Molecular Omics

ROYAL SOCIETY OF CHEMISTRY

RESEARCH ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Mol. Omics*, 2025, **21**, 747

Mapping the *Helichrysum* metabolome: uncovering species-specific chemistry through an AI-guided LC-MS/MS workflow

Motseoa Mariam Lephatsi,[a] Mpho Susan Choene,[a] Abidemi Paul Kappo,[a] Ntakadzeni Edwin Madala[b] and Fidele Tugizimana *[a]

Helichrysum species, of which 35% are native to South Africa, are renowned for their diverse medicinal properties, yet their chemical composition remains largely unexplored. As such, continuous efforts are needed to comprehensively characterize the phytochemistry of *Helichrysum* species which will subsequently contribute to the discovery and exploration of *Helichrysum*-derived natural products for drug discovery. Thus, a computational metabolomics work is reported herein to comprehensively characterize the metabolic landscape of three medicinal species (*H. italicum*, *H. petiolare*, and *H. splendidum*), which are less studied. The metabolites were extracted using hexane, ethyl acetate, and methanol and analyzed on a liquid chromatography-tandem mass spectrometry (LC-MS/MS) system. Different solvents were utilized to increase metabolome coverage in *Helichrysum* species. Spectral data were mined using molecular networking (MN) strategies. The results revealed that multiple extraction methods provide a more comprehensive analysis of the metabolome of the three plants. The measured metabolome of *Helichrysum* species is rich in phenylpropanoids, lipids and lipid-like molecules, pointing to a rich chemistry with potential bioactivities. Comparative analysis of the *H. italicum*, *H. petiolare* and *H. splendidum* metabolomes revealed that the flavonoid glucoside and triterpenoid profiles of the three species differ distinctively. These results expand the knowledge base on the chemistry of *Helichrysum* plants and provide deconvoluted details of the various chemical classes that differentially define the metabolome of the *Helichrysum* plants. Such actionable insights point to *Helichrysum*'s potential as a valuable source of natural compounds with promising medicinal properties.

Received 2nd June 2025,
Accepted 23rd September 2025

DOI: 10.1039/d5mo00118h

rsc.li/molomics

1. Introduction

Helichrysum plants have traditionally been used in various societies across the globe, being highly popular for their medicinal, cosmetic, and fragrant qualities. Traditional medicine has frequently relied on *Helichrysum* species to address health issues such as respiratory disorders, skin conditions, digestive problems, and inflammation.[1] Modern scientific research has directed its focus toward the *Helichrysum* genus to uncover the unexplored chemical diversity within the *Helichrysum* species. The *Helichrysum* genus has been reported to be highly rich in phenolic compounds, and the latter have been attributed as the bioactive constituents[2–4] of this genus. Previous phytochemical investigations of *Helichrysum* species have identified diverse classes of bioactive metabolites, including flavonoids, terpenoids, and phenolic acids. For instance, chlorogenic acids have been consistently reported in *H. arenarium* and *H. italicum*, showing antibacterial and anti-inflammatory properties.[5,6] However, this is a tip of an iceberg, considering what could be the diversity and the size of the chemical constellations and space of *Helichrysum* plants. There are gaps that still exist in the understanding and description of the metabolome of *Helichrysum* plants. This 'dark matter' of the *Helichrysum* metabolome therefore limits the use and exploration of this plant. Furthermore, characterizing metabolic charts of the lesser studied *Helichrysum* species will in turn uncover the untapped potential of the plant, discovering new bioactivities and pharmaceutical applications.

Furthermore, this knowledge gap in the chemistry of *Helichrysum* plants is related to the metabolome coverage with reference to metabolite annotation and identification, which remains a major bottleneck in untargeted metabolomics and in phytochemistry studies. Mass spectrometry (MS) has become the most dominant analytical platform for obtaining structural information of complex mixtures. Despite the advancements in

[a] *Research Centre for Plant Metabolomics, Department of Biochemistry, University of Johannesburg, Auckland Park, Johannesburg 2006, South Africa. E-mail: motseoal@uj.ac.za, ftugizimana@uj.ac.za*
[b] *Department of Biochemistry and Microbiology, University of Venda, Thohoyandou 0950, South Africa. E-mail: Ntakadzeni.Madala@univen.ac.za*

MS technologies, only about 1.8% of spectra in an untargeted metabolomics experiment can be annotated.[7] This implies that the vast majority of information collected by metabolomics is ''dark matter'', chemical signatures that remain uncharacterized.[7,8] Considering that accurate annotation is vital for data interpretation, the only way to overcome this challenge is through the development and integration of artificial intelligence (AI)-driven strategies, computational solutions and methods.[9,10] These strategies include molecular networking, MS2LDA, NAP and SIRIUS which were proven to enhance metabolite annotation, thus elucidating the molecular complexity of *H. splendidum* and providing insights into the chemical space of this plant. Molecular networking, which is housed on the Global Natural Products Social Molecular Networking (GNPS) ecosystem, is a computational strategy that takes advantage of the idea that MS/MS can be used as a proxy for molecular structures.[9,11]

This study therefore involves the use of these computational methods to comprehensively investigate the metabolomes of three *Helichrysum* species namely, *H. petiolare*, *H. italicum*, and *H. splendidum*. Furthermore, pre-analytical parameters, such as metabolite extraction methods, are explored to expand the metabolome coverage.[12,13] Thus, herein we report a comprehensive metabolomic atlas, the first of its kind, of the *Helichrysum* plants. As such, this AI-driven metabolomics study provides actionable insights into the metabolomic landscape of *Helichrysum*, expanding the current knowledge base on the phytochemistry of the plant, and highlighting key metabolite classes and subclasses that define the *Helichrysum* chemical space. Such metabolite profiles are essential drivers of the biological and pharmaceutical activities attributed to the extracts from the *Helichrysum* plants. It is worth noting that while this AI-guided LC-MS/MS workflow was optimized using *Helichrysum* species, the pipeline itself is not taxonomically restricted. The combination and integration of these emerging AI and computational strategies can be applied and explored in mining and interpreting metabolomics spectral data from other plants.

2. Materials and methods

2.1. Chemicals and plant materials

All chemicals used in this study were of pure grade quality and were acquired from various manufacturers. The organic solvent, methanol, was of LC-MS-grade quality and was obtained from Romil (Cambridge, UK), water was purified using a Milli-Q gradient A10 system (Siemens, Munich, Germany) and formic acid was purchased from Sigma Aldrich (Munich, Germany). The different *Helichrysum* seeds were purchased from Seeds for Africa (**https://www.seedsforafrica.co.za**, accessed on 17 March 2023) and the plants were cultivated separately by species to avoid cross-contamination or environmental bias in 4 L pots filled with potting soil mixed with Vita-Veg organic fertilizer (Talborne Organics, Bronkhorstspruit, South Africa). For each of the three *Helichrysum* species (*H. italicum*, *H. petiolare*, and *H. splendidum*), 24 independent plants (3 plants per pot and 8 pots in total) were prepared to ensure reproducibility and were placed under natural light. Each pot (containing 3 plants) was considered as one biological replicate; this means that for each *Helichrysum* species, there were 8 biological replicates (*i.e.*, 8 pots). The plants were harvested at a 4-month growth stage. The stems and leaves of the three *Helichrysum* species (*H. italicum*, *H. petiolare* and *H. splendidum*) were freeze-dried and crushed, and the powdered samples were stored in dried form at room temperature for metabolite extractions.

2.2. Extraction of metabolites from plant materials and sample preparation

One gram (1 g) of the powdered plant material, per biological replicate, was weighed and dissolved in 20 mL of 80% methanol at 4 °C. The mixture was spun overnight in a digital rotisserie tube rotator at 70 rpm. The crude extracts were then centrifuged at 2000 rpm for 30 min at 4 °C, and the resulting supernatants were filtered through a 0.22 μm nylon filter into pre-labeled glass vials with 500 μL inserts. The filtered samples were then stored at 4 °C until further analysis. Thus, twenty-four independent biological replicates (*i.e.*, 8 biological replicates per *Helichrysum* species) were prepared and each biological replicate was analyzed three times (technical replicates) to account for instrumental variability. Furthermore, pooled quality control (QC) samples, prepared by combining equal aliquots from all study samples, were injected at both the start and end of the batch to ensure data validation and system stabilization. Additionally, QC samples were analyzed after every 15 injections to monitor and address any changes in the instrument's response.

2.3. Sample analysis using liquid chromatography-quadrupole time-of-flight tandem mass spectrometry (LC-MS/MS)

The prepared *Helichrysum* extracts were then analyzed on a liquid chromatography-quadrupole time-of-flight tandem MS system (LCMS-9030 qTOF, Shimadzu Corporation, Japan). The chromatographic separation was done using a Shim-pack Velox C18 column (100 mm × 2.1 mm, 2.7 μm) (Shimadzu Corporation, Kyoto, Japan), thermostatted at 50 °C, with an injection volume of 3 μL. A binary solvent system consisting of solvent A: 0.1% formic acid (Merck, Darmstadt, Germany) in Milli-Q water and solvent B: methanol (Romil Pure Chemistry, Cambridge, UK) with 0.1% formic acid was used as the mobile phase with a flow rate of 0.4 mL min^{-1} to separate analytes over an optimized run length. The gradient was set over 53 min with the following separation conditions: 10% B maintained for 3 min, 10–60% B over 3–40 min, from 40–43 min the conditions were maintained at 60% B, then the gradient was changed to 90% B between 43 and 45 min and maintained at 90% B for 3 min. The gradient returned to initial conditions between 48 and 50 min, which was followed by a 3-min column equilibration time.

