

## Research Article

# Bioprospecting the Metabolome of Plant *Urtica dioica* L.: A Fast Dereplication and Annotation Workflow in Plant Metabolomics

Keshab Bhattarai,<sup>1,2</sup> Babita Paudel,<sup>1</sup> Sujan Dahal,<sup>3</sup> Parasmani Yadav,<sup>3</sup> Niraj Aryal,<sup>1</sup> Bikash Baral,<sup>2</sup> and Hari Datta Bhattarai <sup>1,3</sup>

<sup>1</sup>Center for Natural and Applied Sciences (CENAS), Kathmandu, Nepal

<sup>2</sup>Institute of Biological Resources, Kathmandu, Nepal

<sup>3</sup>Central Department of Botany, Tribhuvan University, Kirtipur, Kathmandu, Nepal

Correspondence should be addressed to Hari Datta Bhattarai; [haridatta.bhattarai@cdb.tu.edu.np](mailto:haridatta.bhattarai@cdb.tu.edu.np)

Received 11 September 2021; Revised 15 February 2022; Accepted 17 March 2022; Published 21 April 2022

Academic Editor: Maulidiani M

Copyright © 2022 Keshab Bhattarai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Plants have a pivotal role in ethnopharmacology, and their preparations are in use globally. However, getting down to the structure requires an effective workflow and mostly requires a time-consuming isolation process. Although bioassay-guided approaches are widely popular, they face a massive problem of rediscovery in recent times, especially in plant metabolomics. Mass spectrometry (MS)-based approach incorporated molecular networking via Global Natural Product Social Molecular Networking (GNPS) is considered here for the benefit of the fast screening of secondary metabolites. This study uses direct crude extracts obtained from various parts of the *Urtica dioica* plant for the characterization of secondary metabolites. The crude extract of the plant initially displayed promising antioxidant and anti-diabetic activities. Then, we employed mass spectrometry-based dereplication to identify the phytochemical components in the extracts. This led to the discovery of 7 unknown and 17 known secondary metabolites, which were further verified with the SIRIUS 4 platform, a computational tool for the annotation of compounds using tandem MS data. On the other hand, chasing the antioxidant activity of methanolic extract of *U. dioica* leaves, we employed a bioassay-guided isolation approach. With this method, we isolated and characterized compound **13**, a known molecule, which possessed strong antioxidant activity without showing much toxicity in the brine shrimp lethality test at the test concentration of 1 mg/mL. With our results, we advocate the MS-based approach as a good starting point for the dereplication of compounds from the complex crude extracts of plants.

## 1. Introduction

*Urtica dioica* L. has been in use for centuries in different parts of the world, especially in Asia and Europe, as food and herbal medicines [1–4]. To date, there are a few published scientific literatures about the utilization of *U. dioica* L. as an herbal medicine. In Nepal, as this plant possesses high therapeutic potential, these plants have been exploited for centuries by local tribes to cure different ailments, such as pain, inflammation, wound healing, and diabetes [5]. Several biological activities such as antimicrobial, antioxidant [6], hypoglycemic activity [7–10] of crude extract of *U. dioica* L., and other genera of *Urtica* have also been reported. Usage of various advanced

techniques such as gas chromatography mass spectrometry (GC-MS) [11] and high-resolution mass spectrometry (HRMS) [12] has been performed for phytochemical profiling of this plant. A total of 35 new compounds have been reported from the plants belonging to the family Urticaceae in the period of 2000–2010 and among them, a new flavone has been mostly reported from *U. dioica* L. [13].

In regard to chemical analysis, mass spectrometry has been a recent favorite analytical tool also used for the analysis of natural products. Due to its high sensitivity and time advantage, it has broad applicability. HRMS data are utilized in various computational platforms to visualize and interpret the chemical repertoire that can be detected using

MS [14]. In that context, Global Natural Product Social Molecular Networking (GNPS) has been widely popular to analyze MS2 raw data. With an increasing number of articles published each year (Figure 1), molecular networking-based dereplication has become one of the powerful tools for studying a diverse class of low molecular weight molecules in a single attempt [14]. Molecular networking works based on clustering structurally related metabolites having comparatively similar MS2 fragments. This allows instantaneous visual investigation of identical molecules, analogs, or compound classes. The networking system assembles mass spectrometric data by mining the spectral similarity between the available MS2 fragmentation patterns of structurally related metabolites. An individual node in the network corresponds to one consensus MS2 spectrum connected to other nodes with common fragmentation patterns of edges. Similarly, MS2 raw data can also be utilized in another *in silico* platform called SIRIUS 4, where the authors have claimed to provide annotation efficiency over 70% in the complex metabolomics data sets [15]. This technique has been successfully applied for the overall evaluation of untargeted and targeted secondary metabolites inhabited from diverse natural sources such as plants, animals, bacteria, fungi, and sponges [16–18]. This study has also relied on MS-based molecular networking to partially characterize compounds from methanolic extracts of leaves, stems, and roots of *U. dioica* L.

Here, we present the workflow (Figure 2), which begins with the acquisition of MS data and guides through the annotation of compounds in the crude extracts. While on the other hand, bioassay-guided isolation of the active compounds has been performed. The dereplication of relatively abundant known and unknown secondary metabolites has been performed to determine the plant species as a potential source of novel natural products.

## 2. Materials and Methods

**2.1. Plant Collection.** *U. dioica* L. was harvested from Kathmandu, Nepal (27.6863°N, 85.2949°E). The plant was authenticated by comparing the authentic voucher specimen deposited in Tribhuvan University Central Herbarium (TUCH) at Central Department of Botany, Tribhuvan University, Kirtipur, Kathmandu and deposited at the Center for Natural and Applied Sciences (CENAS), Kathmandu, Nepal, under the accession number UKIB\_33.

**2.2. Preparation of Crude Extracts.** Air-dried plant samples leaves, stems, and roots (300 g each) were ground to powder and were extracted with a mixture of water and methanol (10 : 90). The filtrate was dried using a rotary evaporator. The obtained crude extract was dissolved in 200 mL of distilled water and extracted with 600 mL of hexane to remove non-polar fatty compounds. The remaining water phase was further extracted with 600 mL of dichloromethane (DCM) to remove medium polar compounds. In addition, the remaining water phase was freeze-dried and dissolved in methanol. Thus, hexane, DCM, and methanol fractions were prepared.

**2.3.  $\alpha$ -Glucosidase Inhibition Assay.** The  $\alpha$ -glucosidase inhibitory activities were assayed following [19] protocol with slight modifications. An aliquot of 100  $\mu$ L of the test samples of various concentrations (0–100  $\mu$ g/mL) was mixed with 40  $\mu$ L of baker's yeast  $\alpha$ -glucosidase preincubated at 37°C for 10 min. Then, the reaction was initiated with 1 mM of p-nitrophenyl- $\alpha$ -glucopyranoside (PNPG) in 50 mM phosphate buffer, pH 7.0, in a total volume of 1 mL. The reaction was terminated by adding 1.0 mL of 0.5 M Tris solution, and the absorbance of p-nitrophenol (PNP) released from PNPG was measured at 400 nm with a UV-visible spectrophotometer (Shimadzu, Japan). Reaction mixture without test samples was taken as negative control, and the reaction mixture with acarbose was employed as positive control.

**2.4. Antioxidant Activity.** Diphenylpicrylhydrazyl (DPPH) free radical scavenging activity of the samples was analyzed as described previously [20]. One mL of DPPH solution (0.1 mM of DPPH in methanol) was mixed with 3 mL of various concentrations (0–100  $\mu$ g/mL) of the test samples. The mixture was incubated at room temperature (RT) for 30 min, and the absorbance was measured at 517 nm in a UV-visible spectrophotometer. Reaction mixtures without the test sample and with BHA were used as negative and positive controls, respectively. The experiments were performed in triplicate.

