matkar, S. S.; Wrischnik, L. A.; Hellmann-Blumberg, U. Production of Hydrogen Peroxide and Redox Cycling Can Explain How Sanguinarine and Chelerythrine Induce Rapid Apoptosis. *Arch. Biochem. Biophys.* **2008**, *477* (1), 43–52.
(32) Savina, A. A.; Sheichenko, V. I.; Tolkachev, O. N. Study of by-Products Obtained during Sanguiritrine Production fromMacleaya Microcapra. *Pharm. Chem. J.* **1997**, *31* (8), 422–423.
(33) Whaley, A. O.; Whaley, A. K.; Toporkova, V.; Fock, E.; Rukoyatkina, N.; Smirnov, S. N.; Satimov, G. B.; Abduraxmanov, B. A.; Gambaryan, S. Bracteatinine and Isogroenlandicine, Two New Isoquinoline Alkaloids Isolated from Corydalis Bracteata and Their Effect on Platelet Function. *Fitoterapia* **2023**, *171* (105697), 105697.
(34) Sai, C.-M.; Li, B.-J.; Zhang, Z.; Wang, H.-N.; Gong, H.; Lun, X.-Y.; Tian, M.-J.; Yang, M.-Y.; Hua, H.-M. Cytotoxic Alkaloids from the Fruit Pods of Macleaya Microcarpa. *Fitoterapia* **2023**, *164* (105378), 105378.
(35) Deng, A.-J.; Qin, H.-L. Cytotoxic Dihydrobenzophenanthridine Alkaloids from the Roots of Macleaya Microcarpa. *Phytochemistry* **2010**, *71* (7), 816–822.
(36) Takao, N.; Kamigauchi, M.; Okada, M. Biosynthesis of Benzo[*c*]phenanthridine Alkaloids Sanguinarine, Chelirubine and Macarpine. *Helv. Chim. Acta* **1983**, *66* (2), 473–484.
(37) Raymond, J. W. RASCAL: Calculation of Graph Similarity Using Maximum Common Edge Subgraphs. *Comput. J.* **2002**, *45* (6), 631–644.
(38) Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrlund, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M.; Sarvepalli, A.; Zhang, Z.; Fleischauer, M.; Dührkop, K.; Wesner, M.; Hoogstra, S. J.; Rudt, E.; Mokshyna, O.; Brungs, C.; Ponomarov, K.; Mutabdžija, L.; Damiani, T.; Pudney, C. J.; Earll, M.; Helmer, P. O.; Fallon, T. R.; Schulze, T.; Rivas-Ubach, A.; Bilbao, A.; Richter, H.; Nothias, L.-F.; Wang, M.; Orešič, M.; Weng, J.-K.; Böcker, S.; Jeibmann, A.; Hayen, H.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X.; Pluskal, T. Integrative Analysis of Multimodal Mass Spectrometry Data in MZmine 3. *Nat. Biotechnol.* **2023**, *41* (4), 447–449.
(39) Camunas-Alberca, S. M.; Bartolini, F.; Cermakova, H.; Sánchez-Illana, Á.; Guaita, D. P.; Cañadas-Miquel, P.; Gil-de-la-Fuente, A.; Galano, J.-M.; Durand, T.; Gradillas, A.; Others. Decoding the Oxylipin Chemical Space Using Ion Identity Molecular Networking. **2025**.