Metabolite fingerprinting of *Urospatha sagittifolia* (Araceae) tubers at different growth stages by multi-platform metabolomics and molecular networking

*Jefferson V. Pastuña[a], Cristian D. Quiroz-Moreno[b], Evencio J. Medina[c], Jessica L. Cooperstone[b,d], Matteo Radice[e], Maria C. Peñuela-Mora[f], José R. Almeida[g], Noroska G. Salazar Mogollón[g]**

[a]Laboratorio de Productos naturales, Universidad Regional Amazónica Ikiam, Km 7 Via Muyuna, Tena, Napo, Ecuador.

[b]Department of Horticulture and Crop Science, The Ohio State University, Columbus, OH 43210, USA.

[c]School of Industrial, Computer and Aeronautical Engineering, University of León, Campus de Vegazana, 24071, León, Spain.

[d]Department of Food Science and Technology, The Ohio State University, Columbus, OH 43210, USA.

[e]Department of Biosciences, Biotechnology and Environment, University of Bari Aldo Moro, Via E. Orabona, 4, 70125, Bari, Italy

[f]Grupo de ecosistemas tropicales y cambio global, Universidad Regional Amazónica Ikiam, Km 7 Via Muyuna, Tena, Napo, Ecuador.

[g]Biomolecules Discovery Group, Universidad Regional Amazónica Ikiam, Km 7 Via Muyuna, Tena, Napo, Ecuador

*Email: noroska.salazar@ikiam.edu.ec; Phone: (06)-299-9160/ (06)-370-0040 ext: 308

Preprint not peer reviewed

Abstract

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Urospatha sagittifolia is a medicinal plant with antivenom, antihemorrhagic and anti-inflammatory properties commonly used by vulnerable Amazon indigenous communities to tackle venomous snakebites and their life-threatening consequences. Despite its pharmacological potential as a rich source of drug candidates, its metabolomic profile remains unknown. In this context, this study integrated mass spectrometry-based metabolomics, multivariate analysis, and molecular networks to uncover the chemical composition of *U. sagittifolia* and its dynamic metabolic changes during three plant growth stages (seedling, juvenile and adult). In general, 50 metabolites were identified in *U. sagittifolia* tubers by LC-MS (43 metabolites) and GC-MS (seven metabolites). Multivariate analysis using LC-MS showed that the relative concentrations of most of the identified metabolites were higher in seedlings or juveniles than in adults. On the other hand, GC-MS analysis showed that methyl palmitate and methyl stearate were the most abundant in the early growth stages, whereas allantoic acid and palmitic acid prevailed as the plant matured. In summary, this is the first metabolomics-centered mining of *U. sagittifolia* compositional diversity focusing on chemical-level variability. These valuable findings offer a temporal view of metabolic changes during plant growth stages, which is useful for future bioprospecting, biological screening, and purification of metabolite-based therapeutics.

Keywords: Metabolomics, Exploratory analysis by multi-platforms, metabolomic fingerprinting

Highlights

- Explore the metabolite changes in *Urouspatha sagittifolia* tubers across various growth stages
- Utilize a multi-platform approach involving metabolomics and molecular networking for comprehensive analysis of the metabolomic profile of *Urouspatha sagittifolia*
- Exploratory analysis in the investigation of metabolites in *Urouspatha sagittifolia*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

1. Introduction

Natural product-based research has been a successful niche with wide opportunities and significant contributions to the challenging drug discovery process.[1,2] The classical screening of rich and diverse natural libraries relying on bioactivity-guided fractionation is time-consuming and expensive leading to insufficient investment from the pharmaceutical market. However, recently interest in natural products has come onto the scene again due to several factors, such as failure of conventional treatments, antimicrobial resistance, the Coronavirus pandemic and the development of *in silico* and omics approaches.[3,4] High biodiversity areas, such as the Amazon, are of primary interest for bioprospection programmes and many initiatives have focused on pharmacologically active plant-derived natural products.[5]

Plants represent one of the most well-studied natural sources of therapeutic agents. Their biochemical content, physiology and synthesis of metabolites are highly influenced by multiple factors beyond genetics, such as lighting, temperature, soil parameters, carbon dioxide availability, ozone, geography, altitude, seasonal shifts, and climate change.[6,7] In addition, the phenological stage of plants constitutes another important driver. A deeper understanding of the compositional variability is crucial to design strategies to modulate the synthesis and further pharmaceutical applications.[8–10] To gain insights into the chemical diversity of plants in different growth stage, metabolomics-driven approaches have emerged as useful tools.

Metabolomics strategies has evolved to identify a snapshot of the secondary metabolites, their intermediates and the content of molecules that can be extracted, separated and characterized in a complex mixture obtained from natural products.[11] Untargeted metabolomics combines modern analytical techniques and chemometric analysis and provides a promising platform for the metabolic variation of a natural system under diverse conditions or between individuals.[12] The integrated use

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

of these tools has been promoted by the resurgent growing interest in finding novel chemical scaffolds from plant extracts for drug discovery initiatives. Historically, phytomedicine is widely used in indigenous communities in Latin America. A classical use consists in the management of clinical manifestations associated with snakebite envenoming, which remains an important medical emergency and the rural tropics.[13,14] *Urospatha sagittifolia* is one of the plant extracts commonly employed for this purpose.

Urospatha sagittifolia (Rudge) Schott (Araceae) is a shallow-water emergent plant mainly distributed in the Amazon region from Venezuela, Brazil, Ecuador and Peru.[15] An *in vitro* study has confirmed the antivenom potential of the ethanolic extract of this plant characterized by neutralization of venom-induced lethality, haemorrhage and inflammation.[16] The traditional colorimetric screening of the phytochemicals from the extract of *U. sagittigolia* pointed out the presence of alkaloids, saponins, terpenoids, tannins, coumarins and phenolics compounds. However, to date, there are no reports available exploring the profile chemical diversity in different growth stages of the plant, leaving much to clarify its composition and temporal changes in secondary metabolite production.

With this in mind, to identify variability in the chemical constituents of *U. sagittifolia*, we considered three growth stages based on plant height: seedling (0.20 m-height), juvenile (0.20-1.15 m-height) and adult (> 1.15 m-height with reproductive organs). We employed mass spectrometry-based tools to unveil the chemical profile. Multivariate and molecular networking analysis were used to extensively interrogate our data, allowing us to report the chemical variation of *U. sagittifolia* for the first time based on its growth stage.

2. Materials and methods

2.1 Urouspatha sagittifolia tissue collection

Samples at different growth stages of *U. sagittifolia* were collected in Puerto Misahualli (1°3´20.49"S 77°40´0.65" W), Napo, Ecuador at 460.6 meters above sea level. *U. sagittifolia* botanical sample was

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

deposited in Herbario Nacional del Ecuador (QCNE) herbarium with voucher number QCNE-030-2021. Taking it account the previous study demonstrating the neutralizing abilities of metabolites from *U. sagittifolia*, we chosen the same Amazon geographic region sample collection. Three developmental stages were considered in this study based on the specimen height: (a) seedlings (0-0.20 m), (b) juvenile (> 0.20-1.15 m), and (c) adults (> 1.15 m with present inflorescences). For the specimens classified as adults, we collected samples only if they had inflorescences present, and we did not discriminate by flower stage (as it has been reported that the same inflorescence can have flowers at different stages of development) [33]. Since this species is used by local communities in traditional medicine applications, we reduced our sample collection to three replicates per development stage to not jeopardize the availability of this resource for local communities.[16] In summary, we collected a total of nine samples: three seedlings, three juveniles, and three adults. After plants were identified in the field, they were measured using a measuring tape starting from the soil level to the highest aerial part of the plant. Plants were carefully extracted from the native substrate and placed on dry ice to promote metabolic quenching. Then, tubers were transported to the laboratory and stored at -80 °C for further analysis.

2.1.1 Sample Preparation.

Tubers at each growth phase were thawed at room temperature, washed thoroughly with distilled water, and heat-dried for seven days at 30 °C (ESCO ISOTHERM, OFA-54-8). The dried roots were uniformly ground using a manual mill until reaching a fine powder. The ground plant tissues were soaked in methanol 1:10 (w:v) for 5 days.[34] The solvent was recovered and then filtered through 15 μm cellulose filter paper (MicroLAB Scientific). The filtered extracts were subsequently evaporated using a rotary evaporator (Buchi, R-300). The resulting extracts were concentrated using a vacuum dryer to completely remove the extraction solvent (Geneva, Mi Vac Duo). Finally, the powdered extracts were stored at -20 °C until further use.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

2.2 Quantification Of Total Phenolic Content

The phenolic content of each extract was analyzed with the Folin-Ciocalteu colourimetric method,[35] using gallic acid as standard: 100 µL (5 mg of extract in 1.5 mL of MeOH: H_2O 1:1 v/v) of the sample was mixed with 200 µL of Folin Reagent (0.1 M) protecting it from light, then 700 µL of Na_2CO_3 (7.5%) was added and left to stand for 120 min. The absorbance of the mixture was measured at 765 nm in a spectrophotometer (Shidmazu, UV-1280), a calibration curve for gallic acid was prepared in the range between 25-150 µg/L and the results were expressed as µg of gallic acid per mg of extract.

2.3 Quantification Of Total Flavonoid Content

The total flavonoid content was determined by the aluminium trichloride ($AlCl_3$) method,[36] using quercetin as a standard: 1 mL of samples (5 mg of extract in 1.5 mL of MeOH) were mixed with 1 mL of 2% $AlCl_3$ and incubated at room temperature leaving to stand for 5 min. The absorbance was measured at 438 nm in a spectrophotometer (Shidmazu, UV-1280). A calibration curve was prepared in the range between 5-50 µg/mL and the results were expressed as µg of quercetin per mg of extract.

2.4 Untargeted Metabolomics Analysis

After conducting the total phenolic and total flavonoid analysis, extracts were reconstituted in H_2O: MeOH 1:1 (v/v), and the sample was diluted using a 1:100 dilution factor. Finally, the diluted samples were filtered using a 0.22 µm PTFE filter to avoid any solids to before mass spectrometry analysis.

2.4.1. Liquid Chromatography-Mass Spectrometry Analysis

Metabolomics data acquisition. *U. sagittifolia* samples were analyzed using an Agilent 1290 Infinity II series UHPLC coupled to an Agilent 6545 quadrupole time of flight mass spectrometer with electrospray ionization (ESI-QTOF-MS) (Agilent, Santa Clara CA) using an optimized method for untargeted metabolomics.[37]

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

An equal aliquot volume of each sample extract was used to create a pooled quality control (QC) sample to conduct a downstream data quality control.[38] Furthermore, solvent and process blanks were also included in the injection list to remove later features detected in blanks. The analysis was performed with an injection volume of 3 µL per sample. Reverse phase chromatography was conducted with a Waters (Milford, MA, USA) Acquity UPLC HSS T3 column (2.1 x 50 mm, 1.8 µm particle size) maintained at 40°. Full-scan spectral data were collected separately in both positive and negative ionization modes for more comprehensive metabolite coverage.[39] Mobile phases consisted of water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B), flowing constantly at 0.5 mL/min. The gradient was set as follows: 0-0.5 min 0% B; 0.5-8 min increase to 100% B; 8-9 min hold 100% B; 9.01-10.0 isocratic at 0% B.