**2.5. Toxicity Assay.** Brine shrimp lethality test was performed as described previously [21]. Briefly, the eggs of *Artemia salina* were hatched in aerated seawater equivalent in light at 25°C. The active larvae ( $n = 100$ ) were selected and treated with various concentrations of test samples (0–1,000  $\mu$ g/mL). The effectiveness of the test samples was monitored after 24 h of treatment by observing the live larvae. The mortality rate of the larvae indicated the toxicity of test samples. Therefore, berberine chloride, a standard anticancer drug, was taken as a positive control, and brine shrimp larvae in salty water were taken as a negative control.

**2.6. Bioactivity-guided Isolation of Antioxidant Compounds.** The water fraction (1 g) of leaf extract was subjected to vacuum liquid chromatography using C<sub>18</sub>ODS column as the stationary phase and methanol–water mixture (0–100%) as a mobile phase. The fraction (280 mg) obtained at 30% of methanol in water showed strong DPPH reducing activity. For isolation of the active compound, the fraction was subjected to high-performance liquid chromatography (HPLC) using a C<sub>18</sub>ODS column (Phenomenex, size 250 mm  $\times$  10 mm, pore size 5  $\mu$ m) with the mobile phase as an acetonitrile–water mixture (2 mL/min). The solvent gradient was run as ACN/H<sub>2</sub>O (0.1% acetic acid); 0 min: 5/95, 25 min: 20/80, 28 min: 100/0, and continued up to 30 minutes. The isolated compound with an absorbance of 280 nm wavelength at 4.6 min displayed strong DPPH reducing activity. The purified compound was subjected ahead for structure elucidation.

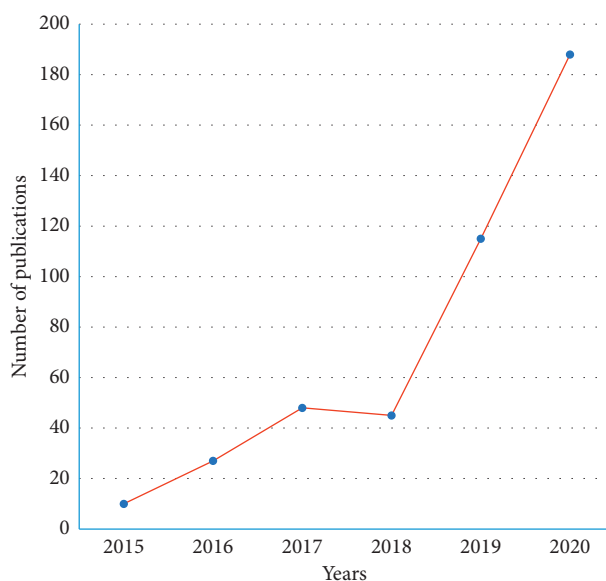


FIGURE 1: A graph showing article numbers published each year since 2016 involving GNPS for compound dereplication.

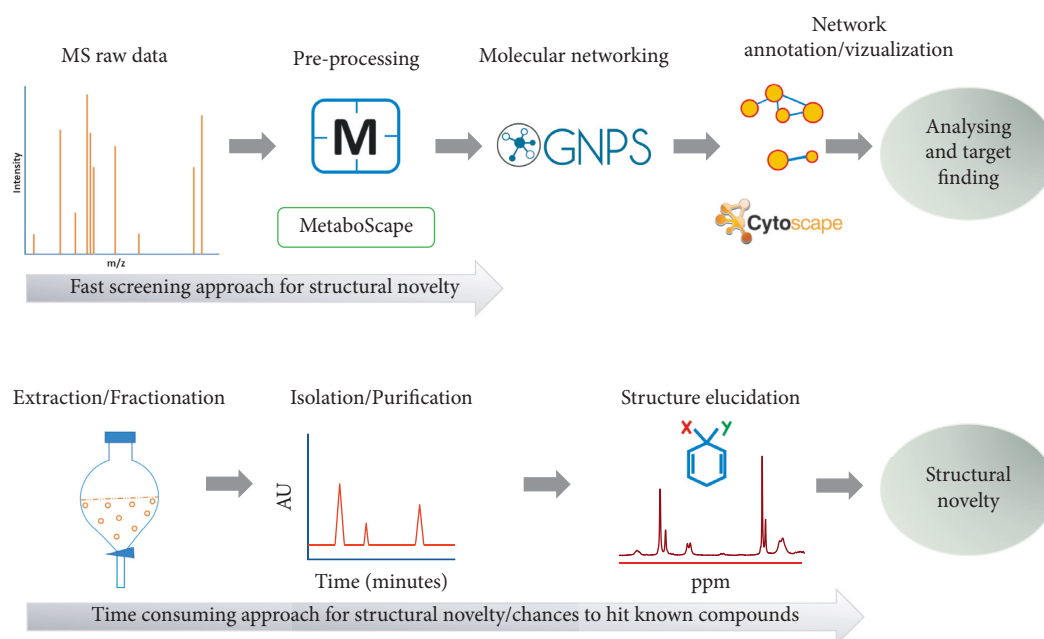


FIGURE 2: Metabolomics workflow showing the molecular networking approach and conventional isolation approach for finding structural novelty.

### 2.7. Identification and Characterization of Compounds.

The molecular structure of the purified compound was elucidated using spectroscopic techniques, such as nuclear magnetic resonance (NMR) spectroscopy (Bruker Avance III HD spectrometer;  $^1\text{H}$ : 400 MHz;  $^{13}\text{C}$ : 100 MHz, COSY, HSQC, and HMBC experiments in  $d_4$ -methanol, 298K), and high-resolution mass spectrometry (HRMS). A 5 mm SMART probe head was used for NMR measurement of the purified compound. Once the structure was fully confirmed, the compound was searched in SciFinder, a database developed and offered by Chemical Abstracts Service (CAS). If the hit is found in the SciFinder, the compound is known, if

not, then the compound is considered as new. Similarly, the molecular formulae for the compounds were initially predicted with Bruker Compass Data Analysis 4.0 software and further assured with SIRIUS 4 platform. These formulae were taken as input in SciFinder. Once the hits were not found within the database, these formulae of the respective compounds were also considered as unknown metabolites.

**2.8. LC-HRMS Based Metabolomics.** The samples of 0.2 mg/mL crude extract from the leaf, stem, and roots of the plant *U. dioica* in HPLC grade methanol were prepared to

measure LC-HRMS. The data obtained from LC-HRMS were first analyzed in Bruker Compass DataAnalysis 4.0 software to assure the proper ionization of the compounds. Only the positively ionized HRMS data were processed through Bruker MetaboScape 3.0 software, omitting the poorly ionized negative mode data. The data were processed here for bucketing all molecular ions in a single file. Some specific parameters such as intensity threshold (counts); 5000, minimum peak length (spectra); 4, minimum peak length (recursive); 3, EIC correlation; 0.8 and mass calibration; true were optimized during bucketing. The adduct ions were confirmed  $[M+H]^+$  as primary ion,  $[M+Na]^+$ ,  $[M+K]^+$ ,  $[2M+H]^+$ ,  $[2M+Na]^+$ ,  $[M+2H]^{2+}$ ,  $[M+2Na]^{2+}$  as seed ions and  $[M-H_2O+H]^+$ ,  $[M+H_2O+H]^+$ ,  $[2M+H_2O+H]^+$ ,  $[2M-H_2O+H]^+$  as common ions during processing. From the raw MS chromatogram, the retention time range of 4–45 minutes and molecular mass range of 200–1500  $m/z$  were considered with the average MS2import method for generating the bucket list of diverse compounds adopting these parameters.