The chromatographic effluents were further analyzed utilizing the qTOF high-definition (HD) mass spectrometry (MS) analytical platform. Negative electrospray ionization (ESI) was used based on preliminary experiments (unpublished) that

showed optimal ion peaks detected across the chromatographic (polarity) range. The following parameters were set: an interface voltage of 4.0 kV, an interface temperature of 300 °C, nebulization and dry gas flow rate of 3 L min^{-1}, a heat block temperature of 400 °C, a DL temperature of 280 °C, a detector voltage of 1.8 kV and a flight tube temperature of 42 °C. Sodium iodide (NaI) was used as a calibration solution to monitor high mass accuracy. MS1 and MS2 (through data-dependent acquisition, DDA) were generated simultaneously for all ions with an *m/z* range between 100 and 1000 Da, surpassing an intensity threshold of 5000 counts. Fragmentation experiments were performed using argon as a collision gas at a collision energy of 30 eV with a spread of 5 eV. To ensure the acquisition of quality spectral data, solvent blanks (50% aqueous methanol) and the pooled quality control (QC) samples were also analyzed in parallel with the sample extracts. The blank samples were used to monitor background noise, whereas the QC samples were used to condition the LC–MS system and to assess the reliability and reproducibility of the analysis. The samples (extracts, blanks and QC) were analyzed in a randomized manner. The LabSolution CS software (Shimadzu Corporation, Kyoto, Japan) was used to operate the LC-MS system.

2.4. Data mining: molecular networking in the GNPS analysis environment

The raw data obtained from the Shimadzu LCMS-9030 qTOF were converted to an open-source format (.mzML) and following this, molecular network (MN) methods were applied to mine the spectral data, particularly the feature-based molecular networking (FBMN) workflow[14] on GNPS (**https://gnps.ucsd.edu**).[15] The mass spectrometry data were first processed with MS-DIAL version 4.9.221218.[16] The parameters used for MS-DIAL data processing included mass accuracy MS1 and MS2 tolerance of 0.05 Da and 0.1 Da respectively, with the MS/MS range of 50–1000 Da; the minimum peak height was 2000 amplitude, the mass slice width was 0.1 Da for peak detection, a Sigma window value of 0.5 was used, and the retention time tolerance was 0.1 min and the MS1 tolerance was set at 0.05. The results were then exported to GNPS for FBMN analysis. The precursor ion mass tolerance was set to 0.05 Da and the MS/MS fragment ion tolerance to 0.05 Da. A molecular network was then created where edges were filtered to have a cosine score above 0.6 and more than 4 matched peaks. Furthermore, edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 100, and the lowest-scoring edges were removed from molecular families until the molecular family size was below this threshold.

The spectra in the network were then searched against GNPS spectral libraries such as GNPS, HMDB, SUPNAT, CHEBI, DRUGBANK and FooDB. All matches maintained between network spectra and library spectra were required to have a score above 0.6 and at least 4 matched peaks. The DEREPLICATOR was used to annotate MS/MS spectra.[17] The molecular networks were then visualized using Cytoscape software version 9.1.[18] All matched and some unmatched nodes were verified or putatively annotated using their empirical formulas generated from accurate mass and fragmentation patterns obtained from MS2 experiments. These were also searched against some common dereplication databases for natural products, such as KNApSAck,[19,20] ChemSpider,[21] PubChem,[22,23] Dictionary of Natural Products[24] and available literature. The molecular network was also explored using DEREPLICATOR for structural annotation. Substructure exploration and annotation were performed using MS2LDA,[25] interface in GNPS, from MotifDB, with MotifSets: Rhamnaceae and GNPS Mass2Motifs included in the search. Default MS2LDA parameters were used, locally providing sufficient insights into the presence of substructure patterns in the mass spectral data. To enhance chemical structural information within the generated molecular network (FBMN), outputs from *in silico* structure annotations from the GNPS Library Search, Network Annotation Propagation (NAP), DEREPLICATOR, and MS2LDA were incorporated into the network using the GNPS MolNetEnhancer workflow.[26] All ions were set to be deprotonated, and the consensus and fusion scores were calculated based on the first 10 candidates. Chemical class annotations were performed using the ClassyFire chemical ontology.[27] The GNPS job links are provided in the SI.

2.5. Relative quantification and pathway analysis

Relative quantification of the identified metabolites was performed by generating heatmaps in MetaboAnalyst version 5. Firstly, the peak intensity table was uploaded onto MetaboAnalyst, and the data were then normalized by median to adjust any systematic differences among the samples. The data were then log-transformed and Pareto-scaled. Following this, heatmaps were generated using the Euclidean distance measure and the Ward clustering method. To focus on patterns from the important features and to visualize the abundance of the different metabolites across the three *Helichrysum* species, only the top 25 features based on the *t*-test/ANOVA were shown. Pathway analysis was performed by inputting the annotated metabolites together with their KEGG IDs onto the MetPA (metabolic pathway analysis) tool, also housed on MetaboAnalyst. The hypergeometric test was used as the enrichment method for the overview-representation analysis, relative-betweenness centrality was chosen for topological analysis, scatterplots were used for the visualization and *Arabidopsis thaliana* (KEGG) was chosen as the pathway library.

3. Results and discussion

3.1. Exploring pre-analytical steps: impact of different extraction methods on metabolome coverage

As the adage goes, "*metabolomics: what you see is what you extract*",[28] there is no single extraction method that can recover all classes of metabolites with high reproducibility and robustness. All extraction methods, individually, provide snapshots of the structurally diverse and highly complex metabolome of any biological system under consideration.[29] As such, to expand the

coverage of the metabolome under consideration, a combination of different extraction methods is devised. Thus, as mentioned above, in this study, we explored three extraction solvents, *i.e.*, methanol, hexane, and ethyl acetate, to comprehensively characterize the metabolome of the three *Helichrysum* plant species. Mining and visualizing the spectral data with molecular networking methods revealed that the three solvents were used to extract different numbers of features (Fig. 1). Methanol showed the highest metabolome coverage across all three plant species with a total of 8607 features (Fig. 1A). Methanol is a commonly used solvent in metabolomics studies due to its ability to extract a diverse range of metabolites. Its polarity allows for the extraction of both polar and non-polar compounds.[30,31] Hexane and ethyl acetate, on the other hand, were used to extract 726 and 2476 total features, respectively (Fig. 1B and C). Thus, the total number of features extracted using the three solvents was 11 809. Hexane is a non-polar solvent used to extract compounds that are not polar, and ethyl acetate, on the other hand, has moderate polarity. The use of the three solvents led to the extraction of different metabolites, which contributed to expanding the coverage of the metabolome, thus illuminating the dark matter of the *Helichrysum* phytochemical space.

To comprehensively delve into the extent to which these three solvents enhance the coverage of the metabolome and thus describe the 'molverse' of *Helichrysum*, different chemical classes extracted using each solvent were elucidated within the generated FBMN (Fig. 1). The NPClassifier tool was used to elucidate and visualize the different chemical classes extracted using the different solvents. NPClassifier is a deep-learning tool used to classify natural products (NPs) based on their spectral data by generating classification information at three levels, including pathway, superclass, and class.[32] This classification information, therefore, reveals the measured chemical diversity by each solvent, allowing for the comprehensive metabolome coverage of *Helichrysum* species. The results showed that more classes of metabolites were extracted with these three extraction methods. Methanol was used to extract a wide range of different metabolite classes including flavonols, flavones, oleanane triterpenoids, cinnamic acids, coumarins, fatty acyl glycosides and cholane steroids (Fig. 1A). Hexane was used to extract classes including sesquiterpenoids, coumarins, flavonols and pregnane steroids (Fig. 1B). Ethyl acetate, on the other hand, was used to extract metabolite classes including phenolic acids, cinnamic acids, monoterpenoids, sesquiterpenoids and coumarins (Fig. 1C). These findings highlight that multiple extraction methods are essential for a more comprehensive analysis of the metabolite profile of a given sample. Leveraging a combination of different extraction solvents therefore expands the coverage of the metabolomic space by capturing a broader range of metabolite classes present in the *Helichrysum* genus.

By using a combination of different solvent extractions such as methanol, hexane and ethyl acetate extractions, a wider range of metabolite classes is captured, elucidating the chemical diversity present in the biological sample. Some overlaps were also observed in the three solvents. All solvents extracted coumarins and terpenoids (Fig. 1D). Interestingly, hexane contributed an additional layer of complexity that could not be deciphered using either methanol or ethyl acetate (Fig. 1D). A class of compounds unique to this solvent was identified as pregnane steroids, suggesting that hexane was used to extract a distinct set of metabolites that are not captured by methanol extraction (Fig. 1B and D). Pregnane steroids are a class of steroids derived from pregnane and they are characterized by a complex vinyl side chain.[33] This class of compounds has been reported to exhibit numerous potent bioactivities such as anti-bacterial, anti-protozoal and cytotoxic activities and has received much attention due to these multiple activities.[33–36] More recently, Song *et al.* (2023)[37] reported on eighteen pregnane steroid glycosides isolated from *Marsdenia tenacissima*. These glycosides were evaluated for their chemo-reversal ability against P-glycoprotein (P-gp)-mediated multidrug resistance (MDR) in the MCF-7/ADR cell line, in which the reversal folds ranged from 2.45–9.01.[37] These compounds therefore possess chemo-reversal ability against multidrug resistance in cancer cells and therefore have potential as therapeutic agents for cancer treatment. The extraction of this class of compound using hexane from *Helichrysum* plants therefore highlights the importance of exploring more than one solvent for an enhanced metabolome coverage.