Settings of the mass spectrometer (MS) were as follows: gas temp 350 °C, gas flow 10 L/min, nebulizer 35 psi, sheath gas temp 375 °C, sheath gas flow 11 L/min, VCap 4500 V, nozzle voltage 500 V, fragmentor 100, skimmer1 45, octopoleRFPeak 750, and scan rate of 2 spectra/s with a mass range of 100-1700 m/z.

Using the same MS settings described above, the MS/MS data of the pooled QC samples was obtained with the method described below. Data were acquired with MassHunter (Agilent Technologies), in iterative data-dependent acquisition mode. These experiments were conducted in positive and negative ionization modes with five injections for each collision energy, 20 and 40 eV. Parameters for AutoMS2 scans were as follows: MS minimum range 40 m/z, MS maximum range 1700 m/z, MS scan rate 3 spectra/s, MS/MS minimum range 40 m/z, MS/MS maximum range 1700 m/z, MS/MS scan rate 1 spectra/s, isolation width narrow (~1.3 amu), and decision engine advanced. Terms for precursor selection were max precursors per cycle 2, threshold (absolute) 10.000, threshold (relative)(%) 0.100, precursor abundance-based scan speed – yes, target 100.000 counts/spectrum, use MS/MS accumulation time limit – no, use dynamic precursor rejection – no, purity stringency 100%, purity cutoff 30%, common isotope model for small molecules, active exclusion enabled –

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

yes, active exclusion excluded after 2 spectra, active exclusion released after 0.12 min, and sort precursors by abundance only.

2.4.1.1 Metabolomics feature identification.

Agilent files (*.d) of raw spectral data were converted to (*.mzML) using ProteoWizard.[40] Raw data was deposited in the database MetaboLights (MTBLS8067).[41] First, MZmine 3.3.0 software[42] was used to find the optimal parameters for peak deconvolution (Supplemental Table S1-7) with a pipeline including mass detection using wavelets (Automated Data Analysis Pipeline, ADAP),[43] including feature detection, isotope grouping, alignment, gap filling, and filtering. Then, the MS-DIAL 4.9.221218v software was set with optimal parameters reached in MZmine (Supporting Information Table S8-S11). Also, the "exclude process blank features" present in samples available MS-DIAL option was used (maximum sample intensity/average blank intensity ratio > 5). Finally, the MS-DIAL feature table (*.txt) was exported for statistical analysis with the notame R package (https://github.com/antonvsdata/notame).[44,45]

2.4.1.1.1.Level 1 metabolite identification.

To report metabolite identification, we used the proposed levels by the metabolomics standard initiave.[23] For level 1 identification, we compared the spectrometric in our samples versus an in-house compound library, using MS/MS and MS spectrometric matching. In both cases, the retention time match was required.

For the first approach, we matched the MS/MS spectrometric collected with the iterative DDA method on QC samples, against an in-house MS/MS library (PhenolicsDB).[46] The parameters used were: retention time tolerance (0.115 min.), minimum MS/MS match score (0.70) (Agilent MS/MS licensed score), and mass tolerance (7 ppm). Due to the limitations of the DDA acquisition method, such as low abundance metabolites not reaching the minimum abundance threshold to trigger MS/MS

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

fragmentation, or isobaric metabolites eluting in a narrow retention time window, we manually inspect for metabolites that are in the in-house library at level MS1.

In the manual inspection step, we used two criteria: (a) retention time matches with the authentic standards, and (b) isotopic pattern matches with at least two isotope ions present.

2.4.1.1.2. Level 2 metabolite identification

Agilent files (*.d) of raw MS/MS data were converted to (*.mzML) format using ProteoWizard applications.[40] Raw spectral data were deconvoluted in MZmine 3.3.0 (Supporting Information Table S5-11) and MS-DIAL 4.9.221218 (Supporting Information Table S3-4). Negative and positive ionization modes were processed together (only in MZmine), and separately for each of two collision energies (20 and 40 eV), to obtain a peak list for putative identity generation, via different library searches algorithm available in Global Natural Products (GNPS) platform.[47] Three GNPS algorithms were used: Classical Molecular Networking, Feature-Based Molecular Networking (FBMN) and Library Search (Supporting Information Table S12).

2.4.1.2 Molecular Networking.

The MS/MS data deconvoluted from MS-DIAL was transferred to the GNPS Molecular Networking server to generate the chemical map in positive (ID=2a3baf635b54492bb5e41d35b2b2bd79) and negative ionization mode (ID=6c9421f32ed142f59239db0df2cc13fb) according to the GNPS documentation.[48] Also, raw data was deposited in the GNPS/MassIVE repository (MSV000090922). The molecular network was generated so that the mass tolerance of precursor ions was 0.005 Da. The mass variances of ion fragments for each group of acquired MS/MS spectra were set as 0.006 Da for clustering. The connections between nodes were formed only if the cosine score was above 0.7 and with a minimum correspondence of 3 peaks in the fragmentation spectrum. The molecular network spectra were then compared to the GNPS spectral libraries, where the same data parameters were applied to the sample spectra.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Besides the feature-based molecular network analysis, we also used MS2LDA[49] and merge polarities pipelines available in GNPS, to propagate the molecular networks. In the case of MS2LDA, we submitted a separate task for each ionization polarity. We selected all MotifDBs available except from the Urine database. For both MS2LDA tasks, the overlap score threshold was set to 0.3, the probability value threshold was set to 0.1, and the topX in node was set to 5. Finally, to connect features that were detected in both ionization polarities, we used the "merge polarity" workflow by providing the task IDs from the feature-based molecular network for each polarity. This workflow accounts for the m/z shift due to the ionization polarity and retention time shift. We used the merged polarity network to propagate annotation based on two criteria: (a) the mass difference connecting two nodes must be 2.012 Da, and (b) the retention time shift must not be greater than 3 seconds (0.05 min), to propagate annotation based on polarity similarity. Finally, the molecular network was pruned and visualized in the Cytoscape software.[50]

2.4.2 Gas Chromatography-Mass Spectrometry Analysis

Methanolic tuber extract samples (5.0 mg) were dissolved in 1.0 mL of analytical-grade methanol. Samples were centrifuged ((Sorvall ST 40R, Thermo Scientific)) at G-force 112 for 3 min. To avoid any insoluble solids,[51] and the resulting supernatant was diluted until 0.1 mg/ml, with 200 ng/ml of caffeine as internal standard. Also, a process blank, solvent blank, and pooled QCs were included for GC-MS analysis.

2.4.2.1 Untargeted metabolomics profile by GC-MS

Chemical composition of the methanolic extract of the species was performed using a gas chromatograph coupled with a mass spectrometer Shimadzu GCMS-QP2020NX fitted with a split/splitless injector and equipped with a capillary column Rtx-5 (30 m x 0.25 mm i.d. x 0.25 µm df). The oven temperature was programmed as follows: the initial column temperature was 70 °C, then increased 6 °C/min to 300 °C for 10 min for a total run time of 48.33 min. The injection port and

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

transfer line were kept at 200 °C and 220 °C, respectively. Ultra-pure helium was used as carrier gas at 1 mL/min initial flow. The analyses were performed operating in the electron impact (EI) ionization mode and full mass spectra were acquired. The MS ionization source was set to 200 °C, and the mass scan range was 50 to 500 Da. All the analyses were performed in triplicate, and an aliquot of each sample extract was pooled to create a pooled QC sample.

2.4.2.1.1 Metabolomics feature identification

Shimadzu files (*.qgd) data were converted to (*.mzML) using ProteoWizard (Supporting Information Table S13). Also, raw data was deposited in GNPS/MassIVE Repository MSV000091437. Raw spectral data were deconvoluted in MS-DIAL 4.9.221218 (http://prime.psc.riken.jp/) with a pipeline including peak detection, alignment, gap filling, and blank filtering (maximum sample intensity/average blank intensity ratio > 7) according to various parameters recorded in Supporting Information Table S13. The feature list table (*.txt) from MS-DIAL was exported for statistical analysis with the notame R package.

Additionally three approaches were adopted for metabolite identification: first, the (*.mgf) alignment result from MS-DIAL was exported to GNPS according to the GNPS documentation;[52] second (*.msp) alignment result from MS-DIAL was exported to NIST MS Search 2.4 software, where mainly, replib and mist_ri libraries of National institute Standard and Technology (NIST, Wiley Registry 12th Edition/NIST 20) were chosen, retention index tolerance used was 20 for Semi-Standard Non-Polar column; third, the MS-DIAL was used, the curated MSP spectral library from RIKEN Center for Sustainable Resource Science (http://prime.psc.riken.jp/; last edited on Aug. 21th, 2022) was load, the retention index tolerance was 20, mass tolerance used was 0.5 Da and EI similarity cut off was 70 %.

2.5 Data Analysis.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Based on the phytochemical analysis of the samples at each growth phase, we conducted a one-way analysis of the variance (ANOVA) followed by Tukey's test ($p > 0.05$) to identify statistical differences in the quantities of flavonoids and phenolic compounds in all the growth stages. Additionally, a principal component analysis (PCA) was conducted using the MS1 features table for each of the three final metabolomics datasets LC–MS (+), LC–MS (−) and GC-MS. All procedures, including Hierarchical Cluster Analysis (HCA) and heatmap for the identified metabolites, were performed using R v.3.6.2 (R development). The detailed scripts and data are available in the GitHub repository (https://github.com/IKIAM-NPL/U_sagittifolia_tubers).

3. Results and discussion

3.1 Total Phenolic And Flavonoid Content.

We found differences in the total phenolic and flavonoid content between *U. sagittifolia* growth stages. Specifically, juveniles exhibited a higher concentration compared to all other growth stages (Table 1). When examining the total flavonoid compounds, there was no noteworthy difference in their quantities between seedlings and adults. In contrast, the total phenolic content differed at each growth stage. In general, the results showed that the total phenolic content was lower in the initial stage, then significantly increased in juveniles, and subsequently decreased in adults. A similar trend was observed in the content of total flavonoids, initially, its content was low in seedlings, then increased substantially in the juvenile stage, but then dropped to seedling levels in the adult stage.

Table 1. Total phenolic and flavonoid content.

Growth stage	Phenolic (µg/mg)	Flavonoids (µg/mg)
Seedling	17.09 ± 0.08^{c}	6.83 ± 0.11^{b}
Juvenile	21.67 ± 0.03^{a}	8.23 ± 0.37^{a}

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Adult	17.65 ± 0.06^{b}	7.14 ± 0.11^{b}

Note: Values are mean ± SD (n = 3), different lowercase letter within the same column indicates significant differences based on Tukey's multiple comparisons ($p < 0.05$)

Our results are consistent with those reported by Li and Hu (2009).[17] These types of secondary metabolites have a wide spectrum of actions. For example, phenolics that are found in roots serve the purpose of chemical communication with the belowground microbiome. Although the plant-microbiome interaction involving phenolic exudates has been studied more in nodule-forming plants, most of these phenolics described in the literature (p-coumaric acid, protocatechuic acid, malic acid and others) are found in *U. sagittifolia*.[18,18–21] Conversely, as the plant matures, it becomes more likely to encounter above-the-ground stressors, such as UV radiation affecting the now-developed aerial tissues.[22] Thus, the requirement of phenolics in aerial tissue could explain the low levels of phenolics detected in adult roots.