After bucketing, the software enlisted the molecular ions observed in the HRMS chromatogram of their related extract sources (leaves, root, and stem) into a column with their respective retention time, adduct ions or parent ions, molecular weight, and abundance (relative ion intensity). The bucket list of these molecular ions was exported in the mascot generic format (MGF) files, which was later unzipped into five different file formats, including Excel and GNPS file types. The GNPS file format was uploaded through the GNPS platform for molecular networking, where the GNPS web server clusters the tandem mass compounds in a molecular network based on minimum matched fragment ions; 5 and minimum cosine score; 0.7. The network data retrieved from the GNPS platform were analyzed and annotated via Cytoscape 3.8.2 [22]. Node and edge attributes were used, where nodes are displayed as a colorful pie chart to visualize the relative abundance of each ion in three sources of extracts (leaf, root, and stem). A single network was generated for the positive ionization modes using the GNPS 2 platform (Supplementary file S1). The list of known and putative unknown metabolites sorted by the relative abundance was generated from molecular networking. These compounds were additionally verified using CSI: FingerID analysis of individual raw MS data in SIRIUS 4 platform via molecular structure database search with  $m/z$  tolerance 20.

**2.9. Statistical Analysis.** All bioassay data are expressed as mean  $\pm$  SD from a minimum of three biological replicates. In addition, 50% inhibition concentration ( $IC_{50}$ ) was calculated by linear regression analysis of the obtained data using Microsoft Excel 2016.

### 3. Results

**3.1. Antidiabetic and Antioxidant Activities.** Methanolic extract of the stem of *U. dioica* L. showed  $\alpha$ -glucosidase inhibition activity in a concentration-dependent manner. However, the dichloromethane and hexane fractions did not

show  $\alpha$ -glucosidase inhibition activity at the test concentration range up to 100  $\mu\text{g/mL}$ . On the other side, the methanol fraction of *U. dioica* L. leaves displayed comparably strong DPPH scavenging activity (Table 1). This result indicated that *U. dioica* L. could be a potent source of antioxidants. Therefore, further work on the isolation and characterization of active compounds was performed using the bioassay-guided liquid chromatography technique.

Compound **13** (5 mg) was obtained as a brownish-yellow gum. The high-resolution electrospray ionization mass spectrometry (HR-ESIMS) data of compound **13** shows a molecular weight of  $m/z$  339.1054  $[M+Na]^+$  with a molecular formula  $C_{14}H_{20}NaO_8$  and five degrees of unsaturation. The  $^1\text{H}$  NMR spectrum of compound **13** (Table 2, Supplementary Figures S27–S32) showed three aromatic proton signals of a metal substituted aromatic ring system and two coupled methylene signals. The  $^{13}\text{C}$  NMR spectrum showed the presence of a benzene ring along with one glucose molecule. In the HMBC spectrum, the key HMBC correlations of H-7 ( $\delta\text{H}$  2.60)/C-1 ( $\delta\text{C}$  130.5), C-2 ( $\delta\text{C}$  117.4), C-6 ( $\delta\text{C}$  123.2), confirmed that an ethoxyl group was linked to C-1 of the benzene ring as shown in Figure 3. In addition, an anomeric proton [ $\delta\text{H}$  4.95 (1H, d,  $J=7.9$  Hz)] and a characteristic proton signal of allose-H-3 at  $\delta\text{H}$  3.97 (1H, t,  $J=2.8$  Hz) were observed in  $^1\text{H}$  NMR spectrum, indicating the presence of an allopyranosyl unit. The site of allopyranosyl on the benzene ring of compound **13** was determined by the long-range HMBC correlation of H-1' ( $\delta\text{H}$  4.95)/C-3 ( $\delta\text{C}$  145.4). Thus, the structure of compound **13** was depicted as 3, 5-dihydroxyphenethanol 3-O- $\beta$ -D-allopyranoside. This compound was reported here for the first time from *U. dioica* L.

Compound **13** displayed antioxidant activity in terms of DPPH scavenging capacity. Its activity was higher than BHA (Table 1). However, it did not show antidiabetic activity and was nontoxic to brine shrimp at the test concentration of 100  $\mu\text{g/mL}$ . This indicates that the compound is nontoxic to eukaryotic cells.

**3.2. Mass Spectrometry and Compound Annotation.** Out of 12,046 molecular ions enlisted in the bucket list from MetaboScape, 2,859 (23.73%) molecular ions were found to have MS2 spectra, which were only reflected as nodes in the molecular network (Supplementary Figure S1). Among them, 1,481 molecular ions are self-looped nodes, 332 molecular ions formed 166 different clusters, each with holding two nodes, and the rest 1,039 molecular ions formed different bigger clusters. Among the ions in the network, the GNPS platform gave the hits of 228 molecular ions as known metabolites based on matching the GNPS library. In the present study, only seven unknown metabolites and 17 known metabolites have been reported based on their relative abundance (Tables 3 and 4). They were further assured with HRMS-based dereplication, CSI: FingerID via SIRIUS 4, and SciFinder as a computational approach for searching molecular structure databases.

Seventeen abundant compounds of known structural fingerprints (Supplementary Figure S2) in the plant extract

TABLE 1:  $\alpha$ -Glucosidase inhibition, DPPH free radical scavenging, and brine shrimp toxicity activities of *U. dioica* L.

Test samples	$\alpha$ -glucosidase inhibition (IC <sub>50</sub> ) ( $\mu$ g/mL)			DPPH free radical inhibition (IC <sub>50</sub> ) ( $\mu$ g/mL)			Brine shrimp toxicity (1 mg/mL)
	ME	DCME	HE	ME	DCME	HE	ME
<i>U. dioica</i> L. (stem)	62.5 $\pm$ 5.2	NA	NA	38.9 $\pm$ 1.9	NA	NT	7%
<i>U. dioica</i> L. (root)	NA	NA	NA	12.1 $\pm$ 0.6	NA	NT	8%
<i>U. dioica</i> L. (leaf)	NA	NA	NA	11 $\pm$ 1.2	26.3 $\pm$ 1.4	NT	7%
Acarbose (+ve standard)	50 $\pm$ 3.6			NT	NT	NT	NT
Potassium dichromate	NT			NT			100%
Compound <b>13</b> from <i>U. dioica</i> L. leaf	NA			4.5 $\pm$ 0.2			NA
BHA	NT			4.8 $\pm$ 0.2			NT

Mean  $\pm$  SD; ME, methanol extract; DCME, dichloromethane extract; HE, hexane extract; IC<sub>50</sub>, 50% inhibition of activity; NA, no activity; NT, not tested; BHA, Butylatedhydroxyanisole.

TABLE 2: NMR spectroscopic data of compound **13** (<sup>1</sup>H: 400 MHz; <sup>13</sup>C: 100 MHz, in CD<sub>3</sub>OD).

	$\delta$ C	$\delta$ H (J in hz)	HMBC	COSY
1	132.2	—	—	—
2	119.9	7.07 (d, 1.3)	39.5, 132.8, 125.2, 146.7	—
3	146.7	—	—	—
4	116.9	6.81 (d, 1.0)	146.7, 146.6, 119.9	—
5	146.6	—	—	—
6	125.2	6.77 (d, 1.0)	146.6, 119.9, 132.8	—
7	39.5	2.69 (t, 7.3)	132.8, 62.5, 132.8, 125.2, 119.9	3.92
8	62.5	3.92 (t, 7.3)	39.5, 132.8	2.69
1'	104.4	4.78 (d, 7.9)	146.7, 77.7	3.47
2'	77.7	3.47 (dd, 7.9, 2.8)	104.4, 74.9	4.78
3'	74.9	3.47 (t, 2.8)	77.7, 78.4	—
4'	78.4	3.35 (dd, 9.8, 2.8)	74.9, 71.4	—
5'	71.4	3.47	78.4, 64.3	—
6'	64.3	3.74	71.4, 78.4	—
		3.46		

were annotated individually based on MS2 data (Supplementary Figure, S3–S19) and fragmentation tree from SIRIUS 4. The MS2 key fragment ions with their annotated fragment motifs were analyzed thoroughly. HRMS data of compound **13** showed MS2 ionization from the precursor ion of sodium adduct  $[M + Na]^+$  and was clustered as self-looped node unlike other known clusters as shown in Figure 4.

