



Charting the *Cannabis* plant chemical space with computational metabolomics

Akhona Myoli¹ · Mpho Choene¹ · Abidemi Paul Kappo¹ · Ntakadzeni Edwin Madala² · Justin J. J. van der Hooft^{1,3} · Fidele Tugizimana^{1,4,5}

Received: 21 February 2024 / Accepted: 2 May 2024
© The Author(s) 2024

Abstract

Introduction The chemical classification of *Cannabis* is typically confined to the cannabinoid content, whilst *Cannabis* encompasses diverse chemical classes that vary in abundance among all its varieties. Hence, neglecting other chemical classes within *Cannabis* strains results in a restricted and biased comprehension of elements that may contribute to chemical intricacy and the resultant medicinal qualities of the plant.

Objectives Thus, herein, we report a computational metabolomics study to elucidate the *Cannabis* metabolic map beyond the cannabinoids.

Methods Mass spectrometry-based computational tools were used to mine and evaluate the methanolic leaf and flower extracts of two *Cannabis* cultivars: Amnesia haze (AMNH) and Royal dutch cheese (RDC).

Results The results revealed the presence of different chemical compound classes including cannabinoids, but extending it to flavonoids and phospholipids at varying distributions across the cultivar plant tissues, where the phenylpropanoid super-class was more abundant in the leaves than in the flowers. Therefore, the two cultivars were differentiated based on the overall chemical content of their plant tissues where AMNH was observed to be more dominant in the flavonoid content while RDC was more dominant in the lipid-like molecules. Additionally, in silico molecular docking studies in combination with biological assay studies indicated the potentially differing anti-cancer properties of the two cultivars resulting from the elucidated chemical profiles.

Conclusion These findings highlight distinctive chemical profiles beyond cannabinoids in *Cannabis* strains. This novel mapping of the metabolomic landscape of *Cannabis* provides actionable insights into plant biochemistry and justifies selecting certain varieties for medicinal use.

Keywords *Cannabis* · Medicinal · LC-MS/MS · Computational tools · Metabolic map · Chemovars · Cultivars

1 Introduction

The *Cannabis* plant is of the Cannabaceae family and is known to grow in various forms of cultivars or strains. *Cannabis* is popularly known for both its recreational use and there is emerging awareness of its medicinal properties (Cicaloni et al., 2022). The taxonomical classification of the *Cannabis* genus has been largely debated by taxonomists and botanists for decades. Majority of scientists classify the *Cannabis* genus as one parent species in the form of *C. sativa* which is subdivided into two sub-species *Sativa* and *Indica*. Others classify *Cannabis* into three separate species based on morphological differences in the form of *C. sativa* (taller plant with thinner fan leaves), *C. indica* (shorter, bushier plant with wider fan leaves) and *C. ruderalis*

✉ Justin J. J. van der Hooft
justin.vanderhooft@wur.nl

✉ Fidele Tugizimana
ftugizimana@uj.ac.za; Fidele.Tugizimana@omnia.co.za

¹ Department of Biochemistry, University of Johannesburg, Auckland Park, Johannesburg 2006, South Africa

² Department of Biochemistry and Microbiology, University of Venda, Thohoyandou, South Africa

³ Bioinformatics Group, Wageningen University, Wageningen 6708 PB, the Netherlands

⁴ International Research and Development Division, Omnia Group, Ltd., Bryanston, Johannesburg 2021, South Africa

⁵ National Institute for Theoretical and Computational Sciences, Johannesburg, South Africa

(wild-looking with a canonical shape, has fewer branches compared to *C. sativa* and *C. indica*) (Watts, 2006; Rupasinghe et al., 2020). Despite all the classification debates associated with *Cannabis*, the plant continues to grow successfully in various environments worldwide. There are currently more than 700 strains of *Cannabis* with trivial names such as Purple Kush, White Widow, Amnesia Haze and Royal Dutch cheese. When drawing focus on the chemical composition of the numerous *Cannabis* strains or cultivars, the chemical classification or differentiation is limited to several categories or chemovars which are based only on the plant's cannabinoid content (Gloss, 2015; Lewis et al., 2018). The cannabinoid chemical class comprises a group of compounds that naturally occur in *Cannabis* varieties. The chemical structure of these compounds is characterized by a C21 terpene phenolic backbone, which can be traced in parent cannabinoids, cannabinoid derivatives, and transformation products. These cannabinoids are further divided into sub-classes, namely; cannabichromene (CBC), cannabidiol (CBD), cannabielsoin (CBE), cannabigerol (CBG), cannabicyclol (CBL), cannabinol (CBN), cannabinodiol (CBND), cannabitrilol (CBT), Δ^8 -trans-tetrahydrocannabinol (Δ^8 -THC), Δ^9 -trans-tetrahydrocannabinol (Δ^9 -THC), and miscellaneous-type cannabinoids (Filipiuc et al., 2021; Radwan et al., 2021; Procaccia et al., 2022). These cannabinoid classes are distributed across different plant tissues where compounds such as Δ^8 -THC and/or Δ^9 -THC are highly expressed by the glandular trichomes of female flowers (Lubell and Brand, 2018; Conneely et al., 2021).