3.2 LC-MS/MS-Based Metabolite Identification

A total of 43 metabolites were identified for the first time in *U. sagittifolia* tubers (Table 2). Between these 43 reported metabolites, 19 were identified at level 1 with the hybrid approach of comparing MS/MS fragmentation patterns between standard versus samples, while also manually checking the match of the protonated/deprotonated molecule with the known retention time of the analytical standard. Then, the remaining 24 metabolites were identified by using the available MS/MS libraries in GNPS (identification level 2). We encountered metabolite annotation challenges, such as having multiple promising candidates with a matching score difference of less than 1 unit of the MS2 score difference between annotation candidates. For example, the case of three quercetin monohexose detected at 2.8 min at 465.1020 m/z $[M+H]^{+}$ and 463.0869 m/z $[M-H]^{-}$. Despite having authentic standards for quercetin 3-glucofuranoside, quercetin 3-galactoside and quercetin 3-glucoside, the reported MS2 scores for these three matches varied by 0.2 units. Consequently, we were unable to confidently differentiate between the candidates due to the narrow margin in MS/MS matching scores.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

For the downstream analysis, features with multiple annotation candidates, the candidate with the highest score was kept for figure labelling purposes.

Furthermore, we also reported the observed molecule mass at each ionization polarity (protonated $[M+H]^+$ and deprotonated $[M-H]^-$) for annotated metabolite. In cases, where a specific metabolite was not detected at a given ionization polarity, we reported it as a dash. Moreover, we found that most annotated compounds were detected in negative polarity, while specific compounds were only detected in positive polarities such as amino acids (e.g., phenylalanine, betaine) and an alkaloid (trigonelline).

Table 2. Phenolic and other chemical compounds were identified in *U. sagittifolia* using UHPLC-MS/MS and GC-MS.

IL	RT/RI	$[M+H]^+$	$[M-H]^-$	Molecular Formula	Error(ppm)	MS2 Score*	Identification
1	0.312	-	179.0567	$C_6H_{12}O_6$	3.35	-	D-(+)-glucose
1	0.328	343.1237	341.1100	$C_{12}H_{22}O_{11}$	2.93	74.00	Sucrose
1	2.358	169.0493	167.0350	$C_8H_8O_4$	-0.35	94.25	Vanillic acid
1	2.692	-	163.0403	$C_9H_8O_3$	0.98	-	p-Coumaric acid
1	2.890	-	193.0507	$C_{10}H_{10}O_4$	-0.05	-	Ferulic acid

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

1	2.865	-	223.0614	$C_{11}H_{12}O_5$	0.57	-	Sinapic acid
1	2.890	-	163.0406	$C_9H_8O_3$	2.82	-	Trans-3-hydroxycinnamic acid/m-Coumaric acid
1	2.848	465,1021	463.0869	$C_{21}H_{20}O_{12}$	-2.97	86.40	Quercetin monohexosa
1	2.770	611.1612	609.1468	$C_{27}H_{30}O_{16}$	1.01	98.95	Rutin
1	2.981	625.1774	623.1651	$C_{28}H_{32}O_{16}$	1.92	96.01	Isorhamnetin-3-O-β-D-Rutinoside/Narcissin
1	0.360	-	133.016	$C_4H_6O_5$	12.63	99.72	Malic acid
1	1.827	-	153.0197	$C_7H_6O_4$	1.90	99.41	3,4 dihydroxybenzoic acid/Protocatechuic acid

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

1	2.407	-	153.0190	$C_7H_6O_4$	-2.67	90.13	2,3-dihydroxybenzoic acid/Pyrocatechuic acid
	2.407		153.0190		-2.67	90.20	2,6-dihydroxybenzoic acid
1	2.435	-	121.0296	$C_7H_6O_2$	0.17	100	4-hydroxybenzaldehyde
1	2.939	-	593.1512	$C_{27}H_{30}O_{15}$	0.00	93.25	Kaempferol-3-O-β-rutinoside/Nictoflorin
1	3.152	-	137.0245	$C_7H_6O_3$	0.04	100	Salicylic Acid
1	4.049	-	315.0509	$C_{16}H_{12}O_7$	-0.63	99.22	Isorhamnetin
2	1.571	284.0989	282.0850	$C_{10}H_{13}N_5O_5$	-2.09	79.75	Guanosine
2	2.965	479.1193	477.1040	$C_{22}H_{22}O_{12}$	-0.52	88.00	Isorhamnetin-3-O-glucoside

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

2	3.248	-	361.1670	$C_{20}H_{26}O_6$	1.90	79.00	2,3-bis[(4-hydroxy-3-methoxyphenyl)methyl]butane-1,4-diol/Secoisolariciresinol
2	2.743	-	165.0557	$C_9H_{10}O_3$	3.22	70.00	3-(3-Hydroxyphenyl)propanoic acid
2	2.893	.	503.2521	$C_{24}H_{40}O_{11}$	5.73	71.00	3,5,5-trimethyl-4-[3-[(6-O-beta-D-xylopyranosyl-beta-D-glucopyranosyl)oxy]butyl]/NCGC00380271-01
2	1.569	-	182.0459	$C_8H_9NO_4$	3.12	88.00	4-Pyridoxic acid
2	5.965	-	295.2278	$C_{18}H_{32}O_3$	1.63	77.00	9-hydroxy-10,12-octadecadienoic acid
2	2.423	-	593.152	$C_{27}H_{30}O_{15}$	2.28	78.00	Vicenin

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

1	2.347	-	177.0196	$C_9H_6O_4$	4.61	86.00	Esculetin
1	2.971	-	433.0783	$C_{20}H_{18}O_{11}$	2.80	73.00	Guajavarin
2	2.141	-	175.0615	$C_7H_{12}O_5$	4.860	72.00	Isopropylmalic acid
2	2.617	-	639.1564	$C_{28}H_{32}O_{17}$	0.43	71.00	Isorhamnetin-3-Glucoside-4'-Glucoside/Isorhamnetin 3,4'-diglucoside
2	2.577	-	563.1412	$C_{26}H_{28}O_{14}$	1.99	78.00	Isoschaftoside
2	2.483	-	172.0984	$C_8H_{15}NO_3$	5.99	82.00	N-Acetyl-L-leucine
2	1.175	-	292.1396	$C_{12}H_{23}NO_7$	-0.09	82.00	N-Fructosyl isoleucine/3-Methyl-2-[[2,3,4-trihydroxy-5-(hydroxymethyl)oxolan-2-yl]methylamino]pentanoic acid

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

2	0.701	-	290.0879	$C_{11}H_{17}NO_8$	1.06	73.00	N-Fructosyl pyroglutamate
2	2.148	-	137.0244	$C_7H_6O_3$	3.88	89.00	Protocatechuic aldehyde
2	0,861	-	180.0667	$C_9H_{11}NO_3$	3.51	75.00	Tyrosine
2	2.185	144.0809	-	$C_{10}H_9N$	-2.95	73.00	2-Naphthylamine/Naphthalen-2-amine
2	0.497	136.0616	-	$C_5H_5N_5$	-5.29	81.00	Adenine/7H-purin-6-amine
2	0.350	118.0859	-	$C_5H_{11}NO_2$	-7.65	73.00	Betaine
2	0.489	244.0929	-	$C_9H_{13}N_3O_5$	-1.83	77.00	Cytidine
2	1.679	166.0871	-	$C_9H_{11}NO_2$	-1.23	88.00	Phenylalanine/2-amino-3-phenylpropanoid acid

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

2	0.630	170.0812	-	$C_8H_{11}NO_3$	-3.05	79.00	Pyridoxine
2	2.748	303.0500	-	$C_{15}H_{10}O_7$	-1.58	71.00	Robinetin
2	0.335	138.0550	-	$C_7H_7NO_2$	-3.65	78.00	Trigonelline
2	1925	-	-	$C_{17}H_{34}O2$	-	88.49	Methyl palmitate•
2	2123	-	-	$C_{19}H_{38}O_2$	-	88.44	Methyl Stearate•
2	1958	-	-	$C_{16}H_{32}O_2$	-	71.45	Palmitic Acid•
2	1104	-	-	$C_{11}H_{24}$	-	98.47	Undecane•
2	3343	-	-	$C_{29}H_{50}O$	-	63.50	γ-Sitosterol•
2	-	-	-	$C_4H_8N_4O_4$	-	79.77	Allantoic acid•
2	3284	-	-	$C_{29}H_{48}O$	-	71.20	Stigmasterol•

Where IL identification level according to Summer et al. (2007) [23], RT retention time in minutes, RI experimental Kovats retention index (value > 1925) and • metabolite identified by GC-MS. *For metabolites identified at level 1, the MS2 score reported referred to the Agilent MassHunter MS/MS score, while for metabolites identified at level 2, the MS2 score refers to the GNPS-modified cosine score.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Additionally, it was possible to differentiate between the distinct growth stages by analyzing the reduced subspace that is defined by the principal component analysis (PCA). For the PCA analysis of both polarities, we used the complete feature table and replaced the identities of the features that we could annotate at levels 1 and 2 (Figure 1). The score plots for positive and negative polarity and the loading plots with the annotated features are presented in Figures 1A, 1B, 1C, and 1D, respectively.

In the score plot for positive and negative polarity (Figure 1A and 1B), we see a clear overlap of the pooled quality control samples, indicating good performance of the mass spectrometer as well as consistency in the data cleaning process. It is worth noting that both polarities display a similar projection in the score plot; thus, we will discuss both score plots simultaneously. Two clear clusters were captured by the first principal component (PC1), clustering juvenile and seedling tubers in a similar group (positive PC1), while adult tubers are clustered in the negative domain of the first principal component. Moreover, one seedling replicate was placed closer to the adult tuber, which suggests that the metabolomics profile of this replicate is more similar to the adult stage than the seedling stage. On the other hand, the PC2 separated two main groups consisting of seedlings and adults, and another composed of juveniles.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