Similarly, seven unknown metabolites annotated via MS2 fragmentation pattern comprised of four different clusters, as shown in Figure 5. These clusters selectively hold three peptides, one phenylalanine peptide, one fatty acyl glycoside, and two putative compounds (Table 4). These annotations were further assured by importing MS2 data to our SIRIUS 4 workflow (Supplementary Figures S20–S26 and S33–S39).

In the peptide class, all three members **1'**, **2'**, and **3'** compounds lose a common mass of 147 Da (C<sub>9</sub>H<sub>9</sub>NO), indicating a typical phenylalanine motif. Compounds **1'** and **2'** further lose 99 Da (C<sub>5</sub>H<sub>9</sub>NO), which possibly refers to valine as a fragment. Fragment ions at  $m/z$  288.1709 and 328.1666 from compounds **1'** and **2'**, respectively, represent further loss of proline motif, while fragment ion  $m/z$  316.1659 from compound **3'** indicates further loss of leucine

or isoleucine motif and 97 Da (C<sub>5</sub>H<sub>7</sub>NO). The common fragment at  $m/z$  259.1445 (C<sub>15</sub>H<sub>19</sub>N<sub>2</sub>O<sub>2</sub>) of all these compounds represents the characteristic fragment as a dipeptide motif, so here we conclude that the compounds might belong to the peptide class.

Compound **6'** gives a characteristic fragment at  $m/z$  163.0603 (C<sub>6</sub>H<sub>11</sub>O<sub>5</sub>), possibly referring to a fructan monomer unit. The compounds **6'** and **7'** have a common fragment at  $m/z$  69.0335 (C<sub>4</sub>H<sub>5</sub>O), indicating a furan ring as a stable modified arrangement of a fructan monomer. The rest of the fragments seem ambiguous, suggesting the compounds having an unknown structure.

In the phenylalanine-peptide class, compound **5'** shows a strong fragment at  $m/z$  445.1761 with continuous loss of ammonia (17 Da) and carbonyl (28 Da) motifs. The fragment further loses carbonyl and water units continuously to yield stable fragments at  $m/z$  417.1812 and 399.1709, respectively. The mutual fragments at  $m/z$  462.2030, 445.1761, and 427.1655 lose 147 Da from their individual skeletons to give stable fragments at  $m/z$  315.1342, 298.1078, and 280.0973, respectively, which is the crucial evidence of phenylalanine embedded in the structure. Another stable fragment at  $m/z$  269.1287 arises from the original molecule after the successive loss of three carbonyl motifs, one ammonia, one

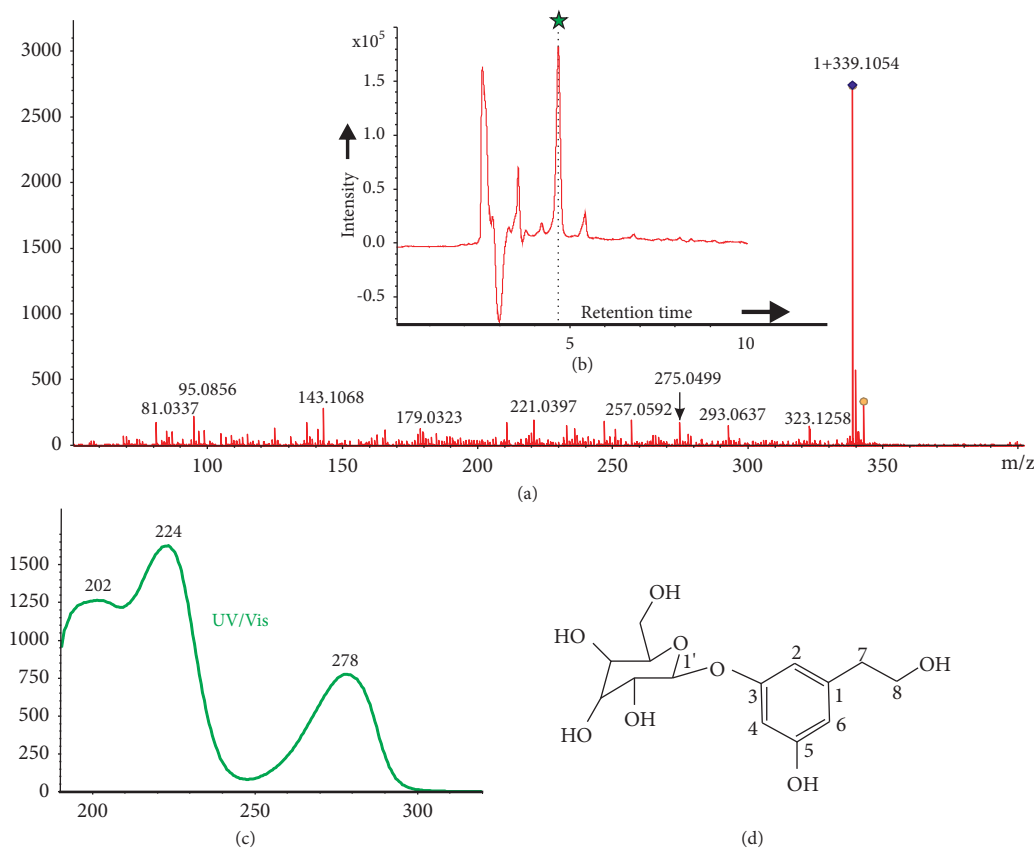


FIGURE 3: Structure of compound **13** showing chromatographic, spectroscopic, and spectrometric features. (Note. (a) MS2 ionization spectra from HRMS data, (b) UV chromatogram in HPLC system with eluted peak ( $R_T$  4.6 min), (c) maximum UV absorbance curve, and (d) elucidated chemical structure based on NMR spectroscopy).

water, and other subunits, which possibly indicate at least two peptide bonds present in the original compound. This classified the compounds possibly belonging to the phenylalanine peptide family.

In the compound class of fatty acyl glycosides, the stable fragment at  $m/z$  209.1539 refers to the polyketide skeleton ( $C_{13}H_{21}O_2$ ) with 4-ring double bond equivalence. This polyketide fragment further loses the water units, giving successive fragments at  $m/z$  191.1437 ( $C_{13}H_{17}O$ ). The fragment resembled dehydrated glucose at  $m/z$  163.0603 ( $C_6H_{10}O_5$ ), with further loss two water subunits one after another to yield stable fragments at  $m/z$  145.0499 ( $C_6H_9O_4$ ) and 127.0391 ( $C_6H_7O_3$ ), respectively. These fragmentation patterns indicate the compound-bearing polyketide chain and glycosyl unit, comprising the family fatty acyl glycosides.

#### 4. Discussion

Two major methods for the dereplication of bioactive secondary metabolites have been emphasized independently in this work from the plant *U. dioica* L. Among these two, the bioactivity-guided approach is a conventional method and was found to be time-consuming. The isolation of compound **13** was done with a bioassay-guided isolation approach. Upon further analysis, its structure was found to be reported earlier from *Centaurea urvillei* DC. subsp. *urvillei*

with anticancer activity against breast cancer cell lines [36] and *Veronica thymoides* subsp. *pseudocinerea* [40] with antioxidant activity. However, compound **13** was reported for the first time from the genus *Urtica*. The isolation of other compounds from bioactive fractions was not achievable in our study, perhaps they were low UV active or lost to column during fractionation. So it is not always the case in bioassay-guided isolation that every bioactive component is achievable with the limited technical resources.