Type I *Cannabis* cultivars are characterized by high THC levels and are used for both medicinal and recreational purposes. Type III contains high levels of CBD, which has been acknowledged for its therapeutic benefits more than type I. This has also led to the emergence of additional chemovars such as type II *Cannabis* which is known to have equal levels of both THC and CBD compounds (Lewis et al., 2018). Nonetheless, some of the reported medicinal applications and benefits of *Cannabis* include alleviating nausea and pain in cancer patients receiving chemotherapy, and alleviating neurological symptoms such as anxiety, stress and sleeping problem (Moltke & Hindocha, 2021). Additionally, studies have shown promising evidence of the therapeutic effects of the *Cannabis* plant (including drugs and essential oils derived from the plant) in the suppression of diseases such as cancer, Alzheimer's disease, and Huntington's disease (Kogan & Mechoulam, 2007; Odiaka et al., 2022). Although *Cannabis* is expected to have a rich, complex, and diverse chemical composition, so far, the only bioactive chemical constituents that have been well-studied from the plant are cannabinoids (Sarma et al., 2020). We note that this is unsurprising given the effects that cannabinoids have on people as well as the medical use. In general, the relative

abundance of aroma-defining terpenes/terpenoids in *Cannabis* are highly regarded by consumers, medical scientists, and practitioners (Ferber et al., 2020; Duggan, 2021; LaVigne et al., 2021). Studies have also shown that in *Cannabis* crude extracts, cannabinoids work synergistically with other specialized metabolites, such as other terpenes/terpenoids, to produce a heightened pharmacological effect compared to isolated compounds (entourage effect) (Maayah et al., 2020; Goerl et al., 2021).

Nonetheless, the cannabinoid chemical class is still considered as the major chemical class and is distributed in varying amounts across all the different plant parts. The typical anatomy of the *Cannabis* plant is comprised of the leaves, flowers, stem, and roots. When considering the plant parts or plant tissues of *Cannabis*, cannabinoids are widely known to be abundant in the flowers (inflorescence) of the plant compared to the other plant tissues (Jin et al., 2020). Consequently, several pharmaceutical companies (supported by numerous scientific studies) exclusively focus on isolated compounds or crude extracts from *Cannabis* flowers (inflorescence) for the development of new drugs due to the high cannabinoid content found in the flowers (Sarma et al., 2020; Balant et al., 2021). However, research shows that the use of leaf or seed in traditional medicine (is often more important than the use of inflorescence for the treatment of certain ailments (Balant et al., 2021). This suggests that *Cannabis* or its other plant tissues can be alternative sources for drug discovery. Therefore, the comprehensive exploration of the full phytochemical space of *Cannabis* and its different plant parts remains important as this can aid in identifying cultivars or specific plant parts that are more useful for specific ailments (Aizpurua-Olaizola et al., 2016; Namdar et al., 2018; Balant et al., 2021).

Nevertheless, the cannabinoid-centred approach continues to dominate the research field, and this has limited the exploration of the full potential of *Cannabis* chemistry. Furthermore, with an increasing attention to *Cannabis*, globally, and a growing momentum for legalization and commercialization of the plant (in some countries this is currently ongoing), there is a need to comprehensively characterize the chemistry of *Cannabis* (and its various plant tissues) beyond cannabinoids to understand its full chemical composition (Lowe et al., 2021; Pattnaik et al., 2022). The increasing scientific efforts to characterize and explore the metabolome of *Cannabis* encompass the utilization of omics sciences, particularly metabolomics, which provides distinctive avenues to investigate the plant's metabolism and unravel the biochemical mechanisms driving the synthesis of various specialized metabolites within its tissues. Furthermore, metabolomics can facilitate and accelerate the search for novel bioactive compounds from plant crude

extracts (Aliferis & Bernard-Perron, 2020; Vásquez-Ocmín et al., 2021; Li et al., 2022).

The application of metabolomics in *Cannabis* research and development (R&D), coined the term “cannabonomics”, is still in its infancy and there is a need for the optimization of bioanalytical protocols, instrument analysis and metabolite databases to improve the robustness of this approach (Aliferis & Bernard-Perron, 2020). Nonetheless, cannabonomics has been suggested for the classification of *Cannabis* into chemical varieties or “chemovars,” to emphasize the unique overall biochemical profile of the *Cannabis* plant of interest (Ladha et al., 2020). Such chemical profiling or metabolic phenotyping can confirm the composition and quality of the chemovar of interest, with variations indicating its medicinal applications. Thus, the study reported herein is a computational metabolomics work to elucidate a *Cannabis* metabolomic atlas of two cultivars for which sufficient seed material was available: Amnesia Haze and Royal Dutch Cheese cultivars (both type I chemovars). We mapped their metabolome content beyond cannabinoids using a suite of computational metabolomics strategies including feature-based molecular networking, substructural discovery method (MS2LDA), in silico tools (e.g., NAP and DEREPLICATOR+), and MolNetEnhancer. The study contributes to ongoing *Cannabis* R&D efforts, particularly the comprehensive characterization of the plant chemistries of these two drug-type cultivars (a first of its kind) that have not been investigated or discussed in literature but form part of the many acclaimed *Cannabis* strains that have been generally hailed for their therapeutic properties against diseases such as cancer. Hence, this study is also part of the ongoing efforts to determine the anti-cancer properties of *Cannabis* chemovars and aims to contribute and build on the evaluation of the bioactivity elicited by the varying chemical composition of *Cannabis* chemovars.

2 Experimental procedure

2.1 Plant cultivation and harvesting

Cannabis seeds, Amnesia haze (a hybrid genetically comprised of 80% *C. sativa* & 20% *C. indica*) and Royal dutch cheese (a hybrid genetically comprised of 70% *C. indica* & 30% *C. sativa*), were purchased from Marijuana SA (Pty) Ltd (Cape Town, South Africa). The seeds were grown into plants in Freedom farms premium classic growing medium (Freedom Farms Horticulture Technologies, Cape Town, South Africa) at 24 °C, 70% humidity and a 24-hour light schedule for germination and vegetative stages. The flowering stage for both cultivars was initiated at 27 °C, 30% humidity and a 12-hour light/12-hour dark cycle.

Seven-week-old plants in their vegetative stages were harvested for leaves per cultivar and 16-week-old plants in their flowering stages were harvested for flowers. The harvested plant-tissues were freeze-dried and crushed with a blender to powder form and stored at room temperature until metabolite extraction.