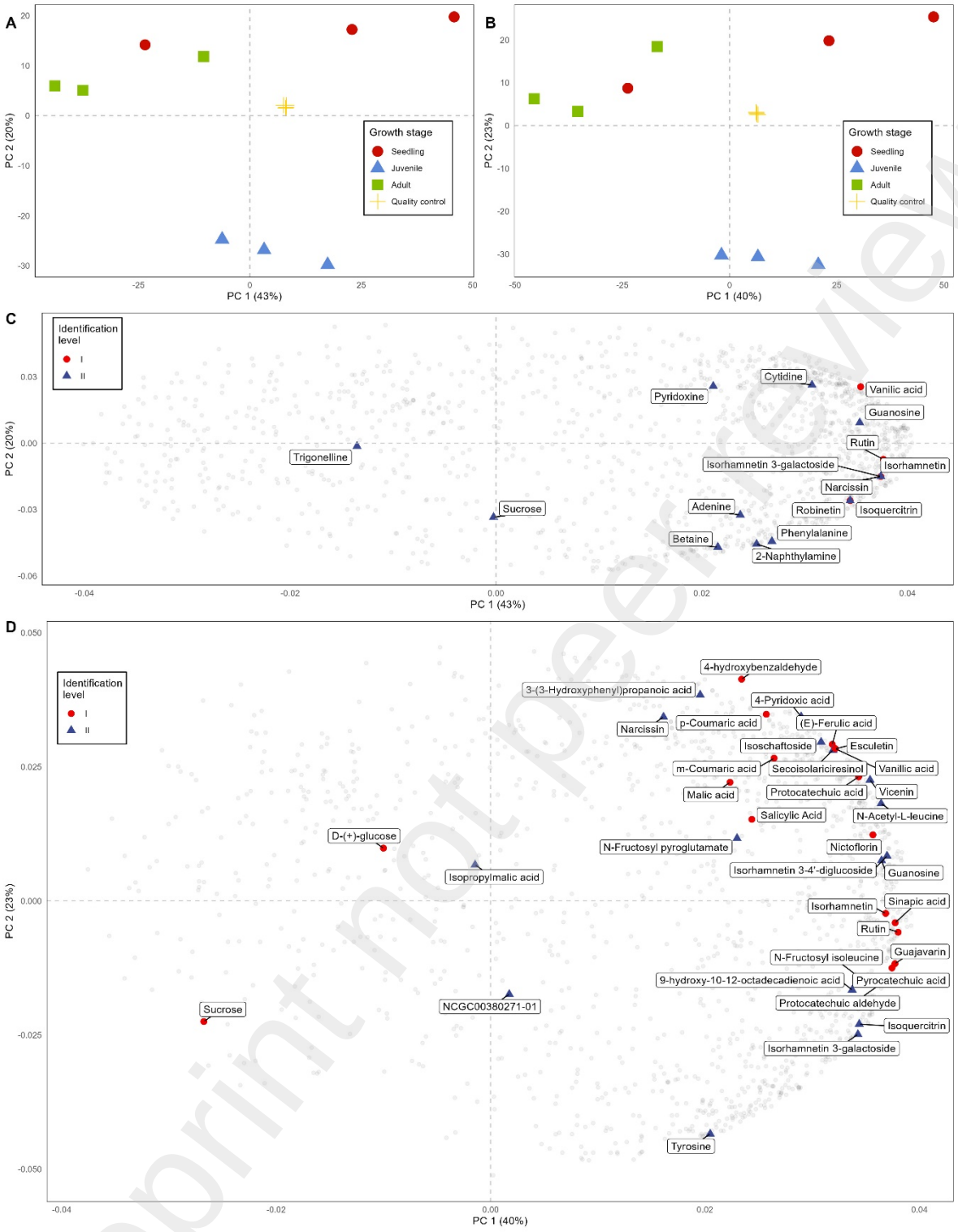

Figure 1. Principal component analysis (PCA) score and loadings plots the deconvoluted feature table at MS1 from different growth stages of *U. sagittifolia*. (A) Score plot in positive ionization mode, (B) Score plot in negative, (C) Loading plot in positive and (D) loading plot in negative.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

The loadings matrix obtained from the PCA analysis for both, positive and negative polarities, are shown in Figure 1C and 1D, respectively. We present the annotated features at level 1 as red circles, while features annotated at level 2 are displayed as blue triangles. It is worth noting that the grey circles in the loading plots represent the remaining features that we could not identify. Similar behavior can be distinguished in the location of the annotated features, as most features are located in the positive domain of the first principal component (PC1 > 0). After a closer look at the annotated features identified by LC-MS, we discovered that features located in the positive domain of the first PC have a greater relative abundance difference between groups, while features located near the origin of the PC1 tend to be more similar or present low variation across groups.

To interrogate the metabolite expression levels across the growth stages, we subset the annotated features and subjected them to heatmap visualization and hierarchical clustering analysis (HCA) (Figure 2). The HCA analysis was conducted sample-wise and metabolite-wise to reveal common clusters across samples and metabolites. While we decided to use a clustering cutoff of two groups sample-wise (shown as 1 and 2), three groups cutoff was used to create metabolite clusters (shown as a, b, and c); the cutoff threshold is displayed as a red dotted line above dendrograms in Figure 2. We also included the metabolite classification legend using a classifier at the superclass level.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

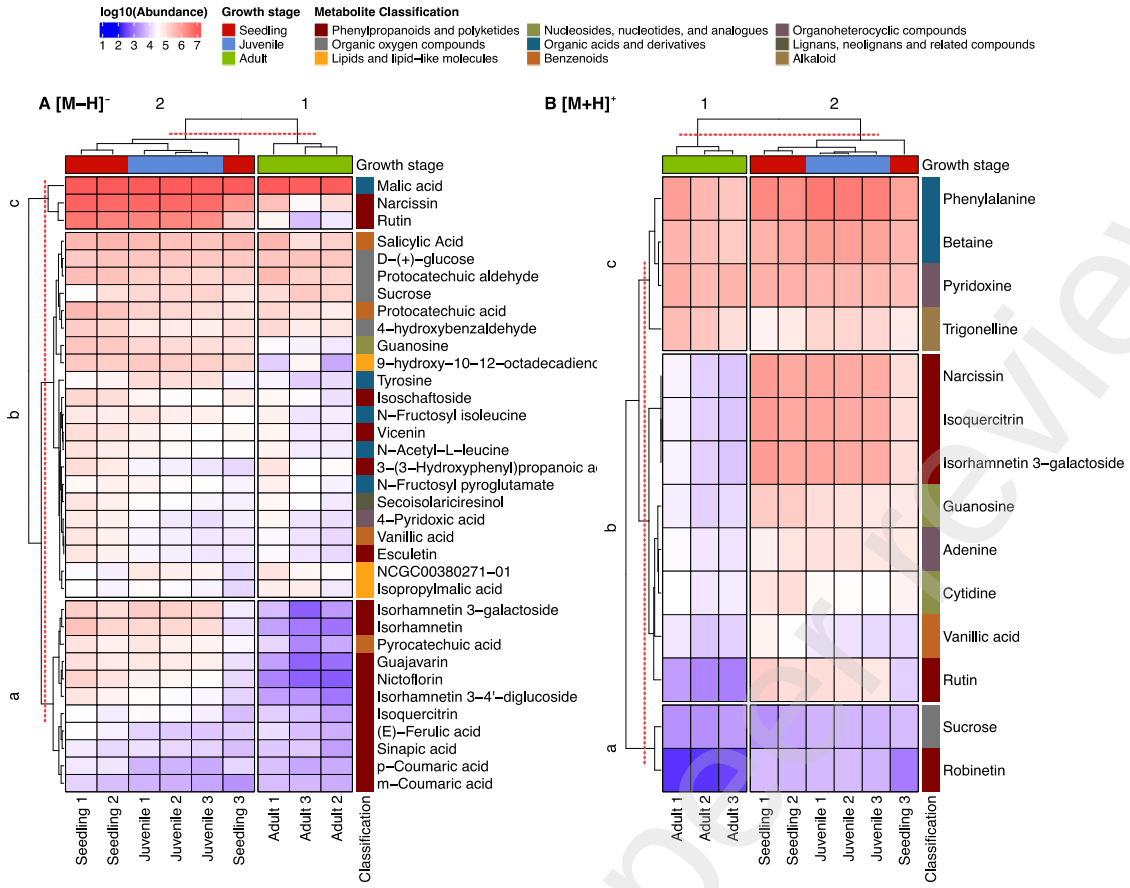

Figure 2. Heatmap and HCA analysis of the identified metabolites using LC-MS and LC-MS/MS in negative (A) and positive polarity (B). Legends were included to indicate the growth stage (seedlings, juveniles, and adults), and compound classification according to classifier. Cold colors (blue scale) represent a low metabolite abundance, while warmer colors (red scale) indicate a higher abundance.

In contrast with the suggested grouping found with PCA, the HCA of the annotated metabolites. HCA suggests a different group clustering where seedlings and juveniles are more similar, as they are consistently clustered together, while adult tuber is clustered in a separate node. However, the clustering grouping is not perfect as the seedling and juvenile tubers are not allocated in single clusters, suggesting that seedling 3 has a metabolite expression that more closely resembles the juvenile profile.

With the assistance of metabolite-wise HCA, we were able to identify common patterns in the metabolite distribution. In terms of relative abundance. In both polarities metabolites, we observe

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

three delineated clusters where cluster (c) is generally compromised of the most abundant metabolites, cluster (b) consists of a combination of abundant and less abundant metabolites, whereas cluster (a) encompasses the least abundant metabolites. To denote a particular grouping of metabolite clusters, we can utilize the metabolite and sample-wise clusters as coordinates.

In negative polarity (Figure 2A), it is depicted that metabolite cluster (c) accounted by malic acid, narcissin, and rutin are the most abundant metabolites at the early growth stage (seedling and juvenile), and decreases in the late growth stage (adult). Cluster (b) is the largest cluster found in this polarity with 21 metabolites from salicylic acid to isopropylamalic acid, while cluster (a) has the least abundant metabolite, exhibiting the lowest abundance in the late growth stage (Figure 2A clusters 1a).

In contrast with the negative polarity where cluster (c) presents different abundance by group, the positive polarity cluster (c), Figure 2B, suggests a constitutive metabolite expression throughout the tuber development. In the same vein as negative polarity, clusters (b) and (a) present in positive polarity show a differential metabolite expression. Cluster (b), starting with narcissin and ending with rutin, presents a mixed metabolite abundance distribution, where in the adult phase these metabolites are the least abundant. Finally, cluster (a) with sucrose and robinetin, is in general a cluster with the least abundant metabolites. It is worth noting that some metabolites present in both polarities, such as rutin detected the highest relative abundance in negative polarity, but moderate to low abundance in positive polarity. These different relative abundances in different polarities can be attributed to ionization efficiency, where in this case, rutin ionizes better in negative polarity.[24,25] However, metabolites with different relative abundances present a consensus behavior in expression profiles regardless of different ionization polarities.

Upon a closer examination of the metabolite profile of *U. sagitifolia* (Figure 2), it exhibits two clear patterns of metabolite expression in tubers. First, we noticed metabolites that are expressed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

consistently across all three growth stages with metabolites such as malic acid (Figure 2A cluster c), salicylic acid, glucose, sucrose, protocatechuic acid (Figure 2A cluster b), as well as phenylalanine, betaine, pyridoxine (Figure 2B cluster c). The second cluster shown in Figure 2A (cluster b), refers to metabolites that are expressed in different abundances at different growth stages, where the majority of them are expressed in the seedling and juvenile growth phase.

If we consider the phenylpropanoid pathway, the early phase of polyphenol biosynthesis, phenylalanine is the precursor for phenylalanine ammonia-lyase (PAL) to synthesize cinnamate, which subsequently is synthesized to p-coumaric acid by cinnamate 4-hydroxylase. This metabolite can be directed to produce p-coumaric acid derivates which can later contribute to lignin creation for root cell walls, or be allocated to polyphenol biosynthesis.[22,26] According to our results, phenylalanine is a housekeeping metabolite in relatively high abundance which could be used to produce p-coumaric acid. The observed low relative intensity of p-coumaric acid and the high relative abundances of a series of polyphenols shown in Figure 2A, suggest an effective allocation of p-coumaric acid as a precursor for polyphenol biosynthesis. However, we cannot discard the possibility of phenylalanine allocation to p-coumaric acid production as it might also be allocated to other pathways such as protein and lignin synthesis.

3.3 Molecular Networking By LC-MS/MS

To complement the multivariate analysis to better understand the metabolite similarly based on its fragmentation patterns, we use molecular networking using GNPS, which provides information on structural similarities between phytochemicals.

The nodes in the network correspond to metabolomics features. In cases where we could identify the feature identity, we include the node name with the identified metabolite. Metabolites are linked to other nodes based on their similarity in terms of chemical structure, fragmentation, and retention times.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

The feature-based molecular networking algorithm available in GNPS was used for molecular network analysis. Results showed 539 nodes in negative and 393 nodes in positive ionization polarity (Supporting Information Figure S1 and S2). In negative ionization polarity, 55 molecular clusters with at least 2 connected nodes, and 101 unconnected nodes were detected, while in positive mode, 24 molecular clusters with at least two connected nodes and 56 unconnected nodes were generated.