While, on the other hand, the HRMS dereplication approach combined with molecular networking predicts possible structural hits within a short time, allowing someone to prioritize the crude fractions from very early on. This priority could be based on structural novelty, compound size, and abundance. This can save us more time and resources from the repetitive isolation of known chemical entities. Although NMR-based structure elucidation is the gold standard in NP chemistry, there are still limitations to this technique. Here also, compound **13** was already predicted as 3,5-dihydroxyphenethanol-3-O-D-allopyranoside via SIRIUS 4 platform and later confirmed via spectroscopic analysis after extensive isolation.

In our study, we also find that the quality of MS2 data is key for structural fingerprint findings. These fingerprints are utilized by publicly available tools like SIRIUS 4.

TABLE 3: Predicted known metabolites annotated and dereplicated with tandem MS data from *U. dioica* L.

Comp No.	<sup>12</sup> C m/z	Parent ion	Molecular formula	Error ppm	R <sub>T</sub> (min)	Proposed structure	Annotation <sup>#</sup> (spectral data source) <sup>&lt;</sup>	Total ions intensity	Leaf ext. int.	Root ext. int.	Stem ext. int.	Ref.
1	565.1553	[M+H] <sup>+</sup>	C <sub>26</sub> H <sub>38</sub> O <sub>14</sub>	0.1	13.8	Vicenin 3	CCMSLIB000004696062*	3360987	1383074	274552	1703361	[23]
2	360.1078	[M+H] <sup>+</sup>	C <sub>18</sub> H <sub>17</sub> NO <sub>7</sub>	0.1	15.2	Clovamide	CCMSLIB00000845552*	3241983	23828	2980171	237984	[24]
3	579.1716	[M+H] <sup>+</sup>	C <sub>27</sub> H <sub>30</sub> O <sub>14</sub>	1.4	14.8	Violanthin	CCMSLIB000004695807*	1228234	519225	52679	656330	[25]
4	595.1668	[M+H] <sup>+</sup>	C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>	1.7	12.7	Vicenin 2	CCMSLIB000005745586*, PR303371 <sup>λ</sup> , PR310993 <sup>λ</sup>	1037798	375851	80058	581889	[26]
5	387.2019	[M+H] <sup>+</sup>	C <sub>19</sub> H <sub>30</sub> O <sub>8</sub>	1.4	13.6	Roseoside	CCMSLIB00000851547*	935573	330113	261540	343920	[27]
6	595.1662	[M+H] <sup>+</sup>	C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>	0.8	16.6	Nicotifolin	CCMSLIB00000846155*	886166	205035	28100	653031	[28]
7	265.1434	[M+H] <sup>+</sup>	C <sub>15</sub> H <sub>20</sub> O <sub>4</sub>	0.0	14.6	Curcuminol G	CCMSLIB000004715428*	854626	8106	606021	240499	[29, 30]
8	279.1707	[M+H] <sup>+</sup>	C <sub>15</sub> H <sub>22</sub> N <sub>2</sub> O <sub>3</sub>	1.4	19.6	Phe-leu	CCMSLIB000003134492*	646071	202418	48413	395240	[31]
9	263.1281	[M+H] <sup>+</sup>	C <sub>15</sub> H <sub>18</sub> O <sub>4</sub>	1.1	15.5	Austricin	—	544856	327953	58519	158384	[32]
10	343.1179	[M+H] <sup>+</sup>	C <sub>19</sub> H <sub>18</sub> O <sub>6</sub>	0.7	15.6	Tetramethoxy flavone	—	521797	2555	423546	95696	[33]
11	285.2215	[M+H] <sup>+</sup>	C <sub>20</sub> H <sub>28</sub> O	0.7	18.3	Retinene	—	478968	10869	456532	11567	[34]
12	227.1281	[M+H] <sup>+</sup>	C <sub>12</sub> H <sub>18</sub> O <sub>4</sub>	1.2	15.9	12-Hydroxyjasmonic acid	CCMSLIB00000856124*	390475	103512	53281	233682	[35]
13	339.1054	[M+Na] <sup>+</sup>	C <sub>14</sub> H <sub>20</sub> O <sub>8</sub>	1.7	9.9	3, 5-dihydroxyphenethanol 3-O-β-D	—	318849	160121	63498	95230	[36]
14	341.1384	[M+H] <sup>+</sup>	C <sub>20</sub> H <sub>20</sub> O <sub>5</sub>	0.1	15.1	Flavaprenin	—	313196	10593	284531	18072	[37]
15	322.2018	[M+H] <sup>+</sup>	C <sub>18</sub> H <sub>27</sub> NO <sub>4</sub>	1.5	15.8	O-hydroxycapsaicin	—	256745	150679	4759	101307	[38]
16	536.1762	[M+H] <sup>+</sup>	C <sub>25</sub> H <sub>29</sub> NO <sub>12</sub>	0.0	15.0	Davallioside A	—	205433	0	181536	23897	[39]
17	743.2035	[M+H] <sup>+</sup>	C <sub>32</sub> H <sub>38</sub> O <sub>20</sub>	0.8	14.2	Quercetin 3-(2G-xylosylrutinoside)	CCMSLIB00000845960*	149503	30820	7076	111607	[30]

\*GNPS library spectrum ID, # SIRIUS 4 verified, λ Massbank spectral access.

TABLE 4: Predicted novel compounds annotated and dereplicated with tandem MS data from *U. dioica* L.

Comp No.	$^{12}\text{C}$ m/z	Parent ion	Molecular formula	Error ppm	R <sub>T</sub> (min)	Predicted compound class	CSI: FingerID (score)	Relative ions intensity (leaf_stem_root)	Abundance priority
1	631.3237	[M+H] <sup>+</sup>	C <sub>34</sub> H <sub>42</sub> N <sub>6</sub> O <sub>6</sub>	0.0	17.2	Putative peptide	69.75%	0.002_0.004_0.994 Cluster-1'	Root
2	671.3555	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>46</sub> N <sub>6</sub> O <sub>6</sub>	1.5	21.6	Putative peptide	56.69%	0.004_0.008_0.989 Cluster-1'	Root
3	673.3709	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>48</sub> N <sub>6</sub> O <sub>6</sub>	0.2	20.0	Putative peptide	61.42%	0.557_0.400_0.042 Cluster-1'	Leaf and stem
4	550.2863	[M+H] <sup>+</sup>	C <sub>25</sub> H <sub>43</sub> NO <sub>12</sub>	0.9	15.1	Fatty acyl glycoside	—	0.558_0.407_0.035 Cluster-2'	Leaf and stem
5	490.1973	[M+H] <sup>+</sup>	C <sub>27</sub> H <sub>27</sub> N <sub>3</sub> O <sub>6</sub>	0.2	13.7	Phenylalanine-peptide derivative	50.68%	0.004_0.006_0.990 Cluster-3'	Root
6	432.1714	[M+H] <sup>+</sup>	C <sub>15</sub> H <sub>29</sub> NO <sub>13</sub>	3.3	5.0	Putative compound	—	0.373_0.203_0.424 Cluster-4'	Leaf, stem, and root
7	446.1874	[M+H] <sup>+</sup>	C <sub>16</sub> H <sub>31</sub> NO <sub>13</sub>	1.5	5.1	Putative compound	51.79%	0.096_0.545_0.359 Cluster 4'	Stem and root