2.2 Metabolite extraction and standard preparation

Fifty milligrams (50 mg) of the plant materials were weighed and extracted in 1 mL of 80% methanol (Chemlab, UK). The samples (leaves and flowers per respective cultivar) were spun overnight in a digital rotisserie tube rotator at 70 rpm. The crude extracts were then centrifuged at $5678\times g$ in a benchtop fixed-angle centrifuge (Thermo Fisher, Johannesburg, South Africa). After centrifugation, the supernatants were filtered using 0.22 µm nylon filters into glass vials with 500 µL inserts. In this study, seven independent replicates for each sample group were weighed and prepared. The prepared samples were then stored at 4 °C until analysis. For chemical baiting, the vitexin standard was prepared in 50% methanol to a concentration of 5 ppm. The standard was treated in the same manner as the samples for experimental procedures.

2.3 Data acquisition – LC-MS/MS analyses of crude extracts

Samples (methanol extracts) were analyzed on a liquid chromatography–quadrupole time-of-flight mass spectrometry (LC-qToF-MS) instrument (LCMS-9030, Shimadzu Corporation, Kyoto, Japan). For chromatographic separation, a sample volume of 3 µL was injected on a Shim-pack C18 column (100 mm × 2.1 mm, 2.7 µm) (Shimadzu Corporation, Kyoto, Japan) thermostatted at 55 °C. In addition to the stationary reverse phase, the chromatography was carried out with a binary mobile phase, applying a gradient elution method. The solvent system comprised solvent A consisting of 0.1% formic acid in Milli-Q water (both HPLC grade, Merck, Darmstadt, Germany) and solvent B being methanol (UHPLC grade, Romil SpS, Cambridge, UK) with 0.1% formic acid, with a flow rate of 0.4 mL/min. The gradient elution was performed as follows, B referring to organic composition (i.e., solvent B): 5% B for 3 min, 5–40% B over 3–5 min, 40–95% B over 5–12 min, then 95% B for from 12 min to 18 min. The gradient was changed back to initial 5% B at 18 min and kept to 20 min. The column was re-equilibrated for 3 min.

The effluents from chromatographic separation were further analyzed with the high-definition mass spectrometer, equipped with electrospray ionization (ESI), acquiring both negative and positive spectral data. The mass spectrometer

parameters used were the following: 4.5 kV interface voltage, with interface temperature of 300 °C; 3 L/min flow rate for nebulization and dry gas; DL temperature of 250 °C, and 400 °C for heat block; detector was operated at 1.8 kV voltage. For monitoring the accuracy of acquired mass-to-charge ratio (m/z), sodium iodide (NaI) was used as a calibration solution. Both non-fragmented (MS1) and fragmented (MS2) spectral data were acquired with m/z range of 100–1200 Da. For MS/MS experiments, data-dependent acquisition method was applied, with an intensity threshold of 5000 counts. The collision-induced dissociation (CID) method was applied for the fragmentation of ions, using argon as a collision gas at a collision energy of 30 eV with a spread of 5 eV.

2.4 Data mining: processing and chemometrics analyses

The acquired raw datasets (ESI negative and positive centroid data) were processed using MetaboAnalyst version 5.0 (Pang et al., 2021). The data was processed with UPLC-Q/TOF default parameters using the centWave method. From the default setting, the max peak width was changed to 15. The resulting feature tables with 3560 and 3570 features for (ESI) negative and positive datasets respectively were imported into soft independent modelling of class analogy (SIMCA) version 17.0 software (Sartorius, South Africa) for chemometrics modelling. Principal component analysis (PCA) and hierarchical cluster analysis (HCA) models were computed for data exploration, and group classification and discriminant analysis.

2.5 Molecular networking and metabolite annotation

Prior to computing the feature-based molecular networks (FBMN) through the Global Natural Product Social (GNPS), the acquired raw datasets obtained from the Shimadzu LCMS-9030-qTOF-MS were converted to open-source format (.mzML) files. The mzML files were then uploaded to Mass Spectrometry-Data Independent Analysis (MS-DIAL) platform for data processing. The parameters used for MS-DIAL data processing parameters included mass accuracy MS1 and MS2 tolerance of 0.25 Da and 0.1 Da respectively, with the MS/MS range of 50–800 Da; the minimum peak height was 2000 amplitude, mass slice width of 0.1 Da for peak detection, a sigma window value of 0.5 was used, and retention time tolerance was 0.1 min and MS1 tolerance set at 0.015.

Post MS-DIAL data processing, the GNPS export files, both GNPS MGF files and feature quantification tables, were exported into the GNPS ecosystem using the WinSCP

server for molecular networking. FBMNs were then computed and generated using the GNPS molecular networking workflow. The parameters used for computing FBMN for the leaves and flower spectral datasets included precursor ion mass tolerance of 0.25 Da with fragment ion mass tolerance of 0.25 Da. For spectral similarity, a cosine score cut off was 0.65, with a minimum of 4 matched fragment ions. For annotation, various spectral libraries were searched and these included MassBank, ReSpec and NIST. To visualize and analyze the computed networks, the Cytoscape network visualization software (version 3.8.2) (Shannon et al., 2003; Smoot et al., 2011) was used.

All semi-annotated and some unmatched/unidentified nodes, visualized in Cytoscape, were verified using their empirical formulae calculated from accurate mass and fragmentation patterns. In addition to the manual inspection of annotations, some natural products dereplication databases such PubChem (<https://pubchem.ncbi.nlm.nih.gov>) and Dictionary of Natural Products (<http://dnp.chemnetbase.com/faces/chemical/ChemicalSearch.xhtml>) were also used. The annotations were also verified and confirmed against available literature. All metabolite annotations were carried out at level 2 and 3 of the Metabolomics Standards Initiative (MSI) (Sumner et al., 2007). All validated annotations (manual and computational) are listed in supplementary Table S1. The networks generated were further explored using network annotation propagation (NAP), where first in silico fragmentation was applied to the generated FBMN where structural searches were performed in databases e.g., GNPS, CHEBI, SUPNAT, and DRUGBANK.