Hexane and ethyl acetate on the other hand were used to extract different metabolites, suggesting that using different solvents for metabolite extraction expands the metabolome coverage of *Helichrysum* plants (Fig. 2). For example, zooming into the different classes, such as phenylpropanoids and terpenoids, which were both extracted with the two solvents, hexane

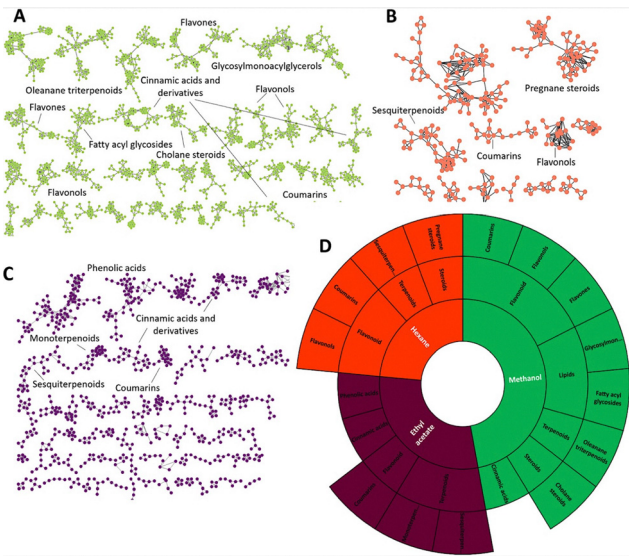

Fig. 1 Class chemical space of *Helichrysum* species using different solvents, highlighting the expansion of metabolome coverage. The three solvents were used to extract different metabolite classes, showing the expansion of the metabolome with some overlaps also observed. (A) Methanol extraction, (B) hexane extraction, (C) ethyl acetate extraction and (D) sunburst plot summarising the different classes covered by each solvent.

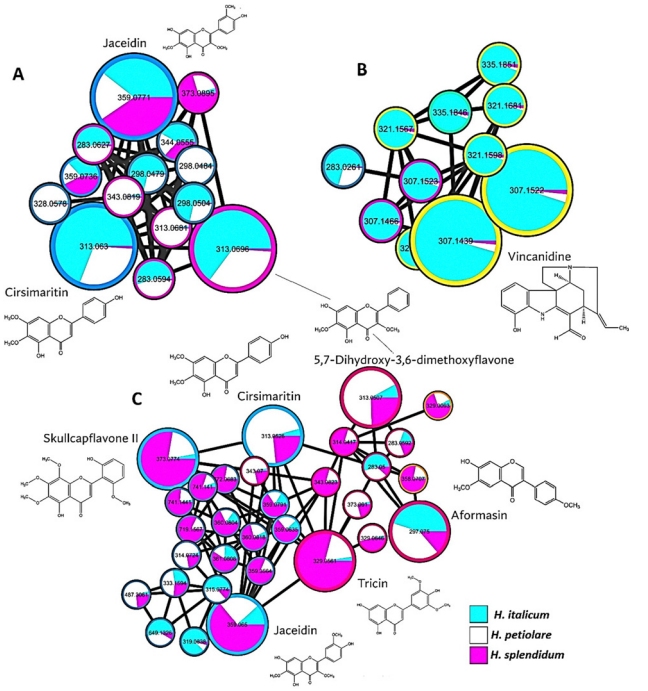

Fig. 2 Feature-based molecular networks showing identified classes and metabolites from different solvents. The hexane extraction and ethyl acetate extraction yielded distinct classes of metabolites, highlighting the specificity of each solvent. The FBMNs provide a visual representation of the identified metabolites. (A) and (B) Hexane extraction and (C) ethyl acetate extraction together with the annotated metabolites.

was used to extract compounds such as vincanidine (Fig. 2A and B). In contrast, ethyl acetate solvent allowed for the extraction of skullcapflavone II and aformasin (Fig. 2C). Some overlaps were also identified between the three solvents. Common to the three extraction methods were jaceidin, cirsimaritin and 5′,7-dihydroxy-3,6-dimethoxyflavone which all belong to the methylated flavonoid class. Methylated flavonoids have been reported to possess numerous bioactivities such as antioxidant activity, anti-inflammatory effects, anticancer potential, neuroprotective effects and anti-microbial activity.[38,39] Jaceidin, for example, has been reported to exhibit *in vitro* cytotoxicity and *in vivo* anti-tumor effect of Ehrlich's ascites carcinoma in mice.[40] It is worth noting that the overlapping metabolites were not extracted at the same levels. For example, cirsimaritin was present in low levels in *H. petiolare* when hexane was used as a solvent (Fig. 2A). The opposite was however observed in the case of ethyl acetate, where *H. petiolare* was the species that had the highest level of cirsimaritin (Fig. 2C). Similar to the findings of this study, Nawaz *et al.* (2020)[41] reported on the effect of solvent polarity on the extraction yield from *Phaseolus vulgaris*, in which highly polar solvents resulted in a high extract yield as compared to non-polar ones. Additionally, the differences in all the three species (*H. italicum*, *H. petiolare*, and *H. splendidum*) also highlight that a metabolite profile is influenced by the type of solvent used during the extraction process.

The intriguing interplay between the solvents underscores the intricate balance required in metabolomics research, where each solvent choice serves as a strategic tool, unveiling specific facets of the *Helichrysum* metabolome. The use of three different solvents in extracting the metabolites from *Helichrysum* plants allowed for the expansion of the metabolome coverage (Fig. 1). With different solvents, it was possible to see additional classes of metabolites which was not possible by using a single solvent for extraction. Exploring various solvents has uncovered the existence of secondary metabolites in the *Helichrysum* genus that have not been reported before, such as vincanidine, cirsimaritin, skullcapflavone II and aformasin, offering significant knowledge about the plant's chemical diversity. As such, these findings reaffirm the importance of exploring multiple extraction solvent systems to ensure comprehensive metabolome coverage in plant metabolomics studies, since no single extraction method can indeed extract or cover the entire metabolome (due to the inherent complexity of the metabolome).[12,42,43]

MolNetEnhancer was used to elucidate the metabolome coverage of the *Helichrysum* plants, and only methanol extracts were used for an epistemological clarity of discussions. MolNetEnhancer reveals molecular families, subfamilies and structural nuances between family members and provides a more comprehensive metabolite assignment at various molecular levels (from wide chemical classes to structurally diverse scaffolds and candidate structures).[26] As such, MolNetEnhancer provides a comprehensive overview of the chemical space present in MS experiments. Thus, the use of MolNetEnhancer to visualize the overall chemical space of the three species of *Helichrysum* (Fig. 3) revealed the presence of distinct chemical superclasses within their metabolomic landscapes. As infographically depicted (Fig. 3), the measured metabolome of *Helichrysum* species is rich in phenylpropanoids, lipids and lipid-like molecules. Functionally, phenylpropanoids and polyketides (Fig. 3) are oxy-prenylated secondary metabolites that

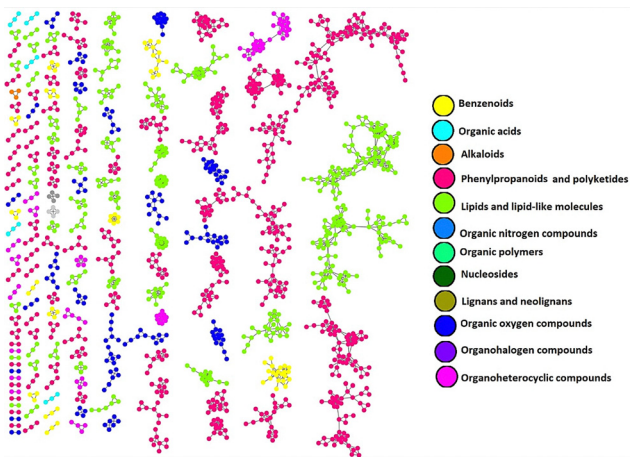

Fig. 3 *Helichrysum* chemical space visualization with molecular networking. MolNetEnhancer network analysis of spectral data from *Helichrysum* methanol extracts. The network shows different superclasses identified, which define the *Helichrysum* chemical space: lipid and lipid-like molecules, phenylpropanoids and organic oxygen-containing compounds. The coloured nodes represent the MS/MS spectra matched to GNPS libraries.

represent a rare group of natural products. Due to their *in vitro* and *in vivo* pharmacological activities, as well as their great therapeutic and nutraceutical potential for the chemoprevention of acute and chronic diseases affecting humans, this group of phytochemicals has become a topic of intense research activity by several teams worldwide over the last two decades.[44] Such studies have revealed that oxy-prenylated secondary metabolites can interact with a range of biological targets at various levels, accounting for their anti-carcinogenic, anti-inflammatory, neuroprotective, immuno-modulatory, anti-hypertensive, and metabolic properties.[45–47]

Lipids on the other hand serve a range of biological roles in plant cells, both structurally and as bioactive compounds. For instance, phospholipids and sphingolipids are cell membrane components that participate in cell signalling; galactolipids are chloroplast membrane components that participate in photosynthesis; and triacylglycerols (TAGs) are used for energy storage.[48,49] Beyond their fundamental cellular functions, some of these lipid molecules have been reported to provide specific health benefits to humans. Depending on their modes of action, plant-derived lipids can stimulate the human immune system, reduce inflammation, improve bone health, eye, and brain function, minimize coronary heart disease, and act as antioxidants and anti-carcinogens.[50–52] Having, thus, uncovered the structurally diverse and complex metabolome of *Helichrysum*, it is possible that there are nuances between *Helichrysum* species. Post MolNetEnhancer analysis for metabolome coverage, the metabolomes of the three *Helichrysum* species were then compared to fish out any differences among the species. The presence of any differences in the metabolic makeup may suggest the existence of unique pathways and compounds in these *Helichrysum* species that perhaps have contributed to the different biological properties of the individual species. This study is therefore of immense significance as the comparison of the three metabolomes could unveil novel bioactive compounds and their application in natural product research or perhaps provide insights for targeted investigations.

3.2. Comparative analysis of the metabolomes of the three *Helichrysum* species: *H. italicum*, *H. petiolare* and *H. splendidum*

Methanol extracts occupy a large fraction of the metabolome compared to other solvents (Fig. 1) and therefore only the methanol extract data were used to investigate the nuances in the metabolomes of the three *Helichrysum* species. Thus, chemometrically, multivariate methods were applied to descriptively highlight trends and groupings within the dataset from methanol extracts, to understand the relationships between and within the samples. The principal component analysis (PCA) model revealed plant species-related sample groupings in the score space (Fig. S1A). Such sample groupings point to the underlying differential metabolic profiles of the three *Helichrysum* plants, *i.e.*, *H. italicum*, *H. splendidum* and *H. petiolare*. Furthermore, the hierarchical cluster analysis (HCA) was applied to further examine the PCA-extracted trends in the data. HCA helps in evaluating natural groupings or distinct subspaces in the metabolite space.[53,54] The computed HCA model revealed that *H. petiolare* and *H. splendidum* are clustered closer compared to *H. italicum* on the dendrogram (Fig. S1B). This points to nuances in the metabolite profiles of the three *Helichrysum* species, with *H. petiolare* and *H. splendidum* being more similar in comparison with the metabolic profiles of *H. italicum*.