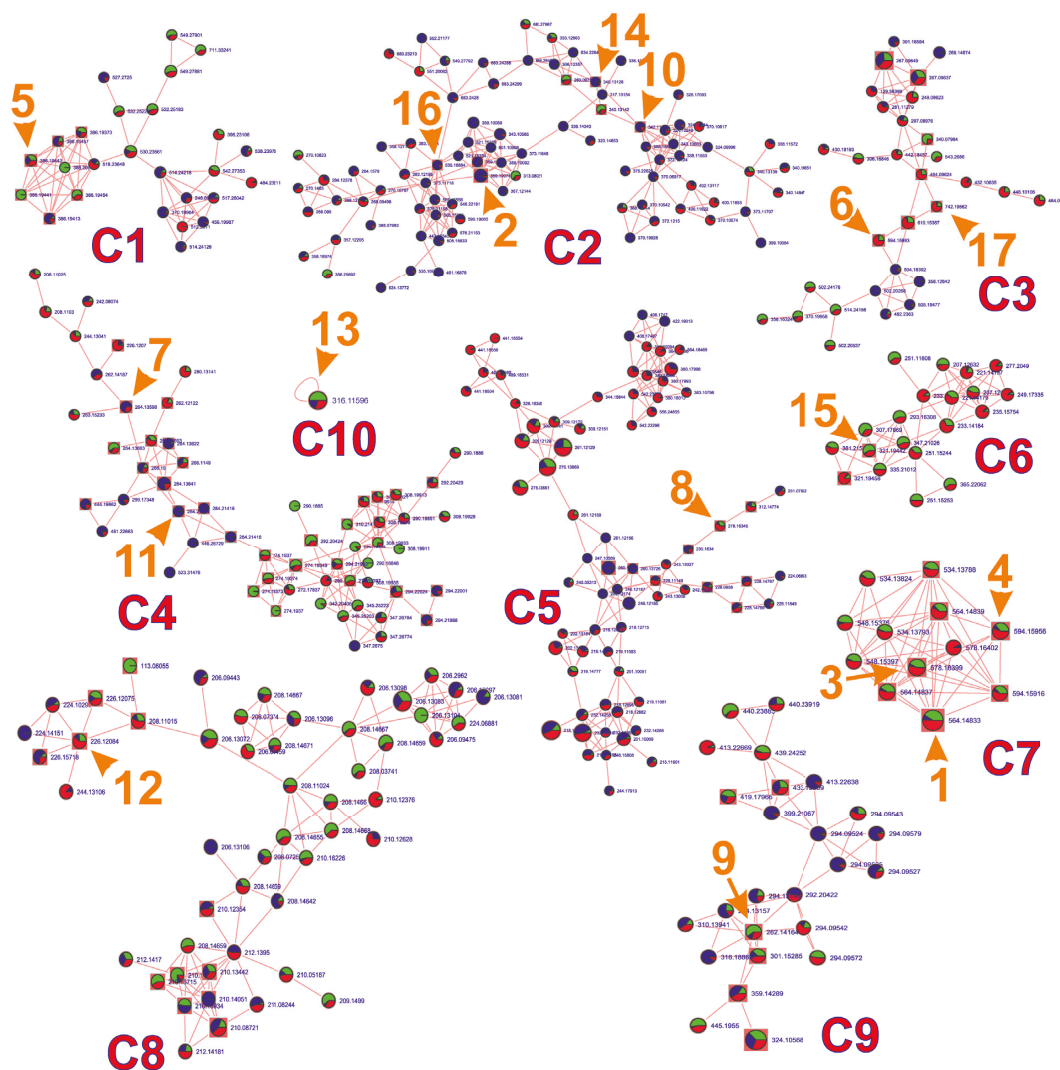


FIGURE 4: Molecular networking of MS2 data from *U. dioica* showing nine clusters including 17 known metabolites. (Note. Each node was identified by the neutral mass of the compound. The three colors green, red, and blue in each node represent the intensity of the compound detected in three different parts of the plant extract; leaf, stem, and root, respectively. The edge thickness in the network represents the similarity extent among the connected nodes).



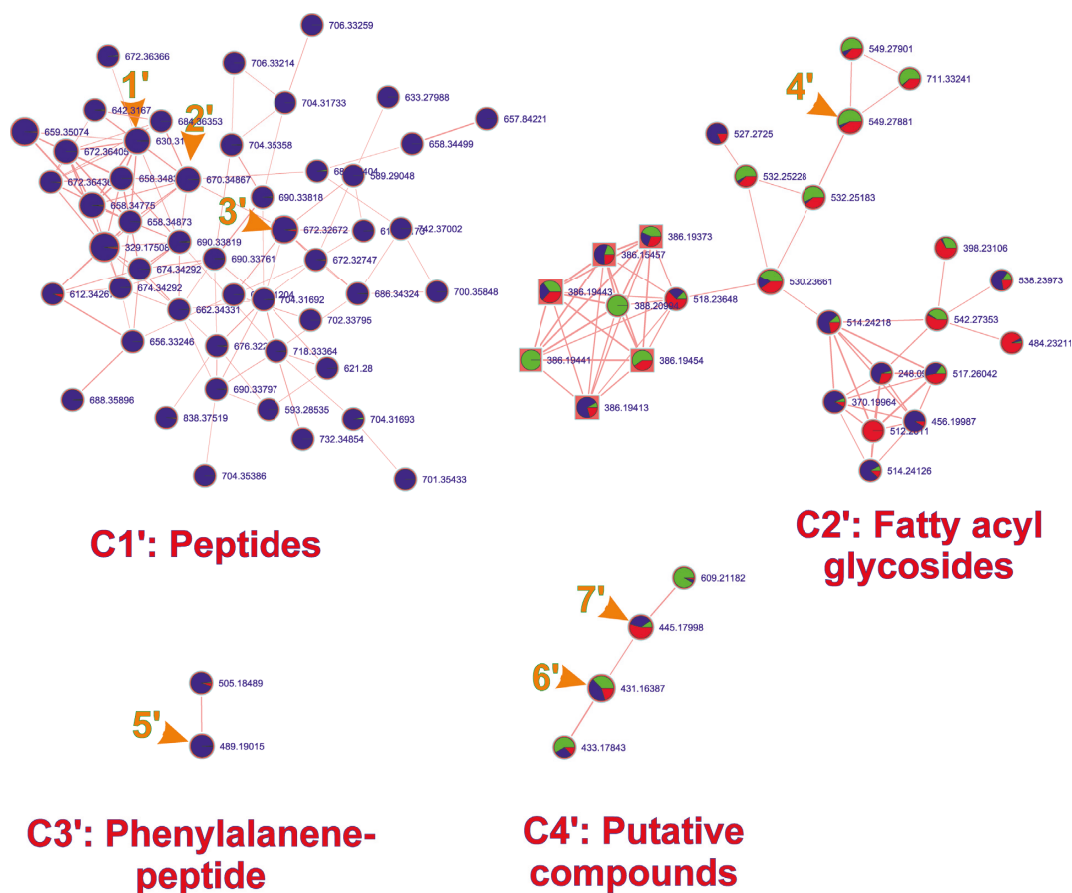


FIGURE 5: Molecular networking of MS2 data from *U. dioica* showing four clusters with seven unknown metabolites. (Note. Each node was identified by the neutral mass of the compound. The ratio of three colors as green, red, and blue in each node represents the relative intensity of the compound detected in three different parts of the plant extract; leaf, stem, and root, respectively. The edge thickness in the network represents the similarity extent among the connected nodes).

Our study is completely in line with the author's claims. However, more advanced algorithms and timely upgrades for such tools are key to improvement. In the past, this platform has been used generally for annotating the molecular formula only [41,42]. In this instance, we used its embedded CSI: FingerID and CANOPUS tools to predict compound structure from several databases, obligating the Natural Products database [43]. Some of the compounds, such as compounds **9**, **10**, **11**, **13**, **14**, **15**, and **16** have no match with GNPS spectral library database but with the Natural Product database embedded with SIRIUS 4. The rest of the compounds were verified with their GNPS spectrum ID (Table 3). The lack of GNPS spectral matching for these compounds might be due to the misrepresentation of precursor ions. For instance, compound **13** has the precursor ion as  $[M + Na]^+$  with  $m/z$  339.1054, although the GNPS recognized it as  $[M + H]^+$  ion and mislead the spectral match. Compound **4** has also matched to Massbank spectral datasets given to their Massbank access number (Table 3). This further supports the annotation made by GNPS and SIRIUS 4.