To perform network annotation propagation (NAP) (da Silva et al., 2018) jobs, both the fusion and consensus scores were used based on the first 10 structural candidates. Using scoring methods, NAP re-ranks the candidate structure list based on the network topology (Ernst et al., 2019; Kang et al., 2019; Nepthali et al., 2022). The two scoring methods, as previously mentioned, utilized by NAP include (a) fusion scoring, which uses MetFrag in silico prediction with MetFusion based on spectral library matches within a molecular family; and (b) consensus scoring, which uses the structural similarity from in silico candidates across the spectral nodes of a molecular family. The FBMNs were also explored using DEREPLICATOR (Mohimani et al., 2018) for peptidic structural annotations.

Substructure annotation and initial exploration was carried out using the MS2LDA (MS2 latent Dirichlet allocation) tool (van der Hooft et al., 2016; Rogers et al., 2019) in GNPS and substructure annotations from MotifDB were included in the analyses. From the default settings, some of the parameters were changed as follows for the ESI negative dataset: Bin width was set at 0.01 (Tof data) and LDA free motifs set at 150. The MotifDB included was Rhamnaceae

Plant and the rest were excluded. For the ESI positive dataset, the parameters were changed as follows: LDA free motifs set at 200. The MotifDB included were GNPS, MassBank and Euphorbia.

Finally, MolNetEnhancer (Ernst et al., 2019) which incorporated outputs from both FBMN and in silico tools such as MS2LDA, DEREPLICATOR, NAP, and the automated chemical classification through ClassyFire, was applied, providing thus a holistic chemical overview of measured metabolomics spectral data, with enhanced structural details for each fragmentation spectrum. The GNPS job links are provided in the Supplementary materials.

2.6 Pathways analysis and relative quantification

The metabolomic pathway analysis (MetPA) tool in MetaboAnalyst, version 5 (Pang et al., 2021) was used to perform functional analysis, particularly pathway analysis. The latter was computed using identifiers of the annotated metabolites, KEGG IDs, as input data. For overrepresentation analysis, the enrichment method, hypergeometric test was used; and for the node importance measure, i.e., topological analysis, the relative betweenness-centrality was employed. For visualization was done using scatterplots and the pathway library used was *Arabidopsis thaliana* (thale cress) (KEGG). Using integrated peak areas, (relative) quantitative analysis was done, and colour-coded heatmap generated in MetaboAnalyst. Pareto-scaling and log-transformation were applied as data pre-treatment methods.

2.7 Cell culture

The MIA PaCa-2 pancreatic cancer and HEK 293 human embryonic kidney were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma Aldrich, USA) supplemented with 10% fetal bovine serum (FBS) (Biogen, UK) and 1% Pen-Strep (penicillin-streptomycin) (Biowhitaker, Germany) at 37 °C in 95% humidity and 5% CO₂. About 75% of confluent cells were harvested and rinsed with 6 ml Dulbecco's Phosphate-buffered Saline (DBPS) (Sigma Aldrich, USA). The cells were then trypsinized with 2 ml of 1 X trypsin and were incubated for 45 s (HEK 293) and 90 s (PaCa-2) at 37 °C in 95% humidity and 5% CO₂. Four millilitres of fresh DMEM were added to both cell lines respectively to stop trypsinization. Into a new 50 ml centrifuge tube, 6 ml of cell solution was transferred into the centrifuge tube, and the cell solution was centrifuged for 4 min at 2200 rpm (779 × g). After centrifugation, the pellet was resuspended in 3 ml of DMEM for cell quantification and further subculturing.

After cell culturing, cell quantification was done using trypan blue and TC20 automated cell counter (Bio-Rad). In

a 1.5 ml microcentrifuge tube, 10 µl of cells in DMEM was gently mixed with 10 µl of trypan blue and was followed by aliquoting 10 µl of the mixture into the cell counter slide compatible with the TC20. Cell viability and cell concentration were read and cells that exhibited 95% viability were used for cytotoxicity and caspase activity assays.

2.8 Extract preparation

The 80% methanol flower extracts of AMNH and RDC were prepared using the protocol reported in Sect. 2.2. The 80% methanol plant extracts were then dried into solid using the rotary evaporator and the two extracts (AMNHF and RDCF) were stored at room temperature until further analysis. The dried extracts were made up to a stock concentration of 100 mg/ml w/v in 100% dimethyl sulphoxide (DMSO) (PanReac AppliChem, Germany) respectively. The extract stock solutions were stored in the dark at 4 °C awaiting further analysis.

2.9 Alamar blue cellular viability assay

The HEK 293 and MIA PaCa-2 were plated in 96 well plates (at 5000 cells per well). The cells were treated with AMNHF and RDCF and were incubated for 24 h with an untreated control, the negative control set as 0.1% DMSO and the positive control set as 100 µM etoposide (Sigma Aldrich, Germany), and a media blank column (plated per column). The AMNHF and RDCF extracts were used to treat the cells in a concentration-based series starting from 100 µg/ml and serially diluting it 2X to 3.125 µg/ml. A volume of 10 µl of the Alamar blue (ThermoFisher, USA) reagent was added to each well following the 24 h treatment and the plate was incubated for 2 h at 37 °C. The fluorescence was measured in each well using a plate reader and the Gen 5 program 530/25 nm excitation and 590/35 nm emission. The analysis was repeated three times.

$$\text{Cell viability \%} = \frac{\text{Fluorescence of treated} - \text{fluorescence of blank}}{\text{Fluorescence of untreated} - \text{fluorescence of blank}} \times 100$$

2.10 10 caspase Glo® 3/7 activity assay

MIA PaCa-2 and HEK 293 cells were plated and treated with AMNHF and RDCF in 33 mm petri dishes in a similar manner as described in Sect. 2.9 with the IC₅₀s and controls. The MIA PaCa-2 cell line was subjected to a caspase

detection assay to determine if caspase-3 and -7 were activated due to the compound treatment. The cells were plated for 24 h in 33 mm petri dishes at 1×10^5 cells/ml. The cells were treated 24 h after plating with 2 ml treatments consisting of an untreated control, the positive and negative controls, and the AMNHF and RDCF treatments with the IC₅₀s of each extract. The assay was repeated three times and was conducted using the Caspase Glo® 3/7 assay kit (Promega, USA) based on the manufacturer's recommendations.