The molecular network was propagated taking two approaches. First, we evaluated the connection between nodes based on the result of the merge polarities pipeline, relating features in different ionization polarities considering precursor ion mass correction and retention time match. The second approach was manually curating the relationship between the unannotated nodes with the nearest annotated node. The result of the propagated molecular network is presented in Figure 3. The black edges connecting the nodes were calculated based on cosine similarities, while purple edges reflect nodes are connected by polarity match. Furthermore, nodes with a circular border represent features detected in negative polarity, while nodes with square borders are features detected in positive polarity (Figure 4).

In terms of polarity match, there were metabolites that we were able to annotate in both ionization polarities based on retention time and MS/MS match. One of these cases is of nictoflorin, where we see a clear mirror plot match where the experimental MS/MS spectrometric matches in retention time as well as the spectrometric product ion against the analytical standard spectrometric (Figure 3). Furthermore, annotating the same metabolite in both polarities was not always the case. There were instances where metabolites were annotated in either negative or positive polarity. For example, rutin was only annotated in negative polarity with a reference spectrometric from the MONA database (GNPS spectrometric ID = CCMSLIB00004719964). The molecular network connects the GNPS annotated rutin node with a feature detected in positive mode (node id = 2879). The retention time difference between these two features is 0.02 minutes and has a mass difference of 2.012 Da, attributed to the correction of the proton mass based on distinct polarities. This node

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

had a match in MS2LDA to a motif that represents a quercetin or a glycosylated quercetin substructure (https://ms2lda.org/motifdb/motif/151174/), suggesting a glycosylation with a quercetin polyphenol backbone. When we investigated the node id = 2879 further, this node was annotated as rutin in positive polarity against the in-house library at the MS1 level, and we proceeded to propagate the molecular annotation. We took the same reasoning to propagate the annotation of isoschaftoside from negative polarity to positive polarity. This polyphenol was annotated in negative polarity based on the MS/MS reference spectrometric from MassBank present in GNPS (GNPS spectrometric ID = CCMSLIB00005778299), and connected with a node with a mass difference of 2.013 Da and a retention time shift of 0.03 min. We present the mirror plot of the experimental spectrometric in negative polarity at 40 eV against the literature spectrometric from MassBank at 50 eV (Figure 3). In this mirror plot, we can see a clear match between the experimental and the literature spectrometric, and the inner spectral differences in ion intensities could be attributed to the different collision energies applied.

On the other hand, when we analyzed the molecular network by cosine similarity, we noted that narcissin is the node with the most connections (6 connections), and the mirror plot shows a good match between the experimental and analytical standard spectrometric (Figure 4). Narcissin was connected with a node that, according to the GNPS annotations is CCG-208435 (GNPS spectrometric ID = CCMSLIB00010113678). The mass difference between these nodes is 35.97 Da. The mass difference between CCG-208435 and narcissin can be attributed to the adduct type since this node was identified as a $[M+Cl]^-$ adduct based on the isotope pattern by GNPS. After we accounted for the mass of chlorine, the monoisotopic mass of CCG-208435 is the same monoisotopic mass as narcissin. However, not only the monoisotopic masses are the same, but both metabolites were detected at the same retention time (Δ rt = 0.008 min), suggesting that CCG-208435 might be the [M+Cl]- adduct of narcissin. To confirm or discard the new annotation of this feature, we compared the authentic standard of narcissin against the MS/MS spectrometric of what GNPS annotated as CCG-208435,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

and we observed a cosine score of 0.97. Therefore, we decided to report this feature as the $[M+Cl]^-$ adduct of narcissin, instead of CCG-208435. In the same vein, we observed a $[2M-H]^-$ adduct, which suggests to be a narcissin dimer, which is in line with retention time and MS/MS fragmentation.

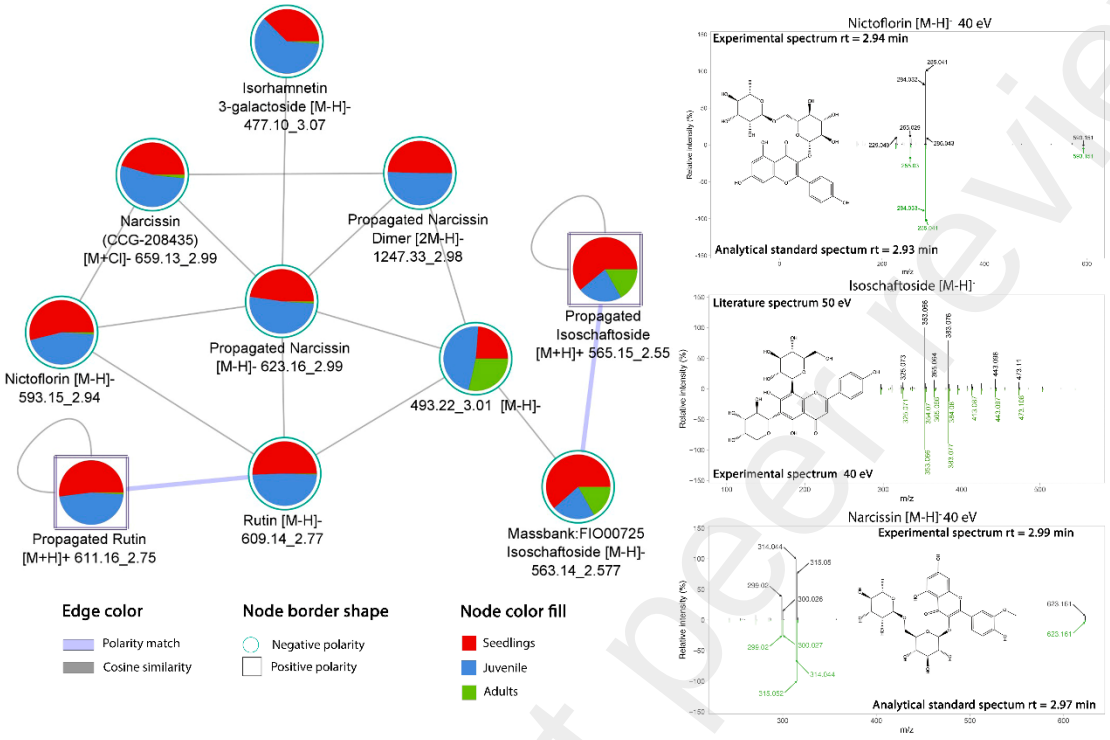

Figure 3. Molecular network results of untargeted metabolomics of *U. sagittifolia* tubers. Feature-based molecular networks differentiated by cosine (black edges) or polarity (purple edges) match (left). Node fill areas represent the relative abundance of each feature at each growth stage. The node border represents at which polarity each node was detected. In the right side, mirror plots for nictoflorin, isoschaftosie and narcissin.

In this sense, molecular networking (Figure 3) facilitated the identification of some flavonoids (of 11 flavonoids identified), contributing to the differentiation among various stages of the plant´s growth (Figure 1). The rutin abundance behavior during the growth stages (Figure 2 and 4) was consistent with a study on *P. coccinea* by Fico et al. (2000), which found that rutin was always present during the pre-flowering growth phase as one of its main constituents.[27] Moreover, it was found that certain compounds such as gujavarin, sinapic acid, and tyrosine were hardly detected in adult plants

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

compared to seedlings and juveniles. Conversely, compounds such as isopropylmalic acid, NCGC00380271-01, D(+)-glucose, and sucrose were found to be higher in the reproductive phase. This might be explained due to the difference in nutrient requirements during the plant's life cycle. Likewise, the analysis uncovered metabolites such as guanosine, cytadine, and vanillic acid at every growth phase, with greater concentrations in the initial development phase. However, metabolites such as quercetin-3-glucoside and isorhamnetin were more prevalent in the juvenile phase. As the plant reached the reproductive phase, the concentration of these elements declined, with trigonelline and sucrose being the most abundant metabolites. The production and accumulation of metabolites are linked to variations in biosynthetic metabolic pathways of medicinal constituents, which could explain the abundance fluctuations observed during different growth stages. An example of this metabolite variation was reported in Calluna vulgaris roots, which had the highest phenolic content during the vegetative pre-reproductive stage as well as in our study (Table 1) while the leaves had the lowest concentrations.[28] However, as the plant matured, a redistribution of phenolic compounds occurred during the floral budding stage. Similarly, perennial medicinal plants usually experience an increase during their growth phase and subsequently decrease once they reach a certain point. For instance, Zhannan et al. (2012) found that chlorogenic acid, hyperin, and quercetin were most abundant in 13-year-old plants, while rutin and quercitrin were highest in 7-year-old plants, and magnolol was highest in 10-year-old plants of Magnolia officinalis bark.[29] This variation could be attributed to the rapid nutrient consumption by developing reproductive organs, which causes the swift reduction of compound diversity until the plant reaches reproductive age, a phenomenon that can be observed in the root metabolite content since roots are closely associated with plant reproductive growth.[29]

3.4 GC-MS-Based Metabolite Identification

Volatile and semi-volatile compounds identified by GC-MS in methanolic extracts are shown in Table 2. Seven compounds were tentatively identified, four were identified by the public MS-DIAL

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

library (http://prime.psc.riken.jp/), two by the NIST library and one by the GNPS platform, most of them through spectrometric similarities and retention index.

Most of the identified compounds were fatty acids and phytosterols and their distribution was relatively similar between juvenile and adult plants. In contrast of seedlings, which is reflected in the score's plots shown in Figure 4.

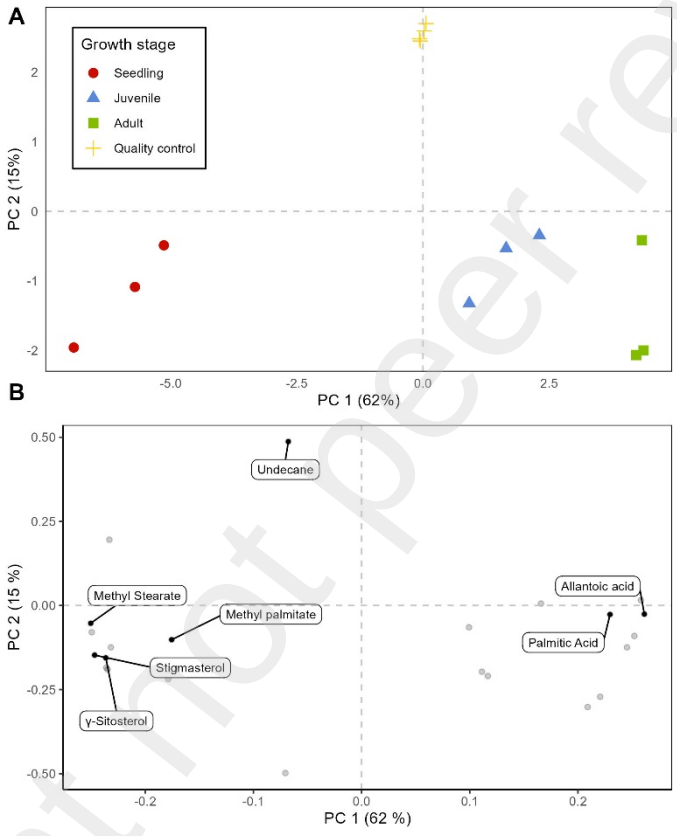

Figure 4. Principal component analysis (PCA) score (A) and loadings (B) plots of GC-MS from different growth stages of *U. sagittifolia*.