To articulate these differential metabolomic landscapes of the three *Helichrysum* species (Fig. S1), spectral data were annotated using molecular networking (MN) strategies. The LC-MS/MS data of methanolic extracts of *H. splendidum*, *H. italicum* and *H. petiolare* were submitted to the GNPS web platform and classical molecular networking was employed. The resulting molecular network consisted of 2971 mass spectral nodes organized into 118 molecular families (two or more connected nodes of a graph) through spectral similarities (Fig. 4). By grouping metabolites into clusters based on their fragmentation spectra, MN provides a comprehensive understanding of the chemical diversity present in *Helichrysum* species. From the putatively annotated metabolites (Table S1), most were identified as flavonoids such as quercetin-3-*O*-rutinoside, quercetin-3-*O*-glucoside and quercetin-3-*O*-glucose-6-acetate (Fig. 4) which were identified in all the three species. Flavonoids are secondary metabolites that are abundant in plants and serve several functions such as regulating cell growth, attracting pollinators, and defending against (a)biotic stresses.[55] For example, flavonoids can serve as signaling molecules and shield against the damaging effects of UV radiation while also aiding in physiological functions associated with drought, cold and heat tolerance.[56,57] Additionally, due to their bioactive properties, flavonoids have been linked to human health benefits. Flavonoids possess anti-inflammatory, anti-cancer, anti-diabetic, anti-bacterial, anti-parasitic and anti-viral properties.[58,59]

The three *Helichrysum* species are also highly rich in chlorogenic acids (CGAs) such as 3,4-dicaffeoylquinic acid and 3-*O*-coumaroylquinic acid (Fig. 4). Chlorogenic acids belong to the polyphenol family and are esters of quinic acid (QA) and one trans-cinnamic acid residue such as caffeic acid (CA), *p*-coumaric

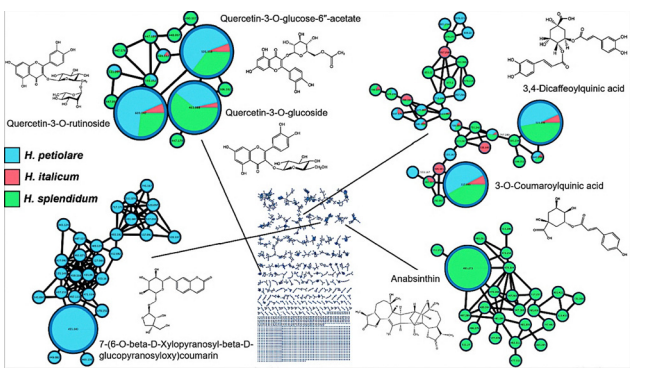

Fig. 4 Classical molecular network of *Helichrysum*. The network highlights flavonoids and chlorogenic acids and their abundance in each species. Different flavonoids and chlorogenic acids are differentially distributed in the three *Helichrysum* species, as depicted in the pie chart in the spectral nodes.

acid (*p*-CoA), and ferulic acid (FA), which are known as caffeoylquinic acids (CQAs), *p*-coumaroylquinic acids (*p*-CoQAs) and feruloylquinic acids (FQAs).[60,61] These compounds have drawn significant attention in natural product research due to their potential health benefits such as anti-inflammatory, antibacterial, anti-viral, anti-parasitic and anti-cancer properties.[62–65] Numerous studies have reported on the presence of chlorogenic acid-type compounds in *Helichrysum* species which serve as anti-HIV activity biomarkers.[66–68] Gradinaru *et al.* (2014)[69] also reported on the anti-bacterial activity of methanol extracts from *Helichrysum arenarium* (L.) Moench subsp. inflorescences against lower respiratory tract pathogens. The authors reported that several caffeic acid conjugates (chlorogenic acid and dicaffeoylquinic acids) which were identified as major constituents of the extract exhibited anti-bacterial activity against methicillin-resistant *Staphylococcus aureus*, penicillin-resistant *Streptococcus pneumoniae* and ampicillin-resistant *Moraxella catarrhalis* isolates.

In addition to the shared metabolic profiles, such as the presence of flavonoids and CGAs in the three *Helichrysum* species, some compounds were specific to a particular species. For instance, 7-(6-*O*-beta-D-xylopyranosyl-beta-D-glucopyranosyloxy) coumarin (a coumarin derivative) and its structurally related compounds were present in *H. petiolare* and absent in the two other species (Fig. 4). Coumarin derivatives are a group of compounds that have garnered attention in the field of natural product research. These compounds, characterized by their nature and fusion of a benzene ring with a pyrone ring, offer a range of biological activities and potential therapeutic applications.[70,71] Coumarins have captured the interest of researchers who have discovered derivatives with properties. This has prompted exploration within the realm of natural product research to uncover their roles and functions. Coumarin derivatives have been reported to exhibit a wide range of biological activities, including anti-inflammatory, antioxidant, anti-coagulant, anti-viral, anti-cancer, and anti-microbial properties.[72–74] These activities make them attractive candidates for natural product-based therapeutics and therefore *H. petiolare* may serve as a candidate for coumarin derivatives with diverse bioactivities.

Interestingly, *H. splendidum* also displayed a distinctive metabolomic profile, setting it apart from the other two species. Anabsinthin, a sesterterpenoid, and its structurally related compounds were identified in *H. splendidum* extracts and were absent in *H. petiolare* and *H. italicum* extracts (Fig. 4). It is worth noting that the presence of these sesterterpenoids only unique to *H. splendidum* highlights the potential for chemical diversity in *Helichrysum* plant species. Sesterterpenoids are a group of terpenes known for their distinctive chemical structures.[75] These compounds exhibit a range of activities, such as suppression of cancer cell growth, inhibition of enzyme activity, modulation of receptor signalling and anti-microbial effects.[76,77] Extensive research has been dedicated to isolating, identifying, and characterizing these compounds leading to the discovery of new sesterterpenoid structures as they are attractive targets in natural product research.[78] Given their structures and bioactivity, this class of compounds are highly appealing targets,

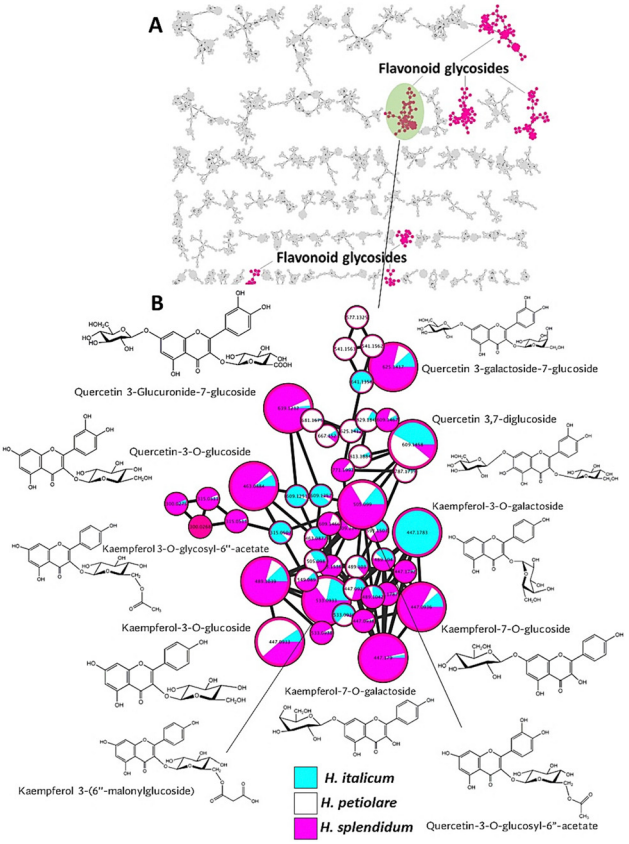

Fig. 5 Molecular network highlighting different flavonoid glycosides in *Helichrysum*. (A) A cluster of flavonoids characterized by a molecular network showing different flavonoid glycosides occupying a large portion of the plants' chemical space. (B) Kaempferol and quercetin are the main backbones of this cluster, with various modifications and substitutions giving rise to different flavonoid glycosides.

and therefore *H. splendidum* may serve as a source for sesterterpenoids.

To further interrogate the *Helichrysum* metabolic landscape, the different subclasses present in the phenylpropanoid and polyketide superclasses were also highlighted using MolNet-Enhancer, and the chemical space was dominantly occupied by flavonoid glycosides (Fig. 5A). Further analysis of these flavonoid glycosides could provide valuable insights into the potential biological activities and medicinal properties of *Helichrysum* species. Additionally, the identification of this subclass offers valuable insights into the potential bioactive compounds present in *Helichrysum* and their potential applications. One particular cluster of interest within the glycoside subclass consisted of compounds such as quercetin-3-glucuronide-7-glucoside, quercetin-3-*O*-glucoside, kaempferol-3-*O*-glucoside, kaempferol-3-*O*-glycosyl-6″-acetate, kaempferol-3-(6″-malonyl-glucoside), quercetin-3-galactoside-7-glucoside, quercetin-3,7-di-glucoside, kaempferol-3-*O*-galactoside, kaempferol-7-*O*-glucoside, kaempferol-7-*O*-galactoside and quercetin-3-*O*-glucosyl-6″-acetate (Fig. 5B). Flavonoid glycosides are naturally occurring compounds that consist of an aglycone linked to a sugar moiety known as a glycosidic bond.[56] The wide range of glycosides includes

compounds like quercetin glycosides and kaempferol glucosides. Different studies have pointed to the health benefits of flavonoid glycosides. For example, the latter have been found to exhibit anti-cancer effects.[58,59]

In addition to flavonoid glycosides, the *Helichrysum* metabolome is characterized by triterpenoids (Fig. 6A), which have also been reported to possess anti-cancer, anti-inflammatory, anti-bacterial and anti-viral activities. The *Helichrysum* triterpenoid profile comprises echinocystic acid, (3β,6β)-Stigmast-4-ene-3,6-diol, 31-norcycloartenol, neotigogenin, oleanolic acid, corosolic acid, ursonic acid, pinicolic acid A, cucurbitacin H, cucurbitacin F and oryzanol A (Fig. 6B). Triterpenoids are derived from squalene and are a subcategory of the larger group of terpenoids, which are among the most diverse and widespread natural compounds in plants.[79] The presence of triterpenoids in *Helichrysum* species, therefore, contributes to its therapeutic potential. For example, echinocystic and oleanolic acids have been reported to inhibit the growth of cancerous cells and inflammation.[80,81] Oryzanol A, on the other hand, has been found to possess antioxidant activity, protecting cells from oxidative damage.[82] Overall, the identification of these triterpenoids in *Helichrysum* species adds to the potential therapeutic value of these plants.