2.11 In silico molecular docking

The PDB files of the target proteins were obtained from the Protein Data Bank (<https://www.rcsb.org/>) database and selected based on their resolution: $<2.0\text{\AA}$ ($1.5\text{--}2.0\text{\AA}$), while the structures of the ligands were retrieved from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>). The files were then prepared in UCSF Chimera for docking by deleting and adding hydrogens to the receptors and ligands respectively. The structures were respectively saved in .pdb and .mol2 formats and converted into rec.pdb and .pdbqt formats using Autodock tools. After predicting the active sites of the receptors, a grid box was created to surround the binding residues present in the site using specific X, Y, and Z dimensions and centres. Setting the results to give five prediction outputs, Raccoon and Autodock Graphical user interface supplied by MGL tools were then used to dock the ligands onto the receptors. The complexes exhibiting the lowest Z-scores i.e., the least energy functions were then selected. The resulting .pdb files were viewed in UCSF Chimera, which was used to identify and label the residues involved in the docking.

3 Results and discussion

3.1 Chemometric analysis of *Cannabis* cultivars

As detailed in the experimental section, a nontargeted LC-MS/MS-based metabolomics approach was applied for profiling the metabolome of leaves and flowers of two cultivars that are used for medicinal purposes: Amnesia haze (*C. sativa* dominant) and Royal dutch cheese (*C. indica* dominant). The methanol extracts from leaf and flower tissues were analyzed on an LC-MS/MS platform. The processed spectral data was explored and evaluated by applying unsupervised chemometrics methods, namely, principal component analysis (PCA) and hierarchical clustering analysis (HCA) modelling. The generated models revealed certain structures within the data, such as tissues- and cultivar-related sample groupings, pointing to underlying differential

metabolite profiles which were further investigated (Figure S1).

3.2 Global chemical profiling – key constituents of *Cannabis* metabolome

The chemical profiling of plant extracts such as *Cannabis*, achieved through GNPS-based molecular networking (MN) tools, can assist in bettering and laying the foundation for the differentiation of *Cannabis* cultivars into chemovars based on their chemical class content. In doing so, metabolite biomarkers can be identified for each chemovar, and this would aid in understanding the non-generic medicinal or bio-active potential elicited by each *Cannabis* chemovar or plant-tissue thus narrowing down the medicinal applications of the chemovar of interest. Thus, in this study, the metabolic inventory of the two cultivars was investigated. The obtained spectral data were submitted to mass spectral molecular networking (MN) through the GNPS ecosystem, a scalable workflow that digitalizes the diversity and distribution of metabolites in plants (Kang et al., 2019).

Feature-based molecular networking (FBMN) enabled the semi-automated putative annotation of metabolites with confidence level 2 (or level 3) annotations as defined in the proposed minimum reporting standards of the metabolomics standards initiatives (MSI) (Sumner et al., 2007) (Fig. 1). FBMN represents a computational strategy that facilitates the visualization and compound annotation of complex, high-resolution untargeted LC-MS/MS metabolite data from natural extracts (Hammerle et al., 2021). This type of MN is designed to distinguish structural isomers by incorporating features such as chromatographic retention times which enhances metabolite annotation and thus the dereplication of metabolites whilst also retaining semi-quantitative information to perform statistical analyses (Quinn et al., 2017; Nothias et al., 2018, 2020). The computed FBMN of spectral data from leaf samples (from both cultivars) comprised 2914 nodes (metabolite features with an MS/MS spectrum) where 997 nodes clustered into 128 molecular families (MF) based on spectral similarity while 1917 nodes formed singletons (stand-alone nodes that are not incorporated into a MF) (Fig. 1A). Searching spectral libraries in GNPS, 166 nodes (from the 2914 nodes) were putatively annotated, with varying relative distributions across the cultivars, which point to diverse chemistries in the leaves of these *Cannabis* cultivars. For the samples from the flower tissues, the computed FBMN was of 2829 nodes of which 119 nodes were putatively annotated and 998 nodes clustered to form 172 molecular families while 1831 formed singletons (Fig. 1B). The singletons (selfloop nodes) positioned at the bottom of both networks (Fig. 1A-B) represented spectra that were not