Principal component analysis of the volatile compounds revealed that PC1 and PC2 account for 62% and 15% of the explained variance, respectively. Through analyzing the loadings plots (Figure 5B), it was determined there were three groups in the PC1 space (Figure 4). Seedlings showed higher levels of fatty acids, including methyl palmitate, methyl stearate, and undecane, which were only detected during this period. while the remaining components were found in all growth phases.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Additionally, stigmasterol and sitosterol were found in higher concentrations in seedlings, while allantoic acid and palmitic were present in higher concentrations in juveniles and adults.

This discovery aligns with the fact that lipids are essential for promoting growth in the early stages of the plant's life cycle.[17] Similarly, Pariyani et al. (2020) reported higher concentrations of fatty acids in the young roots of four different plant species.[30] As the plant matures, allantoic acid and palmitic acid become to be produced at higher levels. This shift can be explained since low molecular weight organic acids play a crucial role, as they can be released into the soil, mobilizing nutrients, and increasing the plant's ability to survive and grow normally. This process is particularly significant in environments with low nutrient availability.[31] Therefore, understanding the accumulation pattern of metabolite in plant growth and its development can help determine the most appropriate harvest periods for medicinal plants.[32]

Conclusions

Motivated by the promising ability of by *U. sagittifolia* root extracts to neutralize the *Bothrops atrox* venom-induced toxicity, we investigated the metabolomics profile of the tubers at three different growth stages.[16] In this study, we used untargeted metabolomics assisted by gas and liquid chromatography coupled with mass spectrometry to unveil the phytochemical diversity of *U. sagittifolia* roots, used as traditional medicine by local communities to treat snakebite cases. Furthermore, we used chemometrics and molecular networking-based algorithms to understand better the similarities and differences across the chemical profile based on the growth stages and propagate the metabolite molecular network annotation.

Overall, we found 50 compounds, such as polyphenols (no flavonoids), flavonoids, lipids, alkaloids, sugars, and others. Most of the flavonoids are found at the highest levels in seedlings, while sugars are at the highest abundance in adults, and polyphenols (no flavonoids) are in juveniles. This

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

metabolomic characterization will allow subsequent studies about obtaining pure molecules or extracts that could present further biological activities.

Although our findings suggest a diversity in phenolic content, the phytochemical screening conducted by Vera-Palacios et al. (2022) also reported high levels of alkaloids. Therefore, future studies to explore this phytochemical class, with specific sample preparation and metabolomics methods are required. Between our limitations, we can also note that we report metabolite levels in terms of relative abundance based on peak height, and not in absolute quantities, which might not be comparable across different chemical classes as they present different ionization efficiencies that can intricate the comparison between different phytochemical classes.

CRediT authorship contribution statement

Jefferson V Pastuña: Methodology, Writing, data curation – original draft. **Cristian D. Quiroz-Moreno**: Methodology, review & editing. **Evencio J. Medina**: Methodology validation. **Jessica L. Cooperstone**: validation, writing – review & editing. **Mateo Radice**: review & editing, **Maria C. Peñuela-Mora:** review & editing. Jose R. Almeida: writing review & editing. **Noroska Gabriela Salazar Mogollón:** Supervision, Project administration, review & editing.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Suplementary materials

Supplemental Table S1-13, MZmine and MS-DIAL parameters used in gas and liquid mass spectrometry data process (XLSX)

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Figure S1, LC-MS/MS molecular networking in positive ionization mode; Figure S2, LC-MS/MS molecular networking in positive ionization mode (PDF)

Acknowledgment

We thank Pablo Leonardo Corella, a student at the Universidad Regional Amazónica Ikiam for technical assistance

Funding Sources

This research was funded by Agencia Española de Cooperación Internacional para el Desarrollo (AECID), grant number 2018/SPE/000040019 (project I+P+i: Investigación de Excelencia, Posgrado e Igualdad.

References

[1] A.G. Atanasov, S.B. Zotchev, V.M. Dirsch, the International Natural Product Sciences Taskforce, C.T. Supuran, Natural products in drug discovery: advances and opportunities, Nature Reviews Drug Discovery. 20 (2021) 200–216. https://doi.org/10.1038/s41573-020-00114-z.

[2] A.L. Harvey, R. Edrada-Ebel, R.J. Quinn, The re-emergence of natural products for drug discovery in the genomics era, Nature Reviews Drug Discovery. 14 (2015) 111–129. https://doi.org/10.1038/nrd4510.

[3] M. Boozari, H. Hosseinzadeh, Natural products for COVID-19 prevention and treatment regarding to previous coronavirus infections and novel studies, Phytotherapy Research. 35 (2021) 864–876. https://doi.org/10.1002/PTR.6873.

[4] S.E. Rossiter, M.H. Fletcher, W.M. Wuest, Natural Products as Platforms to Overcome Antibiotic Resistance, Chemical Reviews. 117 (2017) 12415–12474. https://doi.org/10.1021/acs.chemrev.7b00283.

[5] A. Skirycz, S. Kierszniowska, M. Méret, L. Willmitzer, G. Tzotzos, Medicinal Bioprospecting of the Amazon Rainforest: A Modern Eldorado?, Trends in Biotechnology. 34 (2016) 781–790. https://doi.org/10.1016/j.tibtech.2016.03.006.

[6] S. Ahmed, T.S. Griffin, D. Kraner, M.K. Schaffner, D. Sharma, M. Hazel, A.R. Leitch, C.M. Orians, W. Han, J.R. Stepp, A. Robbat, C. Matyas, C. Long, D. Xue, R.F. Houser, S.B. Cash, Environmental Factors Variably Impact Tea Secondary Metabolites in the Context of Climate Change, Frontiers in Plant Science. 10 (2019) 1–16. https://doi.org/10.3389/FPLS.2019.00939/BIBTEX.

[7] M. Marrelli, Medicinal Plants, Plants 2021. 10 (2021) 1355. https://doi.org/10.3390/plants10071355.

[8] F. Campos-Arguedas, G. Sarrailhé, P. Nicolle, M. Dorais, N.J.B. Brereton, F.E. Pitre, K. Pedneault, Different Temperature and UV Patterns Modulate Berry Maturation and Volatile Compounds Accumulation in Vitis sp., Frontiers in Plant Science. 13 (2022) 1–17. https://doi.org/10.3389/FPLS.2022.862259/BIBTEX.

[9] S. Dragović, V. Dragović-Uzelac, S. Pedisić, Z. Čošić, M. Friščić, I.E. Garofulić, Z. Zorić, The mastic tree (Pistacia lentiscus l.) leaves as source of BACs: Effect of growing location, phenological stage and

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

extraction solvent on phenolic content, Food Technology and Biotechnology. 58 (2020) 303–313. https://doi.org/10.17113/FTB.58.03.20.6662.
[10] Z.D. Jakovljević, S.M. Stanković, D.M. Topuzović, Seasonal variability of Chelidonium majus L. secondary metabolites content and antioxidant activity, EXCLI Journal. 12 (2013) 260–268. https://doi.org/10.17877/DE290R-5473.
[11] S. Gonda, Special Issue: Plant Metabolomics, Metabolites. 10 (2020) 1–4. https://doi.org/10.3390/metabo10110467.
[12] A. Ribbenstedt, H. Ziarrusta, J.P. Benskin, Development, characterization and comparisons of targeted and non-targeted metabolomics methods, PLOS ONE. 13 (2018) e0207082. https://doi.org/10.1371/journal.pone.0207082.
[13] A.A.X. Adrião, A.O. dos Santos, E.J.S.P. de Lima, J.B. Maciel, W.H.P. Paz, F.M.A. da Silva, M.B. Pucca, A.M. Moura-da-Silva, W.M. Monteiro, M.A. Sartim, H.H.F. Koolen, Plant-Derived Toxin Inhibitors as Potential Candidates to Complement Antivenom Treatment in Snakebite Envenomations, Frontiers in Immunology. 13 (2022). https://doi.org/10.3389/FIMMU.2022.842576.
[14] I. Gómez-Betancur, V. Gogineni, A. Salazar-Ospina, F. León, Perspective on the Therapeutics of Anti-Snake Venom, Molecules. 24 (2019). https://doi.org/10.3390/MOLECULES24183276.
[15] T.B. Croat, O.O. Ortiz, M. Cedeño-Fonseca, D. Scherberich, A new Urospatha Schott (Araceae: Lasioideae) from the Caribbean Slope of Panama and Colombia, Feddes Repertorium. 131 (2020) 196–202. https://doi.org/10.1002/fedr.202000013.
[16] A.L. Vera-Palacios, J.D. Sacoto-Torres, J.A. Hernández-Altamirano, A. Moreno, M.C. Peñuela-Mora, D. Salazar-Valenzuela, N.G.S. Mogollón, J.R. Almeida, A First Look at the Inhibitory Potential of Urospatha sagittifolia (Araceae) Ethanolic Extract for Bothrops atrox Snakebite Envenomation, Toxins. 14 (2022) 1–18. https://doi.org/10.3390/toxins14070496.
[17] J. Li, Z. Hu, Accumulation and Dynamic Trends of Triterpenoid Saponin in Vegetative Organs of Achyranthus bidentata, Journal of Integrative Plant Biology. 51 (2009) 122–129. https://doi.org/10.1111/J.1744-7909.2008.00764.X.
[18] T. Rudrappa, K.J. Czymmek, P.W. Paré, H.P. Bais, Root-Secreted Malic Acid Recruits Beneficial Soil Bacteria, Plant Physiology. 148 (2008) 1547–1556. https://doi.org/10.1104/pp.108.127613.
[19] S.M. Mandal, D. Chakraborty, S. Dey, Phenolic acids act as signaling molecules in plant-microbe symbioses, Plant Signaling & Behavior. 5 (2010) 359–368. https://doi.org/10.4161/psb.5.4.10871.
[20] G. Seneviratne, H.S. Jayasinghearachchi, Phenolic acids: Possible agents of modifying N2-fixing symbiosis through rhizobial alteration?, (n.d.).
[21] S. Mandal, M. Mandal, A. Das, B. Pati, A. Ghosh, Stimulation of indoleacetic acid production in a Rhizobium isolate of Vigna mungo by root nodule phenolic acids, Arch Microbiol. 191 (2009) 389–393. https://doi.org/10.1007/s00203-008-0455-6.
[22] R. Marchiosi, W.D. Dos Santos, R.P. Constantin, R.B. De Lima, A.R. Soares, A. Finger-Teixeira, T.R. Mota, D.M. De Oliveira, M.D.P. Foletto-Felipe, J. Abrahão, O. Ferrarese-Filho, Biosynthesis and metabolic actions of simple phenolic acids in plants, Phytochem Rev. 19 (2020) 865–906. https://doi.org/10.1007/s11101-020-09689-2.
[23] L.W. Sumner, A.E. Alexander, A. Ae, D.B. Ae, M.H.B. Ae, R. Beger, C.A. Daykin, A.E. Teresa, W.-M. Fan, A.E. Oliver, F. Ae, R. Goodacre, A.E. Julian, L. Griffin, A.E. Thomas, H. Ae, N. Hardy, A.E. James, H. Ae, R. Higashi, A.E. Joachim, K. Ae, A.N.L. Ae, J.C. Lindon, A.E. Philip, M. Ae, A.W.N. Ae, M.D. Reily, J.J.T. Ae, M.R. Viant, L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, Á.R. Higashi, R. Higashi, O. Fiehn, R. Goodacre, Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI), Metabolomics. 3 (2007) 211–221. https://doi.org/10.1007/s11306-007-0082-2.
[24] M.-J. Motilva, A. Serra, A. Macià, Analysis of food polyphenols by ultra high-performance liquid chromatography coupled to mass spectrometry: An overview, Journal of Chromatography A. 1292 (2013) 66–82. https://doi.org/10.1016/j.chroma.2013.01.012.
[25] S. Piccolella, G. Crescente, L. Candela, S. Pacifico, Nutraceutical polyphenols: New analytical challenges and opportunities, Journal of Pharmaceutical and Biomedical Analysis. 175 (2019) 112774. https://doi.org/10.1016/j.jpba.2019.07.022.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