In terms of activity, the strong  $\alpha$ -glucosidase inhibition activity of methanolic stem extract can be justified as we annotated vicenin 2, roseoside, and Quercetin 3-

(2G-xylosylrutinoside) in abundance that are already known for their strong antidiabetic properties [29, 30, 44]. However, to thoroughly verify this assumption and these compounds, one needs to isolate them and confirm their structure by spectroscopic analysis. In addition of that, the further isolation, characterization, and bioassay of those seven unknown compounds are worthy to carry out.

## 5. Conclusion

*U. dioica* L. is considered a traditionally important plant in Nepalese culture. It carries medicinal importance and has been consumed at the local level in the form of vegetables as well. It is prioritized due to its antidiabetic activity. Our study has successfully annotated compounds that hold  $\alpha$ -glucosidase activity, which ultimately supports these community-based claims. Moreover, this plant is associated with cosmetic abilities, which were also verified in our study. Further work to isolate new annotated compounds is still necessary. However, our work sets up a vital chemical foundation for this plant.

## Abbreviations

BHA: Butylated hydroxyanisole  
 DPPH: 2,2-diphenyl-1-picrylhydrazyl  
 GNPS: Global natural product social molecular networking  
 HPLC: High-performance liquid chromatography  
 HRMS: High-resolution mass spectrometry  
 NMR: Nuclear magnetic resonance  
 NP: Natural Product.

## Data Availability

We provide supporting and necessary data for the publication of the article. All supporting data are presented in the article and additional files. The sample voucher specimen is deposited in the plant herbarium of Tribhuvan University and can be accessed upon reasonable request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Authors' Contributions

KB and HDB designed the experiments which were conducted by KB, BP, SD, and PV. NA and BB helped in data analysis and drafting the manuscript. All authors read and approved the final manuscript. Keshab Bhattarai and Babita Paudel have equal contribution in the manuscript.

## Acknowledgments

KB and HDB are grateful to Prof. Dr. Harald Gross, Institute of Pharmaceutical Sciences, University of Tübingen, Germany, for providing NMR and HRMS measurement facilities. The work was supported by the IFS Research Grant (F-6243-1) to KB and TUPIO (HERP DLI-7B) research grant to HDB.

## Supplementary Materials

The supplementary files for the manuscript are provided in the supporting information attached herewith. (*Supplementary Materials*)

## References

- [1] M. Asadi-Samani, M.-T. Moradi, L. Mahmoodnia, S. Alaei, F. Asadi-Samani, and T. Luther, "Traditional uses of medicinal plants to prevent and treat diabetes; an updated review of ethnobotanical studies in Iran," *Journal of Nephrothology*, vol. 6, no. 3, pp. 118–125, 2017.
- [2] S. A. Dar, F. A. Ganai, A. R. Yousuf, M.-u.-H. Balkhi, T. M. Bhat, and P. Sharma, "Pharmacological and toxicological evaluation of *Urtica dioica*," *Pharmaceutical Biology*, vol. 51, no. 2, pp. 170–180, 2013.
- [3] J. A. González, M. García-Barriuso, and F. Amich, "Ethnobotanical study of medicinal plants traditionally used in the Arribes del Duero, western Spain," *Journal of Ethnopharmacology*, vol. 131, no. 2, pp. 343–355, 2010.
- [4] P. M. Guarrera and V. Savo, "Perceived health properties of wild and cultivated food plants in local and popular traditions of Italy: a review," *Journal of Ethnopharmacology*, vol. 146, no. 3, pp. 659–680, 2013.
- [5] R. M. Kunwar, L. Mahat, R. P. Acharya, and R. W. Bussmann, "Medicinal plants, traditional medicine, markets and management in far-west Nepal," *Journal of Ethnobiology and Ethnomedicine*, vol. 9, p. 24, 2013.
- [6] D. Kregiel, E. Pawlikowska, and H. Antolak, "*Urtica* spp.: ordinary plants with extraordinary properties," *Molecules*, vol. 23, no. 7, Article ID 1664, 2018.
- [7] B. Farzami, D. Ahmadvand, S. Vardasbi, F. J. Majin, and S. Khaghani, "Induction of insulin secretion by a component of *Urtica dioica* leave extract in perfused Islets of Langerhans and its in vivo effects in normal and streptozotocin diabetic rats," *Journal of Ethnopharmacology*, vol. 89, no. 1, pp. 47–53, 2003.
- [8] G. Kavalalı, H. Tuncel, S. Göksel, and H. H. Hatemi, "Hypoglycemic activity of *Urtica pilulifera* in streptozotocin-diabetic rats," *Journal of Ethnopharmacology*, vol. 84, pp. 241–245, 2003.
- [9] A. Ranjbari, M. A. Azarbayjani, A. Yusof et al., "In vivo and in vitro evaluation of the effects of *Urtica dioica* and swimming activity on diabetic factors and pancreatic beta cells," *BMC Complementary and Alternative Medicine*, vol. 16, no. 1, p. 101, 2016.
- [10] M. Shokrzadeh, A. Mirshafa, N. Yekta Moghaddam, B. Birjandian, and F. Shaki, "Mitochondrial dysfunction contribute to diabetic neurotoxicity induced by streptozotocin in mice: protective effect of *Urtica dioica* and pioglitazone," *Toxicology Mechanisms and Methods*, vol. 28, no. 7, pp. 499–506, 2018.
- [11] R. Kraus and G. Spiteller, "Phenolic compounds from roots of *Urtica dioica*," *Phytochemistry*, vol. 29, no. 5, pp. 1653–1659, 1990.
- [12] M. A. Farag, M. Weigend, F. Luebert, G. Brokamp, and L. A. Wessjohann, "Phytochemical, phylogenetic, and anti-inflammatory evaluation of 43 *Urtica* accessions (stinging nettle) based on UPLC-Q-TOF-MS metabolomic profiles," *Phytochemistry*, vol. 96, pp. 170–183, 2013.
- [13] W. Jian, Y. Hong-Xing, T. Yong-Zhen, Y. Pei, T. Rui, and L. Chun-bi, "An overview on the progress of chemical constituents and bioactivities of plants in Urticaceae during 2000–2010," *Chinese Herbal Medicines*, vol. 3, pp. 9–16, 2011.
- [14] T. Hautbergue, E. L. Jamin, R. Costantino et al., "Combination of Isotope Labeling and molecular networking of tandem mass spectrometry data to reveal 69 unknown metabolites produced by *Penicillium nordicum*," *Analytical Chemistry*, vol. 91, no. 19, pp. 12191–12202, 2019.
- [15] R. A. Quinn, L.-F. Nothias, O. Vining, M. Meehan, E. Esquenazi, and P. C. Dorrestein, "Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy," *Trends in Pharmacological Sciences*, vol. 38, no. 2, pp. 143–154, 2017.
- [16] T. Hautbergue, E. L. Jamin, L. Debrauwer, O. Puel, and I. P. Oswald, "From genomics to metabolomics, moving toward an integrated strategy for the discovery of fungal secondary metabolites," *Natural Product Reports*, vol. 35, no. 2, pp. 147–173, 2018.
- [17] T. P. Cantrell, C. J. Freeman, V. J. Paul, V. Agarwal, and N. Garg, "Mass spectrometry-based integration and expansion of the chemical diversity harbored within a marine sponge," *Journal of the American Society for Mass Spectrometry*, vol. 30, no. 8, pp. 1373–1384, 2019.
- [18] A. Martín-Blázquez, C. Díaz, E. González-Flores et al., "Untargeted LC-HRMS-based metabolomics to identify novel