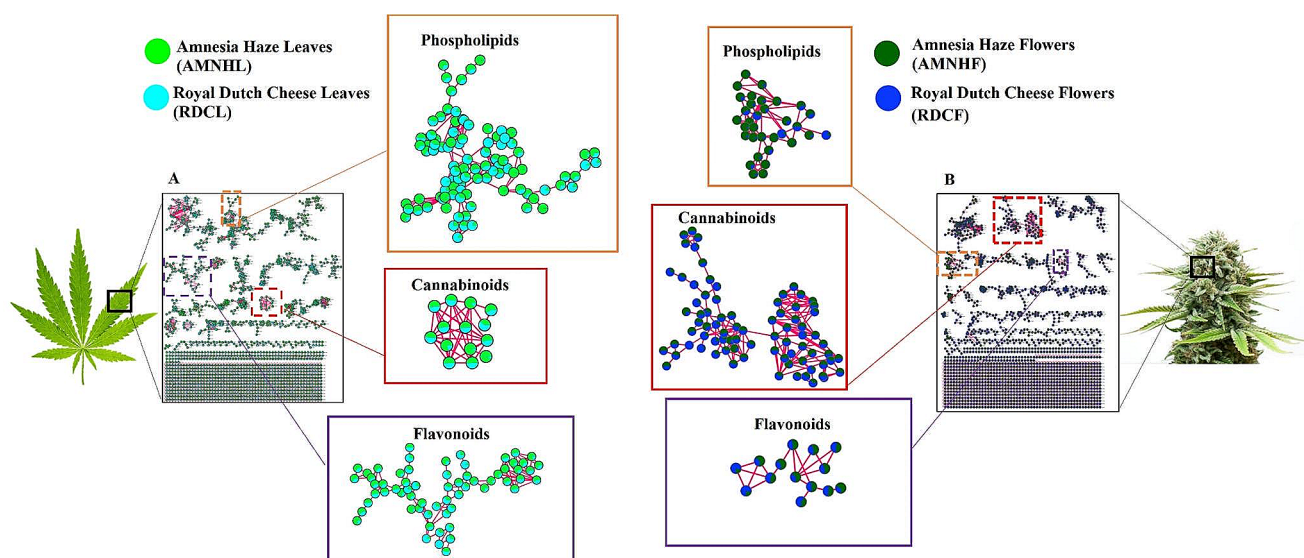


Fig. 1 Feature-based molecular networking (FBMN): Molecular networks representing the general mass spectrometry detected chemical space of *Cannabis* cultivars Amnesia haze (AMNH) and Royal dutch cheese (RDC). (A) mass spectral molecular network for the leaves, L denoting leaves and (B) mass spectral network for the flowers, F denoting flowers. Key: Amnesia haze leaves (AMNHL), Amnesia haze flowers (AMNHF), Royal dutch cheese leaves (RDCL) and Royal dutch cheese flowers (RDCF). Some major molecular families from each network are highlighted, e.g., flavonoids, cannabinoids and phospholipid Clusters.

clustered into molecular families, i.e., low spectral similarity scores.

The molecular families (MFs) in the networks provided a global visualization and insights on the metabolome of the two cultivars, revealing diverse metabolite classes including phospholipids, cannabinoids and flavonoids, which are differentially distributed across the cultivars and across the tissues (Fig. 1 and Figure S2). From the 128 MFs found in the leaf data (where two or more connected nodes are considered as an MF), 94 MFs were comprised of unknown metabolites. However, 34 MFs had at least one metabolite automatically matched to the GNPS libraries, a similar trend that was also observed in the MFs found in the flower MN. Considering that MN is rationally constructed based on spectral similarities (Aron et al., 2020; Vincenti et al., 2020; Yu et al., 2022; Zhang et al., 2023), automated annotation of one node in a cluster can aid in decoding and annotating other structurally similar metabolites or features in the same molecular family. As such, MN improves on metabolite annotation, decoding ‘dark matter’ in spectral data, which subsequently provides an improved coverage on the annotated metabolome. In this study, this is illustrated by Fig. 2A where one spectral node in the MF was putatively annotated through GNPS spectral library matching as CBD. Based on the concept of molecular networking, theoretically this cluster posed as a cannabinoid cluster formed by metabolites with similar structures and fragmentation patterns. This then propagated the manual annotation of unidentified metabolites such as cannabidiolic and cannabidivarinic acids in the cluster (Fig. 2A). The unannotated metabolites in this

cluster also suggest that there may be other unknown CBD related compounds or analogues present.

Furthermore, as may be expected from plant samples, flavonoids were also one of the major metabolite classes found in the *Cannabis* tissues. In this flavonoid cluster, there are spectral nodes (such as m/z 489.10, 505.28, 479.15, 639.23 and 669.31) that could not be annotated through GNPS spectral library matching, yet they are postulated to be structurally related to glycosylated flavonoids (Table S1) due to their grouping with flavonoids such as isoquercitrin (m/z 433.07) and isorhamnetin-3-glucoside (m/z 477.09) (Fig. 2B) (Table S1).

As above mentioned, it is worth noting that applying these computational tools improves annotation as it provides insights into possible molecular families in the measured metabolome. However, there are some limitations, related to similarity scoring algorithms and the complexity of untargeted metabolomics spectral data. To improve on this, the “chemical baiting” concept was explored and applied. The baiting approach is a concept that is predominantly used in protein biochemistry where chemical baits or probes are used to bind or concentrate a diverse range of biomarkers (proteins & peptides, metabolites, lipids & fatty acids, nucleic acids, and post translationally modified peptides) that are either occurring in low concentrations or masked by other dominant resident proteins (Luchini et al., 2010). Nonetheless, within metabolomics approaches based on molecular networking (MN), a known exogenous compound (a bait) can be employed to screen for related endogenous metabolites residing within the same molecular