Interestingly, the triterpenoid profile of the three species differed distinctively. Cucurbitacin H, for example, was only present in *H. splendidum* and not detected in the other two species (Fig. 6B). Cucurbitacins are highly oxidized tetracyclic triterpenoids that interact with various cellular targets and have been reported to have anti-tumour properties.[83] Additionally, these compounds have been demonstrated to have analgesic, anti-inflammatory, anti-microbial and anti-viral activities.[84] *H. splendidum* can therefore be prioritized for the potential application in the aforementioned bioactivities using the triterpenoid profile when compared to the other two species. A unique profile in pentacyclic triterpenoids was also observed across the three species. Ursonic acid and neotigogenin were highly abundant in *H. petiolare*, echinocystic acid and oleanolic acid in *H. splendidum* and pinicolic acid A more abundant in *H. italicum* and *H. splendidum* (Fig. 6B). Pentacyclic triterpenoids are naturally occurring organic compounds that possess a five-ring structure. These compounds are synthesized either by and through the mevalonate or the non-mevalonate pathway and exhibit a wide variety of biological activities[85,86] such as the anti-cancer activity common to all the compounds. For example, ursonic acid and oleanolic acid inhibit the growth of various cancer cells, including lung, liver, ovarian, and prostate cancer by promoting apoptosis and inducing cell cycle arrest, making them attractive agents for the development of anti-cancer drugs.[87,88] The nuances of pentacyclic triterpenoids in the three *Helichrysum* plants therefore highlight the potential of this genus as an anti-cancer agent and suggest that *H. italicum* and *H. splendidum* species may have a higher potential as potent anti-cancer agents compared to other *Helichrysum* species. This metabolic diversity also underscores the intricate biochemical variations that can occur among closely related plant species.

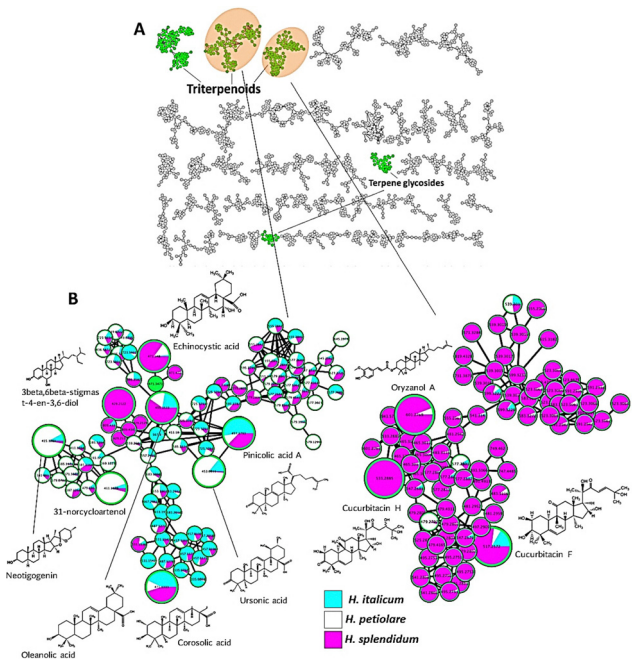

Fig. 6 Molecular network highlighting different triterpenoids in *Helichrysum* methanol extracts. (A) A cluster of terpenoids characterized by a molecular network showing triterpenoids identified in *Helichrysum* species. (B) The abundances differ in each plant species with some containing higher levels of triterpenoids than others or none.

The absence of specific triterpenoids such as cucurbitacin H in *H. italicum*, *H. splendidum* and *H. petiolare* despite their close botanical relationship highlights the potential divergence in their genetic and enzymatic machinery responsible for triterpenoid biosynthesis. Understanding the differences in triterpenoid profiles can provide insights into these plants' medicinal properties and potential therapeutic uses. The discovery of flavonoid glycosides and triterpenoids in the three species of *Helichrysum* marks a step forward in our understanding of the chemical composition of these plants. These results expand the knowledge base on the chemistry of *Helichrysum* plants, providing deconvoluted details of the various chemical classes that differentially define the metabolome of the *Helichrysum* plants. Such actionable insights point to the potential of *Helichrysum* as a valuable source of natural compounds with promising medicinal properties. Each unique chemical compound or group may possess properties that can be utilized in medicinal applications; as such, the chemistry of the genus may hold pharmaceutical potential. The synthesis of complex compounds by plants is an intricate process that gives rise to a wide range of metabolites. These metabolites have diverse modifications which may lead to a myriad of bioactivities, thus denoting their potential as important sources of pharmaceuticals. Among all plant metabolites, flavonoids are a prominent and structurally diverse group of compounds (Fig. 5) produced through a variety of modifications to yield numerous flavonoid derivatives. This structural variety of flavonoids has opened a plethora of pharmaceutical prospects.

3.2.1. Flavonoid decoration in *Helichrysum* species revealed by molecular networking. The *Helichrysum* chemical space is

predominantly occupied by flavonoid glycosides (Fig. 5) which exhibit structural diversity. The generated MolNetEnhancer networks revealed the exact modifications and decorations that take place, resulting in the biosynthesis of various complex flavonoids (Fig. 7). These modifications included methylation, demethylation, dehydration, decarboxylation, and glycosylation (Fig. 7) which are crucial for the formation of diverse flavonoids with varying potential bioactivities. Quercetin-3-oleate (*m/z* 300.227) underwent methylation resulting in the formation of 3-*O*-methylquercetin (*m/z* 315.0611) which in turn was glycosylated and demethylated, resulting in the formation of quercetin-3-*O*-glucoside (*m/z* 455.0554) (Fig. 7). This therefore highlights how chemical modifications result in the formation of different compounds of structural diversity. Molecular networking (MN) allowed for the elucidation of the different chemical decorations that occur in *Helichrysum* plants. MN compares the fragmentation spectra of related molecules based on their structural similarity.[89] The generated MN networks were therefore used to identify clusters of related compounds which allowed for the visual identification of patterns that revealed structural variation. Additionally, since metabolites that are grouped in an MN share a common fragmentation pattern translated into a molecular fingerprint, the different chemical modifications were revealed (Fig. 7).

In the context of glycosylation in flavonoids, the MN generated enabled the identification of glycosylated forms by grouping spectra with similar glycosylation patterns. Quercetin-3-*O*-glucoside (*m/z* 455.0554), quercetin-3-*O*-(6″-malonylglucoside) (*m/z* 549.099) and quercetin-3-*O*-glycosyl-6″-acetate (*m/z* 565.098) were closely structured together (Fig. 7). This revealed the chemical modifications and relationships among the mass spectra, indicating the presence and types of sugar moieties attached to the flavonoid structures. Glycosylation is one of the major modifications that occurs in numerous biological processes, resulting in the formation of a wide range of NPs.[90] This modification is catalyzed by glycosyltransferases (GTs) which transfer a sugar moiety from nucleotide-sugar donors to the aglycones.[91] Donor molecules include uridine 5-diphosphate (UDP) sugars, whereby GTs are referred to as UDP-glycosyltransferases (UGTs).[92] These UGTs are located in the cytosol, where they are responsible for the biosynthesis of a wide range of NPs such as flavonoids, phenylpropanoids and steroids.[93] In the case of flavonoids, the sugars are glucose moieties which accumulate in either mono-, di- or tri-glycosides.[90] Glycosylation of flavonoids has been reported to increase the chemical stability, solubility, and bioavailability of flavonoids, enabling access to plant membrane transportation systems which recognise the glycosylated forms instead of the aglycones.[94]

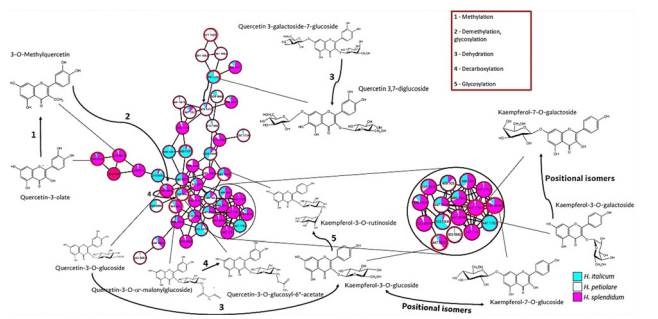

Fig. 7 Flavonoid modifications in *Helichrysum* species. Key flavonoid modifications observed in *Helichrysum* species shown by the chemical structures, highlighting points at which these modifications occur.