- biomarkers of metastatic colorectal cancer,” *Scientific Reports*, vol. 9, no. 1, Article ID 20198, 2019.
- [19] F. A. van de Laar, P. L. Lucassen, R. P. Akkermans, E. H. van de Lisdonk, G. E. Rutten, and C. van Weel, “ $\alpha$ -Glucosidase inhibitors for patients with type 2 diabetes,” *Diabetes Care*, vol. 28, no. 1, pp. 154–163, 2005.
- [20] M. S. Blois, “Antioxidant determinations by the use of a stable free radical,” *Nature*, vol. 181, no. 4617, pp. 1199–1200, 1958.
- [21] B. Meyer, N. Ferrigni, J. Putnam, L. Jacobsen, D. Nichols, and J. McLaughlin, “Brine shrimp: a convenient general bioassay for active plant constituents,” *Planta Medica*, vol. 45, no. 5, pp. 31–34, 1982.
- [22] D. Otasek, J. H. Morris, J. Bouças, A. R. Pico, and B. Demchak, “Cytoscape Automation: empowering workflow-based network analysis,” *Genome Biology*, vol. 20, p. 185, 2019.
- [23] K. Jeong, M. Yang, Y. Jin, E. Kim, J. Ko, and J. Lee, “Identification of major flavone C-glycosides and their Optimized extraction from *Cymbidium kanran* using deep eutectic solvents,” *Molecules*, vol. 22, no. 11, Article ID 2006, 2017.
- [24] C. Sanbongi, N. Osakabe, M. Natsume, T. Takizawa, S. Gomi, and T. Osawa, “Antioxidative polyphenols isolated from *Theobroma cacao*,” *Journal of Agricultural and Food Chemistry*, vol. 46, no. 2, pp. 454–457, 1998.
- [25] A.-P. Carnat, A. Carnat, D. Fraisse et al., “Violarvensin, a new flavone di-C-glycoside from *Viola arvensis*,” *Journal of Natural Products*, vol. 61, no. 2, pp. 272–274, 1998.
- [26] D. R. de Oliveira, C. R. Zamberlam, R. B. Gaiardo et al., “Flavones from *Erythrina falcata* are modulators of fear memory,” *BMC Complementary and Alternative Medicine*, vol. 14, no. 1, p. 288, 2014.
- [27] M. N. Samy, H. E. Khalil, S. Sugimoto, K. Matsunami, H. Otsuka, and M. S. Kamel, “Three new flavonoid glycosides, byzantionoside B 6'-O-sulfate and xyloglucoside of (Z)-Hex-3-en-1-ol from *ruellia patula*,” *Chemical and Pharmaceutical Bulletin*, vol. 59, no. 6, pp. 725–729, 2011.
- [28] C. Sarikurkcü, S. S. Sahinler, and B. Tepe, “*Astragalus gynomolobus*, *A. leporinus* var. *hirsutus*, and *A. onobrychis*: phytochemical analysis and biological activity,” *Industrial Crops and Products*, vol. 150, Article ID 112366, 2020.
- [29] D. Patel, S. Prasad, R. Kumar, and S. Hemalatha, “An overview on antidiabetic medicinal plants having insulin mimetic property,” *Asian Pacific Journal of Tropical Biomedicine*, vol. 2, no. 4, pp. 320–330, 2012.
- [30] M. R. Moloto, A. D. T. Phan, J. L. Shai, Y. Sultanbawa, and D. Sivakumar, “Comparison of phenolic compounds, carotenoids, amino acid composition, in vitro antioxidant and antidiabetic activities in the leaves of seven cowpea (*Vigna unguiculata*) Cultivars,” *Foods*, vol. 9, no. 9, Article ID 1285, 2020.
- [31] A. S. Noguerola, B. Murugaverl, and K. J. Voorhees, “An investigation of dipeptides containing polar and nonpolar side groups by curie-point pyrolysis tandem mass spectrometry,” *Journal of the American Society for Mass Spectrometry*, vol. 3, no. 7, pp. 750–756, 1992.
- [32] M. J. Abad, P. Bermejo, M. Alvarez, J. A. Guerra, A. M. Silván, and M. V. Angel, “Flavonoids and a sesquiterpene lactone from *Tanacetum microphyllum* inhibit anti-inflammatory mediators in LPS-stimulated mouse peritoneal macrophages,” *Planta Medica*, vol. 70, no. 1, pp. 34–38, 2004.
- [33] T. Hôrie, Y. Ohtsuru, K. Shibata, K. Yamashita, M. Tsukayama, and Y. Kawamura, “<sup>13</sup>C NMR spectral assignment of the A-ring of polyoxygenated flavones,” *Phytochemistry*, vol. 47, pp. 865–874, 1998.
- [34] G. Wald, “Carotenoids and the vitamin A cycle in vision,” *Nature*, vol. 134, no. 3376, p. 65, 1934.
- [35] H. Ferraz, L. Longo Jr., and F. Bombonato Jr., “Synthetic approaches to naturally occurring ten-membered-ring lactones,” *Synthesis*, vol. 2007, no. 21, pp. 3261–3285, 2007.
- [36] D. Gülcemal, Ö. Alankuş-Çalışkan, C. Karaalp, A. U. Örs, P. Ballar, and E. Bedir, “Phenolic Glycosides with anti-proteasomal activity from *Centaurea urvillei* DC. subsp. *urvillei*,” *Carbohydrate Research*, vol. 345, no. 17, pp. 2529–2533, 2010.
- [37] A. Keiler, O. Zierau, and G. Kretzschmar, “Hop extracts and hop substances in treatment of menopausal complaints,” *Planta Medica*, vol. 79, no. 7, pp. 576–579, 2013.
- [38] Y.-J. Surh, S. H. Ahn, K.-C. Kim, J.-B. Park, Y. W. Sohn, and S. S. Lee, “Metabolism of capsaicinoids: evidence for aliphatic hydroxylation and its pharmacological implications,” *Life Sciences*, vol. 56, no. 16, pp. PL305–PL311, 1995.
- [39] C.-B. Cui, Y. Tezuka, T. Kikuchi, H. Nakano, and J.-H. Park, “Davallioside A and B, novel flavan-3-ol derivatives with a  $\gamma$ -lactam, from the rhizomes of *Davallia mariesii* Moore,” *Chemical and Pharmaceutical Bulletin*, vol. 38, no. 9, pp. 2620–2622, 1990.
- [40] I. Saracoglu, M. Varel, U. S. Harput, and A. Nagatsu, “Acylated flavonoids and phenol glycosides from *Veronica thymoides* subsp. *pseudocinerea*,” *Phytochemistry*, vol. 65, no. 16, pp. 2379–2385, 2004.
- [41] M. Ludwig, M. Fleischauer, K. Dührkop, M. A. Hoffmann, and S. Böcker, “De novo molecular formula annotation and structure elucidation using SIRIUS 4,” *Methods in Molecular Biology*, vol. 2104, pp. 185–207, 2020.
- [42] K. Dührkop, K. Scheubert, and S. Böcker, “Molecular formula identification with SIRIUS,” *Metabolites*, vol. 3, no. 2, pp. 506–516, 2013.
- [43] K. Dührkop, L.-F. Nothias, M. Fleischauer et al., “Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra,” *Nature Biotechnology*, vol. 39, no. 4, pp. 462–471, 2021.
- [44] M. N. Islam, I. J. Ishita, H. A. Jung, and J. S. Choi, “Vicenin 2 isolated from *Artemisia capillaris* exhibited potent anti-glycation properties,” *Food and Chemical Toxicology*, vol. 69, pp. 55–62, 2014.