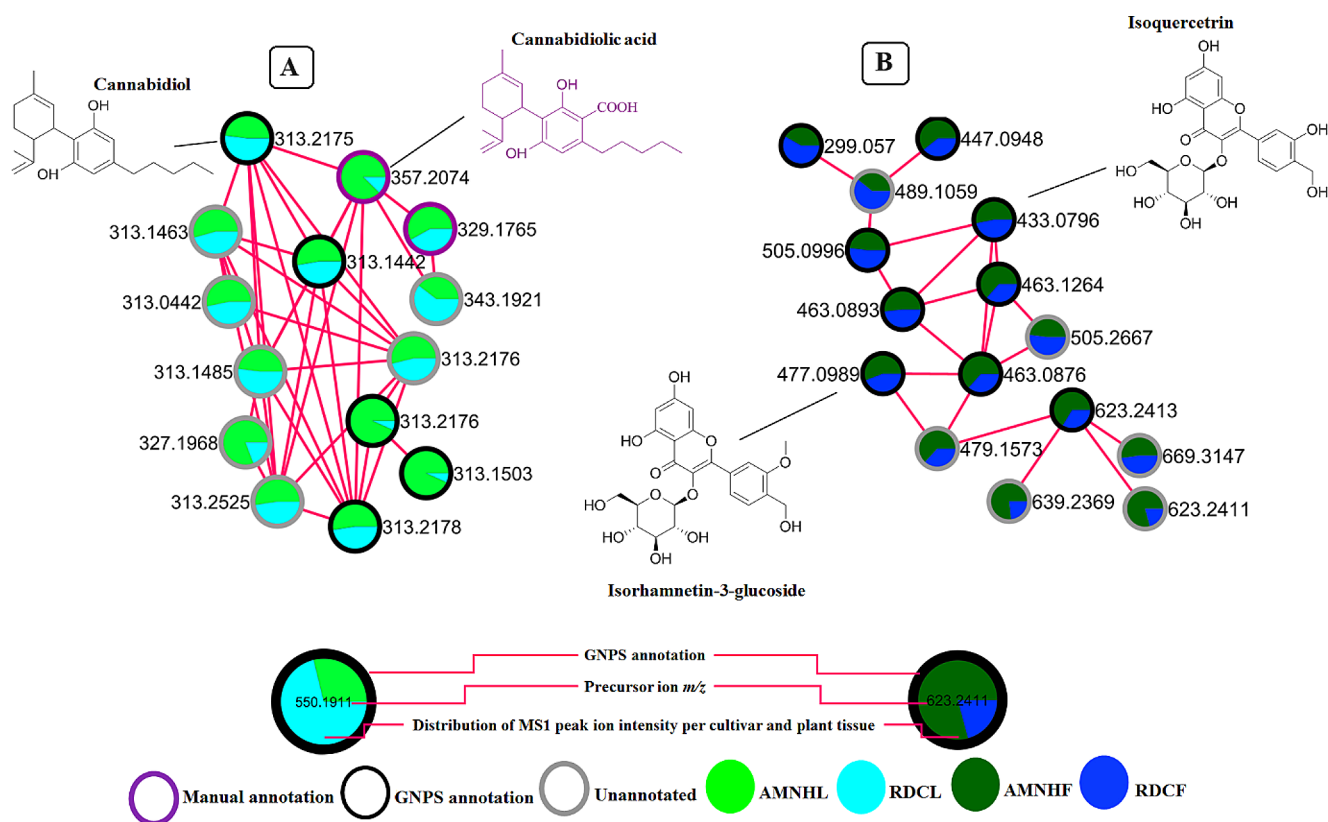


Fig. 2 Molecular clusters extracted from FBMN. A detailed visual representation of the output of feature-based molecular networking highlighting some of the major chemical classes in leaves /flowers of Amnesia haze and Royal dutch cheese. *Key:* Amnesia haze leaves (AMNHL), Amnesia haze flowers (AMNHF), Royal dutch cheese leaves (RDCL) and Royal dutch cheese flowers (RDCF). **(A)** Cannabinoid cluster in the leaves consisting of cannabidiolic (m/z 357.20) and cannabidivarinic acid (m/z 329.17) **(B)** flavonoid cluster in the flowers consisting of Isoquercitrin (m/z 463.12) and isorhamnetin-3-glucoside (m/z 477.09). In addition to the annotation of the metabolites in the clusters, FBMN also aids in visualizing the relative abundance of each metabolite in the extracts

network. Thus, in this study, the vitexin standard was used as a probe to determine the effect of chemical baiting on FBMN (Fig. 3).

The overall observed effect of chemical baiting on FBMN (Fig. 3) showed that using a known reference standard, in this case the vitexin standard, can aid in probing or fishing out known, “unknown knowns” and “unknown unknown” that are structurally similar and thus have similar fragmentation patterns depending on one or more standards used. As such, the use of molecular baits would speed up the annotation process, increasing the accuracy in forming molecular families or clusters (based on structural similarities) in the network. This was shown by the clustering of vitexin and vitexin-2-O-rhamnoside into one molecular family when the vitexin standard was added as a chemical bait.

We also note that metabolite features annotated with cannflavins A and B, flavonoids unique to *Cannabis* cultivars and structurally related to vitexin, were also initially detected in the same MF as vitexin-2-O-rhamnoside before the addition of the vitexin standard (a molecular bait). However, once the vitexin standard was added, the cannflavins

formed their own MF (Fig. 4). It was previously demonstrated that cannflavins (cannflavin A, B and C) are geranylated flavones postulated to be unique to *Cannabis* cultivars (Rea et al., 2019; Bautista et al., 2021; Tomko et al., 2022).

In our study, cannflavins A and B were detected in both the leaves and flowers of AMNH and RDC and reported in Table S1. Cannflavin A (m/z 435.1825) was distributed evenly across the plant tissues of both cultivars, and cannflavin B (m/z 367.1193) was recognisably present in the leaves and flowers of RDC but significantly abundant in the flowers of AMNH (Fig. 4). Using the GNPS analogues library search, the unknown metabolites in the cannflavin cluster were matched to flavone analogs such as afrormosin and 3',3',4'-trimethoxyflavone which are structurally similar to the detected cannflavins.