Furthermore, glycosylation increases structural complexity and diversity which has sparked interest since flavonoid glycosides have been reported to modulate numerous pharmacokinetic parameters such as anti-viral, anti-inflammatory, anti-cancer, and anti-bacterial properties.[95,96] These decorations can also be used as biomarkers to distinguish different species from each other. In this study, kaempferol-3-*O*-glucoside was detected in high levels in *H. splendidum* and *H. petiolare*, but the opposite was observed with the glycosylated form (kaempferol-3-*O*-rutinoside) (Fig. 7). Nengovhela *et al.* (2021)[97] reported on the modifications of the flavonoid chemical in two closely related species *Coccinia grandis* and *Coccinia rehmannii*. In this study, sugar acylation by cinnamic acids such as caffeic acid and coumaric acid was noted in *C. rehmannii* but not in *C. grandis*, revealing that glycosylation patterns do differ in species of the same genus. These findings therefore suggest that the glycosylation patterns of flavonoids can vary in closely related species, resulting in complex and diverse metabolomes which can be correlated to species-specific pharmacological properties.[53,54] Molecular networking comprehensively revealed the flavonoid chemical alterations across *Helichrysum* species, creating a detailed molecular map that reflects the distinctive biochemical profile of the *Helichrysum* genus. Further elucidation of the metabolomic complexity in this genus through pathway analysis therefore describes the annotated chemistry in a biological context and space.

3.3. Pathway analysis: understanding the metabolic landscape of *Helichrysum* species

To situate the identified metabolites in three *Helichrysum* species in a metabolome view context and to identify the most enriched biological pathways based on the measured metabolome, pathway analysis was employed using the metabolic pathway analysis (MetPA) – an integral module of MetaboAnalyst 5.0[98] (Table S2). Pathway enrichment analysis is fundamental for integrating metabolomics data into biological contexts. Comprehensive databases, such as KEGG, contain manually selected pathways exhibiting well-structured functions or biological processes involving several compounds. Enrichment methods such as over-representation analysis (ORA) can be used with these resources. These methods compare the list of metabolites detected in an experiment with the metabolites associated with each pathway in the database to identify significantly enriched pathways. ORA calculates a statistical score, such as a *p*-value, to determine the significance of pathway enrichment.[99] This provides insights into the underlying biological mechanisms and pathways that the measured metabolome impacts. The overview-representation and pathway topological analysis based on a hypergeometric test

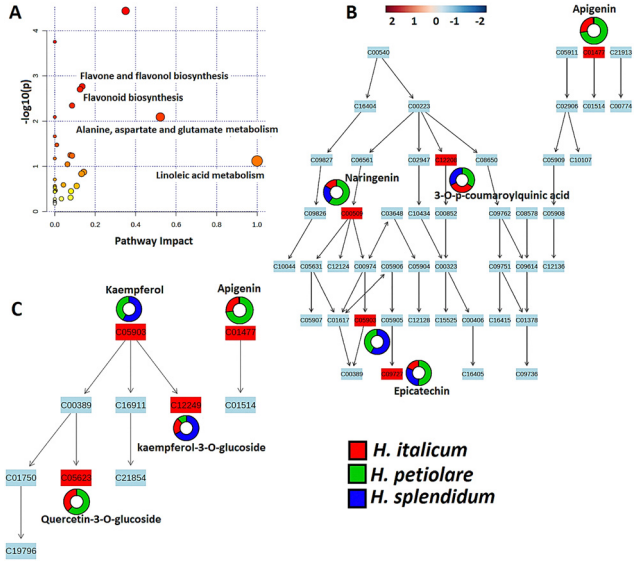

Fig. 8 Summary of pathway analysis in *Helichrysum* species. (A) The graph, "metabolome view," contains all the matched pathways (the metabolome) arranged by *p*-values (pathway enrichment analysis) on the *y*-axis, and the pathway impact values (pathway topology analysis) on the *x*-axis, with the numbers corresponding to the mapped pathways listed in Table S2. The node colour is based on the *p*-value, and the node radius is defined by the pathway impact values. The latter is the cumulative percentage from the matched metabolite nodes. (B) Topology map of flavonoid biosynthesis and (C) flavone and flavonol biosynthesis pathways.

algorithm and relative-betweenness centrality respectively revealed eight significant metabolomic pathways with an impact score >0.1, out of a total of 35 pathways that were impacted in the three *Helichrysum* plant species (Table S2). The significant pathways included the flavonoid biosynthesis, flavone and flavonol biosynthesis, alanine, aspartate and glutamate metabolism and linoleic acid metabolism pathways (Fig. 8A), with flavonoid biosynthesis and the flavone and flavonol biosynthesis pathways showing the highest hits. In the flavonoid biosynthesis (Fig. 8B) and the flavone and flavonol biosynthesis (Fig. 8C) pathways, aglycones such as kaempferol (C05903), apigenin (C01477) and naringenin (C00509) could be mapped onto them. Aglycones are core structures of flavonoids that lack sugar molecules attached to them. These aglycones play a role in the potential of flavonoids and contribute to their wide range of health-promoting effects such as their anti-inflammatory, antioxidant, and anti-cancer properties.[100] The aglycones identified serve as precursors to produce diverse compounds such as flavonoids which have been identified in the *Helichrysum* species, as highlighted in Fig. 5. The presence of these aglycones in *Helichrysum* species therefore suggests their potential therapeutic applications and highlights the importance of understanding the biosynthetic pathways involved in their production.

The alanine, aspartate and glutamate metabolism and linoleic acid metabolism pathways were also significantly impacted in the three *Helichrysum* species (Fig. 8A). Alanine, aspartate, and glutamate metabolism in plants is crucial since it contributes to the production of compounds with therapeutic benefits. These amino acids serve as the foundation for metabolites, like alkaloids, flavonoids and polyphenols that have been recognized for their properties.[101] These amino acids act as building blocks for metabolites like alkaloids, flavonoids and polyphenols which have been reported for their medicinal properties.[101,102] Additionally, these metabolic pathways are associated with the plants' stress response since certain secondary metabolites serve as compounds against herbivores and environmental stressors. Furthermore, specific glutamate-related compounds found in plants may interact with neurotransmitter systems in humans and offer potential benefits for maintaining a healthy nervous system.[103] Linoleic acid metabolism, on the other hand, is a crucial process that significantly impacts plant growth and development and the capacity to adapt to environmental factors. Furthermore, linoleic acid plays a role in producing triacylglycerols (TAGs) which serve as an energy reserve for seed development and germination.[104]

4. Conclusion

The computational metabolomics study reported herein provides a global metabolic chart of *Helichrysum italicum*, *Helichrysum petiolare* and *Helichrysum splendidum* plants, highlighting the diverse and complex chemical composition of these plants. These species are characterized by a wide spectrum of chemical (sub)classes, which can be grouped into superclasses ranging from lipid and lipid-like molecules to organic oxygen compounds. The analysis reveals that the chemical superclasses dominating the metabolomic landscape of these *Helichrysum* species are (i) lipids and lipid-like molecules, (ii) phenylpropanoids and polyketides, and (iii) organic oxygen compounds. This extensive variety of chemical classes not only underscores the biochemical diversity within the *Helichrysum* genus but also contributes to their potential pharmacological properties. Notably, these chemical constituents have been linked to various therapeutic activities, such as anti-inflammatory, antioxidant, anti-cancer, and antimicrobial effects. Collectively, the rich metabolomic landscape within *Helichrysum* plants suggests that they represent a valuable reservoir of bioactive compounds, positioning this genus as a promising source for the development of novel therapeutic agents. The actionable insights derived from this study further emphasize the potential of *Helichrysum* species in the discovery of natural compounds with significant medicinal applications. Although this study maps key *Helichrysum* metabolites with known bioactivities, functional validation remains limited. Bioactivity inferences are based on putative annotations and the literature, without direct experimental assays. Moreover, species selection and extraction constraints may have excluded additional metabolites with pharmacological relevance, necessitating broader bioassays and expanded taxonomic sampling for translational insights.

Author contributions

M. M. L.: methodology, data analysis, data curation, writing of the original draft, and writing – review and editing. F. T.:

supervision, conceptualization, methodology, data analysis, data curation, writing of the original draft, writing – review and editing, and project administration. N. E. M.: supervision, methodology, and editing. M. S. C. and A. P. K.: supervision and editing. All authors read and approved the manuscript.

Conflicts of interest

There are no conflicts to declare.

Data availability

Spectral data are available through the GNPS links provided in the supplementary information (SI). Supplementary information is available. See DOI: **https://doi.org/10.1039/d5mo00118h**.

In addition to these GNPS links, spectral data are submitted to the MassIVE repository: dataset MSV000098231, **https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=a21755d761c445239606e89f1b1cb6d5**.

Acknowledgements

University of Venda's Biochemistry Department is thanked for access to the Shimadzu LCMS-9030 qTOF, and Shimadzu South Africa is thanked for all the support on the instrumentation. The University of Johannesburg is highly thanked for the Commonwealth scholarship to M. M. L.