[26] W. Liu, Y. Feng, S. Yu, Z. Fan, X. Li, J. Li, H. Yin, The Flavonoid Biosynthesis Network in Plants, IJMS. 22 (2021) 12824. https://doi.org/10.3390/ijms222312824.
[27] G. Fico, A. R. Bilia, I. Morelli, F. Tomè, Flavonoid distribution in Pyracantha coccinea plants at different growth phases, Biochemical Systematics and Ecology. 28 (2000) 673–678. https://doi.org/10.1016/S0305-1978(99)00109-X.
[28] A.A. Cucu, G.M. Baci, A.B. Cucu, Ş. Dezsi, C. Lujerdean, I.C. Hegeduş, O. Bobiş, A.R. Moise, D.S. Dezmirean, Calluna vulgaris as a Valuable Source of Bioactive Compounds: Exploring Its Phytochemical Profile, Biological Activities and Apitherapeutic Potential, Plants. 11 (2022) 1–31. https://doi.org/10.3390/PLANTS11151993.
[29] Y. Zhannan, L. Shiqiong, Y. Zhengwen, Comparison analysis of bioactive compounds of Mignolia officinalis Rehd.et Wils from different growth ages, Heilongjiang Medicine Journal. 25 (2012) 553–555.
[30] R. Pariyani, M. Kortesniemi, J. Liimatainen, J. Sinkkonen, B. Yang, Untargeted metabolic fingerprinting reveals impact of growth stage and location on composition of sea buckthorn (Hippophaë rhamnoides) leaves, Journal of Food Science. 85 (2020) 364–373. https://doi.org/10.1111/1750-3841.15025.
[31] J.A. Lucas García, C. Barbas, A. Probanza, M.L. Barrientos, F.J. Gutierrez Mañero, Low molecular weight organic acids and fatty acids in root exudates of two Lupinus cultivars at flowering and fruiting stages, Phytochemical Analysis. 12 (2001) 305–311. https://doi.org/10.1002/PCA.596.
[32] H. Jianping, L. Zongsuo, W. Jingmin, The Relationship Between Mineral Element and the Root Growth and Accumulation of Effective Ingredient in Root of Traditional Herbs, PLANT PHYSIOLOGY COMMUNICATIONS. 39 (2003) 78–82.
[33] D. Barabé, C. Lacroix, M. Gibernau, Floral development of Urospatha: merosity and phylogeny in the Lasioideae (Araceae), Plant Syst Evol. 296 (2011) 41–50. https://doi.org/10.1007/s00606-011-0475-6.
[34] P.G. Vásquez-Ocmín, A. Gadea, S. Cojean, G. Marti, S. Pomel, A.C. Van Baelen, L. Ruiz-Vásquez, W. Ruiz Mesia, B. Figadère, L. Ruiz Mesia, A. Maciuk, Metabolomic approach of the antiprotozoal activity of medicinal Piper species used in Peruvian Amazon, Journal of Ethnopharmacology. 264 (2021) 1–16. https://doi.org/10.1016/J.JEP.2020.113262.
[35] N.L. Lartey, H. Asare-Anane, E.K. Ofori, S. Antwi, J. Asiedu-Larbi, F. Ayertey, L.K.N. Okine, Antidiabetic activity of aqueous stem bark extract of Annickia polycarpa in alloxan-induced diabetic mice, Journal of Traditional and Complementary Medicine. 11 (2021) 109–116. https://doi.org/10.1016/J.JTCME.2020.02.001.
[36] W. Wang, J. Li, H. Zhang, X. Wang, J. Fan, X. Zhang, Phenolic compounds and bioactivity evaluation of aqueous and methanol extracts of Allium mongolicum Regel, Food Science & Nutrition. 7 (2019) 779–787. https://doi.org/10.1002/FSN3.926.
[37] E.A. Bilbrey, K. Williamson, E. Hatzakis, D.D. Miller, J. Fresnedo-Ramírez, J.L. Cooperstone, Integrating genomics and multi-platform metabolomics enables metabolite QTL detection in breeding-relevant apple germplasm, Plant Biology, 2021. https://doi.org/10.1101/2021.02.18.431481.
[38] D. Broadhurst, R. Goodacre, S.N. Reinke, J. Kuligowski, I.D. Wilson, M.R. Lewis, W.B. Dunn, Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies, Metabolomics. 14 (2018) 72. https://doi.org/10.1007/s11306-018-1367-3.
[39] A.-T. Ramabulana, D. Petras, N.E. Madala, F. Tugizimana, Mass spectrometry DDA parameters and global coverage of the metabolome: Spectral molecular networks of momordica cardiospermoides plants, Metabolomics. 19 (2023) 18. https://doi.org/10.1007/s11306-023-01981-4.
[40] M.C. Chambers, B. MacLean, R. Burke, D. Amodei, D.L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, K. Hoff, D. Kessner, N. Tasman, N. Shulman, B. Frewen, T.A. Baker, M.Y. Brusniak, C. Paulse, D. Creasy, L. Flashner, K. Kani, C. Moulding, S.L. Seymour, L.M. Nuwaysir, B. Lefebvre, F. Kuhlmann, J. Roark, P. Rainer, S. Detlev, T. Hemenway, A. Huhmer, J. Langridge, B. Connolly, T. Chadick, K. Holly, J. Eckels, E.W. Deutsch, R.L. Moritz, J.E. Katz, D.B. Agus, M. MacCoss, D.L. Tabb, P. Mallick, A cross-platform toolkit for mass spectrometry and proteomics, Nature Biotechnology. 30 (2012) 918–920. https://doi.org/10.1038/nbt.2377.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

[41] K. Haug, K. Cochrane, V.C. Nainala, M. Williams, J. Chang, K.V. Jayaseelan, C. O'Donovan, MetaboLights: a resource evolving in response to the needs of its scientific community, Nucleic Acids Research. 48 (2020) D440–D444. https://doi.org/10.1093/NAR/GKZ1019.
[42] R. Schmid, S. Heuckeroth, A. Korf, A. Smirnov, O. Myers, T.S. Dyrlund, R. Bushuiev, K.J. Murray, N. Hoffmann, M. Lu, A. Sarvepalli, Z. Zhang, M. Fleischauer, K. Dührkop, M. Wesner, S.J. Hoogstra, E. Rudt, O. Mokshyna, C. Brungs, K. Ponomarov, L. Mutabdžija, T. Damiani, C.J. Pudney, M. Earll, P.O. Helmer, T.R. Fallon, T. Schulze, A. Rivas-Ubach, A. Bilbao, H. Richter, L.F. Nothias, M. Wang, M. Orešič, J.K. Weng, S. Böcker, A. Jeibmann, H. Hayen, U. Karst, P.C. Dorrestein, D. Petras, X. Du, T. Pluskal, Integrative analysis of multimodal mass spectrometry data in MZmine 3, Nature Biotechnology. 41 (2023) 447–449. https://doi.org/10.1038/s41587-023-01690-2.
[43] O.D. Myers, S.J. Sumner, S. Li, S. Barnes, X. Du, One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks, Analytical Chemistry. 89 (2017) 8696–8703. https://doi.org/10.1021/acs.analchem.7b00947.
[44] A. Klåvus, antonvsdata/notame: An R package for non-targeted LC-MS metabolomics, Github. (2023). https://github.com/antonvsdata/notame (accessed June 8, 2023).
[45] A. Klåvus, M. Kokla, S. Noerman, V.M. Koistinen, M. Tuomainen, I. Zarei, T. Meuronen, M.R. Häkkinen, S. Rummukainen, A.F. Babu, T. Sallinen, O. Kärkkäinen, J. Paananen, D. Broadhurst, C. Brunius, K. Hanhineva, "Notame": Workflow for Non-Targeted LC–MS Metabolic Profiling, Metabolites. 10 (2020) 1–35. https://doi.org/10.3390/METABO10040135.
[46] Quiroz-Moreno, Cristian (Daniel), J.L. Cooperstone, PhenolicsDB: phenolics MS/MS (LC-ESI-QTOF) database repository, (n.d.).
[47] M. Wang, J.J. Carver, V.V. Phelan, L.M. Sanchez, N. Garg, Y. Peng, D.D. Nguyen, J. Watrous, C.A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A.V. Melnik, M.J. Meehan, W.T. Liu, M. Crüsemann, P.D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R.D. Kersten, L.A. Pace, R.A. Quinn, K.R. Duncan, C.C. Hsu, D.J. Floros, R.G. Gavilan, K. Kleigrewe, T. Northen, R.J. Dutton, D. Parrot, E.E. Carlson, B. Aigle, C.F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B.T. Murphy, L. Gerwick, C.C. Liaw, Y.L. Yang, H.U. Humpf, M. Maansson, R.A. Keyzers, A.C. Sims, A.R. Johnson, A.M. Sidebottom, B.E. Sedio, A. Klitgaard, C.B. Larson, C.A.P. Boya, D. Torres-Mendoza, D.J. Gonzalez, D.B. Silva, L.M. Marques, D.P. Demarque, E. Pociute, E.C. O'Neill, E. Briand, E.J.N. Helfrich, E.A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J.J. Kharbush, Y. Zeng, J.A. Vorholt, K.L. Kurita, P. Charusanti, K.L. McPhail, K.F. Nielsen, L. Vuong, M. Elfeki, M.F. Traxler, N. Engene, N. Koyama, O.B. Vining, R. Baric, R.R. Silva, S.J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P.G. Williams, J. Dai, R. Neupane, J. Gurr, A.M.C. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B.M. Duggan, J. Almaliti, P.M. Allard, P. Phapale, L.F. Nothias, T. Alexandrov, M. Litaudon, J.L. Wolfender, J.E. Kyle, T.O. Metz, T. Peryea, D.T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K.M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P.R. Jensen, B. Palsson, K. Pogliano, R.G. Linington, M. Gutiérrez, N.P. Lopes, W.H. Gerwick, B.S. Moore, P.C. Dorrestein, N. Bandeira, Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking, Nature Biotechnology. 34 (2016) 828–837. https://doi.org/10.1038/nbt.3597.
[48] R. Schmid, D. Petras, L.F. Nothias, M. Wang, A.T. Aron, A. Jagels, H. Tsugawa, J. Rainer, M. Garcia-Aloy, K. Dührkop, A. Korf, T. Pluskal, Z. Kameník, A.K. Jarmusch, A.M. Caraballo-Rodríguez, K.C. Weldon, M. Nothias-Esposito, A.A. Aksenov, A. Bauermeister, A. Albarracin Orio, C.O. Grundmann, F. Vargas, I. Koester, J.M. Gauglitz, E.C. Gentry, Y. Hövelmann, S.A. Kalinina, M.A. Pendergraft, M. Panitchpakdi, R. Tehan, A. Le Gouellec, G. Aleti, H. Mannochio Russo, B. Arndt, F. Hübner, H. Hayen, H. Zhi, M. Raffatellu, K.A. Prather, L.I. Aluwihare, S. Böcker, K.L. McPhail, H.U. Humpf, U. Karst, P.C. Dorrestein, Ion identity molecular networking for mass spectrometry-based metabolomics in the GNPS environment, Nature Communications. 12 (2021) 1–12. https://doi.org/10.1038/s41467-021-23953-9.
[49] J.J.J. Van Der Hooft, J. Wandy, M.P. Barrett, K.E.V. Burgess, S. Rogers, Topic modeling for untargeted substructure exploration in metabolomics, Proc. Natl. Acad. Sci. U.S.A. 113 (2016) 13738–13743. https://doi.org/10.1073/pnas.1608041113.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

[50] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks, Genome Research. 13 (2003) 2498–2504. https://doi.org/10.1101/GR.1239303.