Notes and references

1 E. Pooley, *Mountain flowers: A field guide to the flora of the Drakensberg and Lesotho*, Flora Publ. Trust, 2003.
2 F. Les, A. Venditti, G. Cásedas, C. Frezza, M. Guiso, F. Sciubba, M. Serafini, A. Bianco, M. S. Valero and V. López, *Ind. Crops Prod.*, 2017, **108**, 295–302.
3 Z. Mao, C. Gan, J. Zhu, N. Ma, L. Wu, L. Wang and X. Wang, *Bioorg. Med. Chem. Lett.*, 2017, **27**, 2812–2817.
4 C. G. Pereira, L. Barreira, S. Bijttebier, L. Pieters, V. Neves, M. J. Rodrigues, R. Rivas, J. Varela and L. Custódio, *J. Pharm. Biomed. Anal.*, 2017, **145**, 593–603.
5 W. B. Adeosun and G. Prinsloo, *S. Afr. J. Bot.*, 2025, **185**, 383–398.
6 W. B. Adeosun, G. K. More, P. Steenkamp and G. Prinsloo, *Front. Mol. Biosci.*, 2022, **9**, 961859.
7 R. R. Da Silva, P. C. Dorrestein and R. A. Quinn, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 12549–12550.
8 S. Rogers, C. W. Ong, J. Wandy, M. Ernst, L. Ridder and J. J. J. Van Der Hooft, *Faraday Discuss.*, 2019, **218**, 284–302.
9 A. T. Aron, E. C. Gentry, K. L. McPhail, L. F. Nothias, M. Nothias-Esposito, A. Bouslimani, D. Petras, J. M. Gauglitz, N. Sikora, F. Vargas, J. J. J. van der Hooft, M. Ernst, K. Bin Kang, C. M. Aceves, A. M. Caraballo-Rodríguez, I. Koester, K. C. Weldon, S. Bertrand, C. Roullier, K. Sun, R. M. Tehan, C. A. Boya P, M. H. Christian, M. Gutiérrez, A. M. Ulloa, J. A. Tejeda Mora, R. Mojica-Flores, J. Lakey-Beitia, V. Vásquez-Chaves, Y. Zhang, A. I. Calderón, N. Tayler, R. A. Keyzers, F. Tugizimana, N. Ndlovu, A. A. Aksenov, A. K. Jarmusch, R. Schmid, A. W. Truman, N. Bandeira, M. Wang and P. C. Dorrestein, *Nat. Protoc.*, 2020, **15**, 1954–1991.
10 M. A. Beniddir, K. Bin Kang, G. Genta-Jouve, F. Huber, S. Rogers and J. J. J. van der Hooft, *Nat. Prod. Rep.*, 2021, **38**, 1967–1993.
11 J. Y. Yang, L. M. Sanchez, C. M. Rath, X. Liu, P. D. Boudreau, N. Bruns, E. Glukhov, A. Wodtke, R. De Felicio, A. Fenner, W. R. Wong, R. G. Linington, L. Zhang, H. M. Debonsi, W. H. Gerwick and P. C. Dorrestein, *J. Nat. Prod.*, 2013, **76**, 1686–1699.
12 R. Verpoorte, Y. H. Choi, N. R. Mustafa and H. K. Kim, *Phytochem. Rev.*, 2008, **7**, 525–537.
13 H. K. Kim and R. Verpoorte, *Phytochem. Anal.*, 2010, **21**, 4–13.
14 L. F. Nothias, D. Petras, R. Schmid, K. Dührkop, J. Rainer, A. Sarvepalli, I. Protsyuk, M. Ernst, H. Tsugawa, M. Fleischauer, F. Aicheler, A. A. Aksenov, O. Alka, P. M. Allard, A. Barsch, X. Cachet, A. M. Caraballo-Rodriguez, R. R. Da Silva, T. Dang, N. Garg, J. M. Gauglitz, A. Gurevich, G. Isaac, A. K. Jarmusch, Z. Kameník, K. Bin Kang, N. Kessler, I. Koester, A. Korf, A. Le Gouellec, M. Ludwig, C. Martin H, L. I. McCall, J. McSayles, S. W. Meyer, H. Mohimani, M. Morsy, O. Moyne, S. Neumann, H. Neuweger, N. H. Nguyen, M. Nothias-Esposito, J. Paolini, V. V. Phelan, T. Pluskal, R. A. Quinn, S. Rogers, B. Shrestha, A. Tripathi, J. J. J. van der Hooft, F. Vargas, K. C. Weldon, M. Witting, H. Yang, Z. Zhang, F. Zubeil, O. Kohlbacher, S. Böcker, T. Alexandrov, N. Bandeira, M. Wang and P. C. Dorrestein, *Nat. Methods*, 2020, **17**, 905–908.
15 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 and N. Bandeira, *Nat. Biotechnol.*, 2016, **34**, 828–837.

16 H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn and M. Arita, *Nat. Methods*, 2015, **12**, 523–526.

17 H. Mohimani, A. Gurevich, A. Shlemov, A. Mikheenko, A. Korobeynikov, L. Cao, E. Shcherbin, L.-F. Nothias, P. C. Dorrestein and P. A. Pevzner, *Nat. Commun.*, 2018, **9**, 4035.

18 P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, *Genome Res.*, 2003, **13**, 2498–2504.

19 K. Nakamura, N. Shimura, Y. Otabe, A. Hirai-Morita, Y. Nakamura, N. Ono, M. A. Ul-Amin and S. Kanaya, *Plant Cell Physiol.*, 2013, **54**, e4.

20 Y. Nakamura, F. Mochamad Afendi, A. Kawsar Parvin, N. Ono, K. Tanaka, A. Hirai Morita, T. Sato, T. Sugiura, M. Altaf-Ul-Amin and S. Kanaya, *Plant Cell Physiol.*, 2014, **55**, e7.

21 H. E. Pence and A. Williams, *J. Chem. Educ.*, 2010, **87**, 1123–1124.

22 Y. Wang, J. Xiao, T. O. Suzek, J. Zhang, J. Wang and S. H. Bryant, *Nucleic Acids Res.*, 2009, **37**, 623–633.

23 S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang and S. H. Bryant, *Nucleic Acids Res.*, 2016, **44**, D1202–D1213.

24 M. Whittle, P. Willett, W. Klaffke and P. van Noort, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 449–457.

25 J. J. J. van der Hooft, J. Wandy, F. Young, S. Padmanabhan, K. Gerasimidis, K. E. V. Burgess, M. P. Barrett and S. Rogers, *Anal. Chem.*, 2017, **89**, 7569–7577.

26 M. Ernst, K. Bin Kang, A. M. Caraballo-Rodríguez, L.-F. Nothias, J. Wandy, C. Chen, M. Wang, S. Rogers, M. H. Medema, P. C. Dorrestein and J. J. J. van der Hooft, *Metabolites*, 2019, **9**, 144.

27 Y. Djoumbou Feunang, R. Eisner, C. Knox, L. Chepelev, J. Hastings, G. Owen, E. Fahy, C. Steinbeck, S. Subramanian, E. Bolton, R. Greiner and D. S. Wishart, *J. Cheminform.*, 2016, **8**, 61.

28 Y. H. Choi and R. Verpoorte, *Phytochem. Anal.*, 2014, **25**, 289–290.

29 O. Yanes, R. Tautenhahn, G. J. Patti and G. Siuzdak, *Anal. Chem.*, 2011, **83**, 2152–2161.

30 F. Maltese, F. van der Kooy and R. Verpoorte, *Nat. Prod. Commun.*, 2009, **4**, 447–454.

31 J. Pezzatti, J. Boccard, S. Codesido, Y. Gagnebin, A. Joshi, D. Picard, V. González-Ruiz and S. Rudaz, *Anal. Chim. Acta*, 2020, **1105**, 28–44.

32 H. W. Kim, M. Wang, C. A. Leber, L. F. Nothias, R. Reher, K. Bin Kang, J. J. J. Van Der Hooft, P. C. Dorrestein, W. H. Gerwick and G. W. Cottrell, *J. Nat. Prod.*, 2021, **84**, 2795–2807.

33 Y. B. Saalmann and M. B. Calford, in *Neuroactive Steroids in Brain Function, Behavior and Neuropsychiatric Disorders*, ed. M. S. Ritsner and A. Weizman, Springer, Netherlands, Dordrecht, 2008, pp. 187–200.

34 S.-K. Wang, C.-F. Dai and C.-Y. Duh, *J. Nat. Prod.*, 2006, **69**, 103–106.

35 J. Q. Reimão, A. E. Migotto, M. H. Kossuga, R. G. S. Berlinck and A. G. Tempone, *Parasitol. Res.*, 2008, **103**, 1445–1450.

36 A. R. Díaz-Marrero, G. Porras, Z. Aragón, J. M. de la Rosa, E. Dorta, M. Cueto, L. D'Croz, J. Maté and J. Darias, *J. Nat. Prod.*, 2011, **74**, 292–295.

37 X. Q. Song, L. L. Tian, T. Ye, H. Liu and H. Zhang, *Phytochemistry*, 2023, **213**, 113787.

38 T. Walle, *Int. J. Mol. Sci.*, 2009, **10**, 5002–5019.

39 L. Wen, Y. Jiang, J. Yang, Y. Zhao, M. Tian and B. Yang, *Ann. N. Y. Acad. Sci.*, 2017, **1398**, 120–129.

40 S. S. Elhady, E. E. Eltamany, A. E. Shaaban, A. A. Bagalagel, Y. A. Muhammad, N. M. El-Sayed, S. N. Ayyad, A. A. M. Ahmed, M. S. Elgawish and S. A. Ahmed, *Plants*, 2020, **9**, 1–17.

41 H. Nawaz, M. A. Shad, N. Rehman, H. Andaleeb and N. Ullah, *Braz. J. Pharm. Sci.*, 2020, **56**, e17129.

42 W. B. Dunn and C. L. Winder, in *Methods in Systems Biology*, ed. D. Jameson, M. Verma and H. V. B. T.-M. E. Westerhoff, Academic Press, 2011, vol. 500, pp. 277–297.

43 T. Hyötyläinen, *RSC Chromatogr. Monogr.*, 2013, 11–42.

44 S. Fiorito, F. Epifano, F. Preziuso, V. A. Taddeo and S. Genovese, *Prog. Chem. Org. Nat. Prod.*, 2019, **108**, 143–205.

45 V. Leláková, S. Béraud-Dufour, J. Hošek, K. Šmejkal, V. Prachyawarakorn, P. Pailee, C. Widmann, J. Václavík, T. Coppola, J. Mazella, N. Blondeau and C. Heurteaux, *J. Ethnopharmacol.*, 2020, **263**, 113147.

46 M. Shahinozzaman, B. Basak, R. Emran, P. Rozario and D. N. Obanda, *Fitoterapia*, 2020, **147**, 104775.

47 W.-Y. Zhou, Z.-H. Xi, N.-N. Du, L. Ye, M.-H. Jiang, J.-L. Hao, B. Lin, G.-D. Yao, X.-X. Huang and S.-J. Song, *Chin. Chem. Lett.*, 2023, 109030.

48 H. U. Kim, *Plants*, 2020, **9**, 871.

49 M. C. Suh, H. Uk Kim and Y. Nakamura, *J. Exp. Bot.*, 2022, **73**, 2715–2720.

50 M. Eggersdorfer and A. Wyss, *Arch. Biochem. Biophys.*, 2018, **652**, 18–26.

51 N. Duhan, S. Barak and D. Mudgil, *Biointerface Res. Appl. Chem.*, 2020, **10**, 6676–6687.

52 D. J. McClements and B. Öztürk, *Foods*, 2021, **10**, 1–17.

53 L. K. Caesar, O. M. Kvalheim and N. B. Cech, *Anal. Chim. Acta*, 2018, **1021**, 69–77.

54 C. Rawlinson, D. Jones, S. Rakshit, S. Meka, C. S. Moffat and P. Moolhuijzen, *Sci. Rep.*, 2020, **10**, 6043.

55 S. L. Rodríguez De Luna, R. E. Ramírez-Garza and S. O. Serna Saldívar, *Sci. World J.*, 2020, **2020**, 1–38.

56 A. N. Panche, A. D. Diwan and S. R. Chandra, *J. Nutr. Sci.*, 2016, **5**, e47.

57 M. C. Dias, D. C. G. A. Pinto, H. Freitas, C. Santos and A. M. S. Silva, *Phytochemistry*, 2020, **170**, 112199.

58 N. Saini, S. K. Gahlawat and V. Lather, *Plant Biotechnology: Recent Advancements and Developments*, Springer, Singapore, Singapore, 2017, pp. 255–270.
59 M. M. Jucá, F. M. S. Cysne Filho, J. C. de Almeida, D. S. Mesquita, J. R. M. Barriga, K. C. F. Dias, T. M. Barbosa, L. C. Vasconcelos, L. K. A. M. Leal, J. E. Ribeiro and S. M. M. Vasconcelos, *Nat. Prod. Res.*, 2020, **34**, 692–705.
60 W. Liu, J. Li, X. Zhang, Y. Zu, Y. Yang, W. Liu, Z. Xu, H. Gao, X. Sun, X. Jiang and Q. Zhao, *J. Agric. Food Chem.*, 2020, **68**, 10489–10516.
61 H. Lu, Z. Tian, Y. Cui, Z. Liu and X. Ma, *Compr. Rev. Food Sci. Food Saf.*, 2020, **19**, 3130–3158.
62 T. M. Hung, M. Na, P. T. Thuong, N. D. Su, D. Sok, K. S. Song, Y. H. Seong and K. Bae, *J. Ethnopharmacol.*, 2006, **108**, 188–192.
63 R. Kurata, M. Adachi, O. Yamakawa and M. Yoshimoto, *J. Agric. Food Chem.*, 2007, **55**, 185–190.
64 Z. Zhao, H. S. Shin, H. Satsu, M. Totsuka and M. Shimizu, *J. Agric. Food Chem.*, 2008, **56**, 3863–3868.
65 X. Tian, L. An, L.-Y. Gao, J.-P. Bai, J. Wang, W.-H. Meng, T.-S. Ren and Q.-C. Zhao, *CNS Neurosci. Ther.*, 2015, **21**, 575–584.
66 S. E. Yazdi, G. Prinsloo, H. M. Heyman, C. B. Oosthuizen, T. Klimkait and J. J. M. Meyer, *S. Afr. J. Bot.*, 2019, **126**, 328–339.
67 S. Emamzadeh Yazdi, H. M. Heyman, G. Prinsloo, T. Klimkait and J. J. M. Meyer, *Front. Pharmacol.*, 2022, **13**, 1–11.
68 W. B. Adeosun, G. K. More, P. Steenkamp and G. Prinsloo, *Front. Mol. Biosci.*, 2022, **9**, 1–13.
69 A. C. Gradinaru, M. Silion, A. Trifan, A. Miron and A. C. Aprotosoaie, *Nat. Prod. Res.*, 2014, **28**, 2076–2080.
70 M. I. Hussain, Q. A. Syed, M. N. K. Khattak, B. Hafez, M. J. Reigosa and A. El-Keblawy, *Biologia*, 2019, **74**, 863–888.
71 S. Jasim and Y. Mustafa, *Iraqi J. Pharm.*, 2022, **18**, 104–125.
72 O. M. Tsivileva, O. V. Koftin and N. V. Evseeva, *Antibiotics*, 2022, **11**, 1156.
73 V. Flores-Morales, A. P. Villasana-Ruíz, I. Garza-Veloz, S. González-Delgado and M. L. Martinez-Fierro, *Molecules*, 2023, **28**, 2413.
74 A. Irfan, L. Rubab, M. U. Rehman, R. Anjum, S. Ullah, M. Marjana, S. Qadeer and S. Sana, *Heterocycl. Commun.*, 2020, **26**, 46–59.
75 D. T. Hog, R. Webster and D. Trauner, *Nat. Prod. Rep.*, 2012, **29**, 752–779.
76 K. Guo, Y. Liu and S.-H. Li, *Nat. Prod. Rep.*, 2021, **38**, 2293–2314.
77 K. Li and K. R. Gustafson, *Nat. Prod. Rep.*, 2021, **38**, 1251–1281.
78 Y. Chen, J. Zhao, S. Li and J. Xu, *Nat. Prod. Rep.*, 2019, **36**, 263–288.
79 S. Perveen and A. Al-Taweel, *Terpenes Terpenoids*, 2018, **1**, 1–12.
80 D. He, G. Hu, A. Zhou, Y. Liu, B. Huang, Y. Su, H. Wang, B. Ye, Y. He, X. Gao, S. Fu and D. Liu, *Front. Pharmacol.*, 2021, **12**, 787771.
81 W. Baer-Dubowska, M. Narożna and V. Krajka-Kuźniak, *Molecules*, 2021, **26**, 4957.
82 I. O. Minatel, F. V. Francisqueti, C. R. Corrêa and G. P. P. Lima, *Int. J. Mol. Sci.*, 2016, **17**(8), 1107.
83 H. S. Tuli, P. Rath, A. Chauhan, A. Ranjan, S. Ramniwas, K. Sak, D. Aggarwal, M. Kumar, K. Dhama, E. H. C. Lee *et al.*, *Biomolecules*, 2023, **13**, 57.
84 E. E. Delgado-Tiburcio, J. Cadena-Iñiguez, E. Santiago-Osorio, L. D. Ruiz-Posadas, I. Castillo-Juárez, I. Aguiñiga-Sánchez and M. Soto-Hernández, 2022, preprint, DOI: **10.3390/ph15111325**.
85 S. Banerjee, S. Bose, S. C. Mandal, S. Dawn, U. Sahoo, M. A. Ramadan and S. K. Mandal, *Egypt J. Chem.*, 2019, **62**, 13–35.
86 M. Malík, J. Velechovský and P. Tlustoš, *Fitoterapia*, 2021, **151**, 104845.
87 G. Fontana, M. Bruno, M. Notarbartolo, M. Labbozzetta, P. Poma, A. Spinella and S. Rosselli, *Bioorg. Chem.*, 2019, **90**, 103054.
88 V. Khwaza, O. O. Oyedeji and B. A. Aderibigbe, *Int. J. Mol. Sci.*, 2020, **21**(16), DOI: **10.3390/ijms21165920**.
89 R. A. Quinn, L.-F. Nothias, O. Vining, M. Meehan, E. Esquenazi and P. C. Dorrestein, *Trends Pharmacol. Sci.*, 2017, **38**, 143–154.
90 P. Tiwari, R. S. Sangwan and N. S. Sangwan, *Biotechnol. Adv.*, 2016, **34**, 714–739.
91 A. Rai, K. Saito and M. Yamazaki, *Plant J.*, 2017, **90**, 764–787.
92 K. Yonekura-Sakakibara, T. Tohge, R. Niida and K. Saito, *J. Biol. Chem.*, 2007, **282**, 14932–14941.
93 J. Ross, Y. Li, E. Lim and D. J. Bowles, *Genome Biol.*, 2001, **2**, 1–6.
94 J. M. Augustin, V. Kuzina, S. B. Andersen and S. Bak, *Phytochemistry*, 2011, **72**, 435–457.
95 J. Xiao, E. Capanoglu, A. R. Jassbi and A. Miron, *Crit. Rev. Food Sci. Nutr.*, 2016, **56**, S29–S45.
96 K. G. Vanegas, A. B. Larsen, M. Eichenberger, D. Fischer, U. H. Mortensen and M. Naesby, *Microb. Cell Fact.*, 2018, **17**, 1–10.
97 N. Nengovhela, P. A. Steenkamp and N. E. Madala, *Natl. Acad. Sci. Lett.*, 2021, **44**, 209–213.
98 Y. Lu, Z. Pang and J. Xia, *Briefings Bioinf.*, 2023, **24**, bbac553.
99 L. Liu and J. Ruan, *Proceedings – 2013 IEEE International Conference on Bioinformatics and Biomedicine*, IEEE BIBM, 2013, vol. 2013, pp. 218–221.
100 T. B. Ribeiro, A. Melo, A. A. Vilas-Boas and M. Pintado, Flavonoids, in *Natural secondary metabolites: from nature, through science, to industry*, ed. M. Carocho, S. A. Heleno and L. Barros, Springer International Publishing, 2023, pp. 73–105.
101 T. M. Hildebrandt, A. Nunes Nesi, W. L. Araújo and H. P. Braun, *Mol. Plant*, 2015, **8**, 1563–1579.
102 C. dos, R. Nunes, M. Barreto Arantes, S. Menezes de Faria Pereira, L. Leandro da Cruz, M. de Souza Passos, L. Pereira de Moraes, I. J. C. Vieira and D. Barros de Oliveira, *Molecules*, 2020, **25**, 3726.
103 Y. Liu, S. Wang, J. Kan, J. Zhang, L. Zhou, Y. Huang and Y. Zhang, *Curr. Neuropharmacol.*, 2019, **18**, 260–276.
104 J. Mercola and C. R. D'Adamo, *Nutrients*, 2023, **15**, 3129.