[51] K. Gopi, K. Renu, G. Jayaraman, Inhibition of Naja naja venom enzymes by the methanolic extract of Leucas aspera and its chemical profile by GC–MS, Toxicology Reports. 1 (2014) 667–673. https://doi.org/10.1016/j.toxrep.2014.08.012.

[52] A.A. Aksenov, I. Laponogov, Z. Zhang, S.L.F. Doran, I. Belluomo, D. Veselkov, W. Bittremieux, L.F. Nothias, M. Nothias-Esposito, K.N. Maloney, B.B. Misra, A.V. Melnik, A. Smirnov, X. Du, K.L. Jones, K. Dorrestein, M. Panitchpakdi, M. Ernst, J.J.J. van der Hooft, M. Gonzalez, C. Carazzone, A. Amézquita, C. Callewaert, J.T. Morton, R.A. Quinn, A. Bouslimani, A.A. Orio, D. Petras, A.M. Smania, S.P. Couvillion, M.C. Burnet, C.D. Nicora, E. Zink, T.O. Metz, V. Artaev, E. Humston-Fulmer, R. Gregor, M.M. Meijler, I. Mizrahi, S. Eyal, B. Anderson, R. Dutton, R. Lugan, P.L. Boulch, Y. Guitton, S. Prevost, A. Poirier, G. Dervilly, B. Le Bizec, A. Fait, N.S. Persi, C. Song, K. Gashu, R. Coras, M. Guma, J. Manasson, J.U. Scher, D.K. Barupal, S. Alseekh, A.R. Fernie, R. Mirnezami, V. Vasiliou, R. Schmid, R.S. Borisov, L.N. Kulikova, R. Knight, M. Wang, G.B. Hanna, P.C. Dorrestein, K. Veselkov, Auto-deconvolution and molecular networking of gas chromatography–mass spectrometry data, Nature Biotechnology. 39 (2020) 169–173. https://doi.org/10.1038/s41587-020-0700-3.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Supplementary Information

Metabolite fingerprinting of *Urospatha sagittifolia* (Araceae) tubers at different growth stages by multi-platform metabolomics and molecular networking

Jefferson V. Pastuña[a], Cristian D. Quiroz-Moreno[b], Evencio J. Medina[c], Jessica L. Cooperstone[b,d], Matteo Radice[e], Maria C. Peñuela-Mora[f], José R. Almeida[g], Noroska G. Salazar Mogollón*

[a]Laboratorio de Productos naturales, Universidad Regional Amazónica Ikiam, Km 7 Via Muyuna, Tena, Napo, Ecuador.

[b]Department of Horticulture and Crop Science, The Ohio State University, Columbus, OH 43210, USA.

[c]School of Industrial, Computer and Aeronautical Engineering, University of León, Campus de Vegazana, 24071, León, Spain.

[d]Department of Food Science and Technology, The Ohio State University, Columbus, OH 43210, USA.

[e]Department of Biosciences, Biotechnology and Environment, University of Bari Aldo Moro, Via E. Orabona, 4, 70125, Bari, Italy

[f]Grupo de ecosistemas tropicales y cambio global, Universidad Regional Amazónica Ikiam, Km 7 Via Muyuna, Tena, Napo, Ecuador.

[g]Biomolecules Discovery Group, Universidad Regional Amazónica Ikiam, Km 7 Via Muyuna, Tena, Napo, Ecuador

*Corresponding authors: *E-mail address*: noroska.salazar@ikiam.edu.ec

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

21

22

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

Table of Content

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

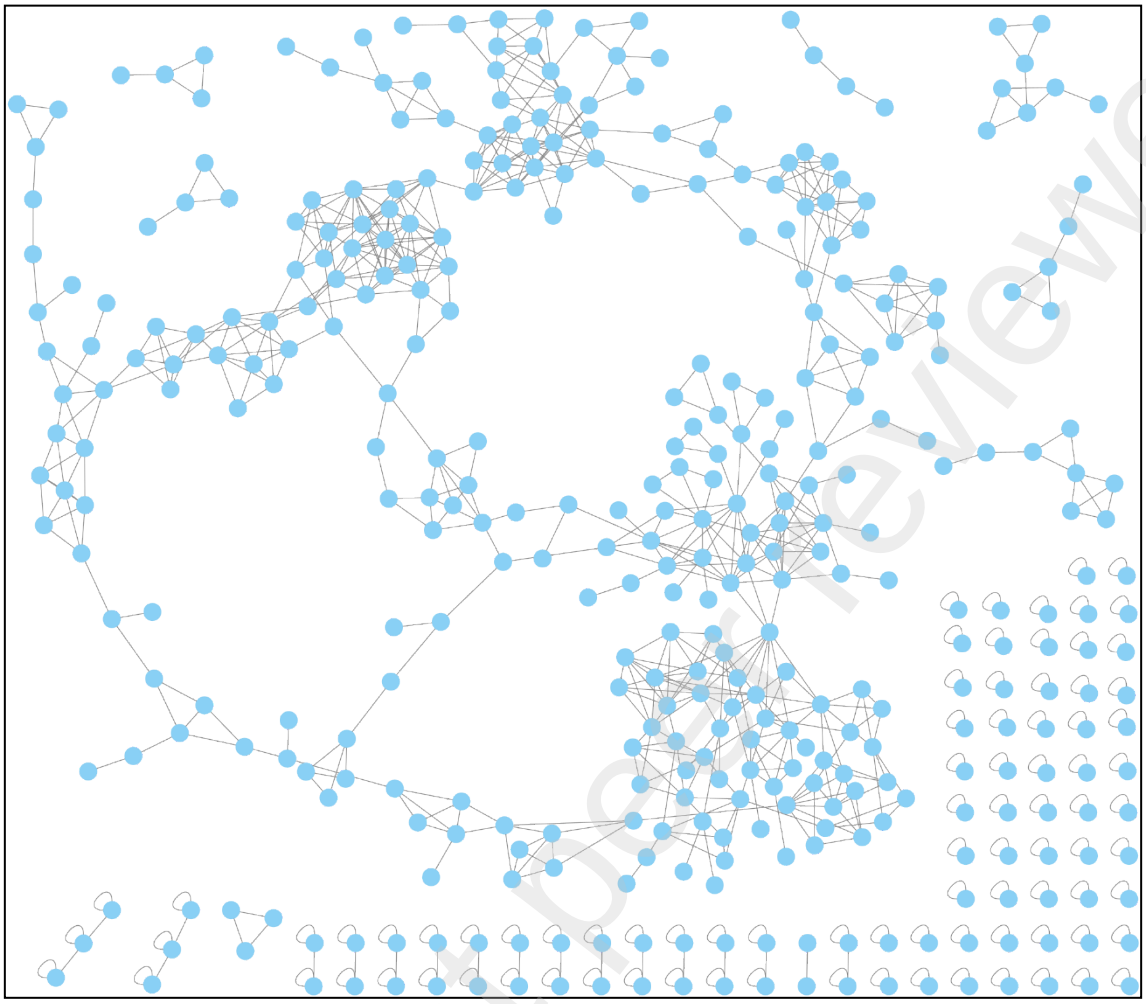

Figure S1. LC-MS/MS Molecular Networking in positive ionization mode.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948

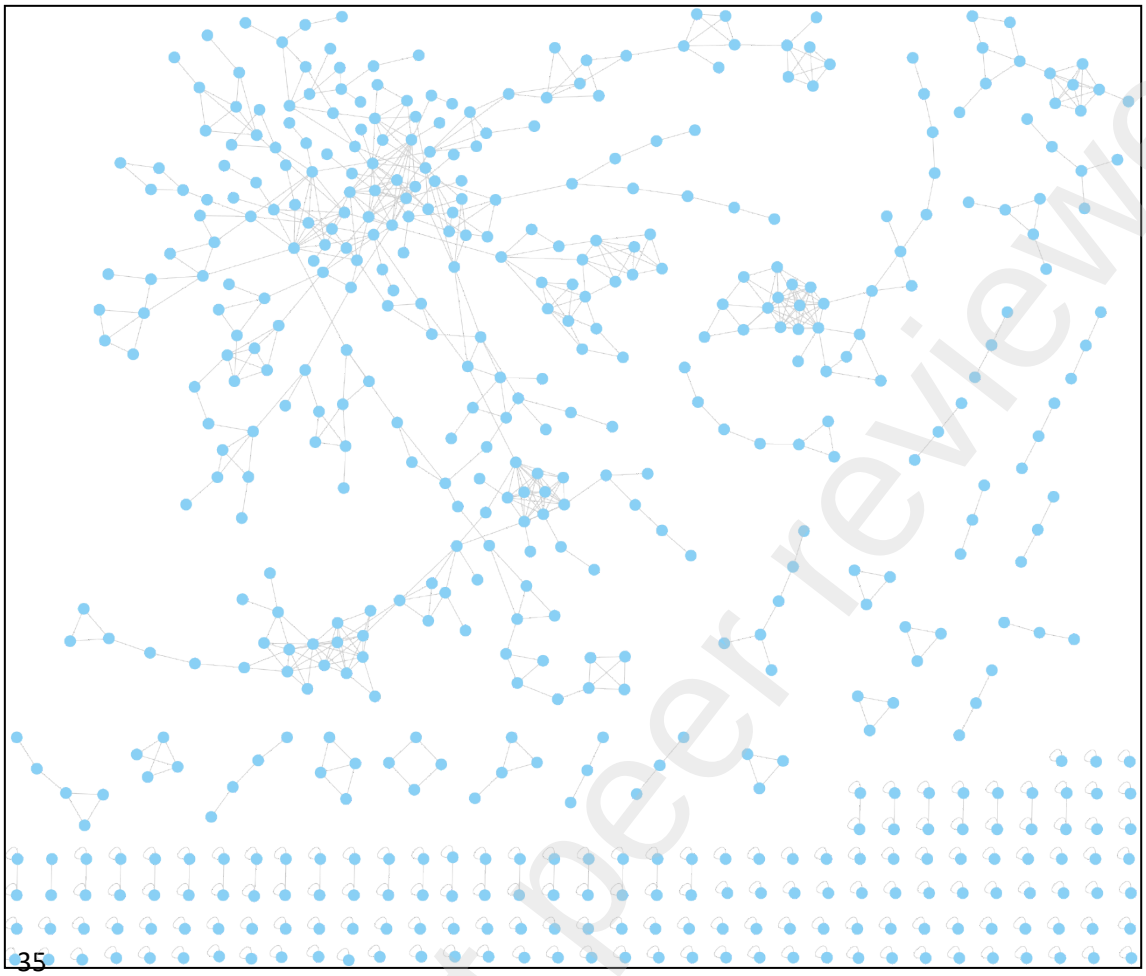

Figure S2. LC-MS/MS Molecular Networking in negative ionization mode.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4680948