Thus, the computed FBMN (Fig. 3) provided spectral clustering and putative annotations of the measured metabolomes from the leaves of the cultivars, revealing actionable insights and key characteristic features of the chemical space of *Cannabis* such as 289 spectral library matches of flavonoids, cannabinoids, phospholipids and more (Table

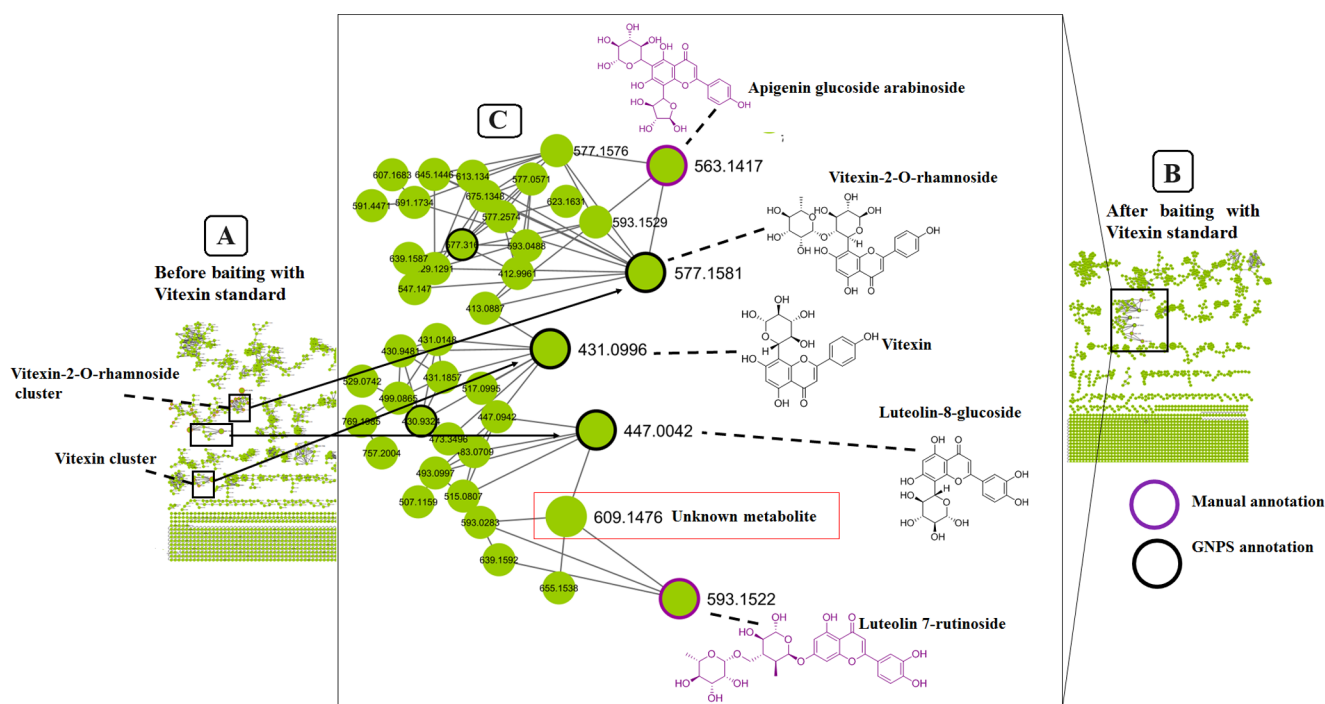


Fig. 3 Chemical baiting in FBMN. The effect of chemical baiting on FBMN using the Vitexin standard as a probe for Vitexin-2-O-rhamnoside and any other metabolites that are structurally similar to vitexin. (A) FBMN before analyzing samples in the presence of the standard, (B) FBMN after analyzing samples with the standard and (C) zoomed-in vitexin cluster from (B)- showing the clustering of the vitexin standard together with vitexin-2-O-rhamnoside, as well as the incorporation of an unknown metabolite and other structurally similar metabolites into one molecular family.

S1), 1265 connected nodes and 192 molecular families. Furthermore, chemical baiting assisted, for instance, in more confident annotation of vitexin-related structures. Moreover, the generated network showed differential distribution of metabolites across the cultivars. For instance, clusters of lipid-like molecules were more abundant in the plant tissues of RDC and flavonoids such as isoquercitrin were more abundant in the tissues of AMNH, while a cluster of hydroxy acids such as citric acid occurred in similar levels across both the cultivars (Figure S2).

To further explore and characterize the metabolome of the *Cannabis* cultivars, in silico annotation tools such as substructure recognition topic modelling through the MS2 Latent Dirichlet Allocation (MS2LDA), and the network annotation propagation (NAP) were performed for all spectral datasets from both cultivars, AMNH and RDC. To illustrate the contribution of these tools to the *Cannabis* metabolite annotation and identification at a scaffold diversity level, the outputs obtained from both AMNH and RDC dataset are reported herein (Fig. 5 and Table S1). The machine learning (ML)-based tool, MS2LDA, allows an unsupervised decomposition of fragment spectra, discovering mass fragmental patterns of co-occurring mass fragments as well as neutral losses (termed ‘Mass2Motifs’, or shortly ‘m2m’) from different MS/MS spectra, which enables extracting information on substructural diversity

within each class of metabolites (van der Hooft et al., 2016). During MS2LDA analyses, sets of previously annotated Mass2Motifs can be added from MotifDB (Rogers et al., 2019), whilst unknown chemistry is captured in Mass2Motifs that are numbered to enable highlighting and further studying them during subsequent analysis (van der Hooft et al., 2020). As such, the generated substructural information highlights the (bio)chemical relationships of the connected compounds. Moreover, these (shared) substructures or scaffolds can point to functional groups and/or core structures of the compounds, potentially revealing common biosynthetic routes in a molecular family (van der Hooft et al., 2016; Beniddir et al., 2021; Ramabulana et al., 2021).

In total, one hundred and eighty-one (181) Mass2Motifs were discovered, and most were common in both the *Cannabis* leaves and flowers datasets (ESI negative) of the two cultivars, respectively. The Mass2Motifs, annotated and unannotated were visualized through the MS2LDA website (ms2lda.org) (Wandy et al., 2018). As an illustration, some of the characterized Mass2Motifs found in the *Cannabis* leaves are highlighted and shown in Fig. 5. These discovered substructural Mass2Motifs are significant in the sense that they can aid in the structural elucidation of “unknown knowns” (metabolites or spectral nodes whose reference spectra are not found in the GNPS spectral libraries and therefore represented as

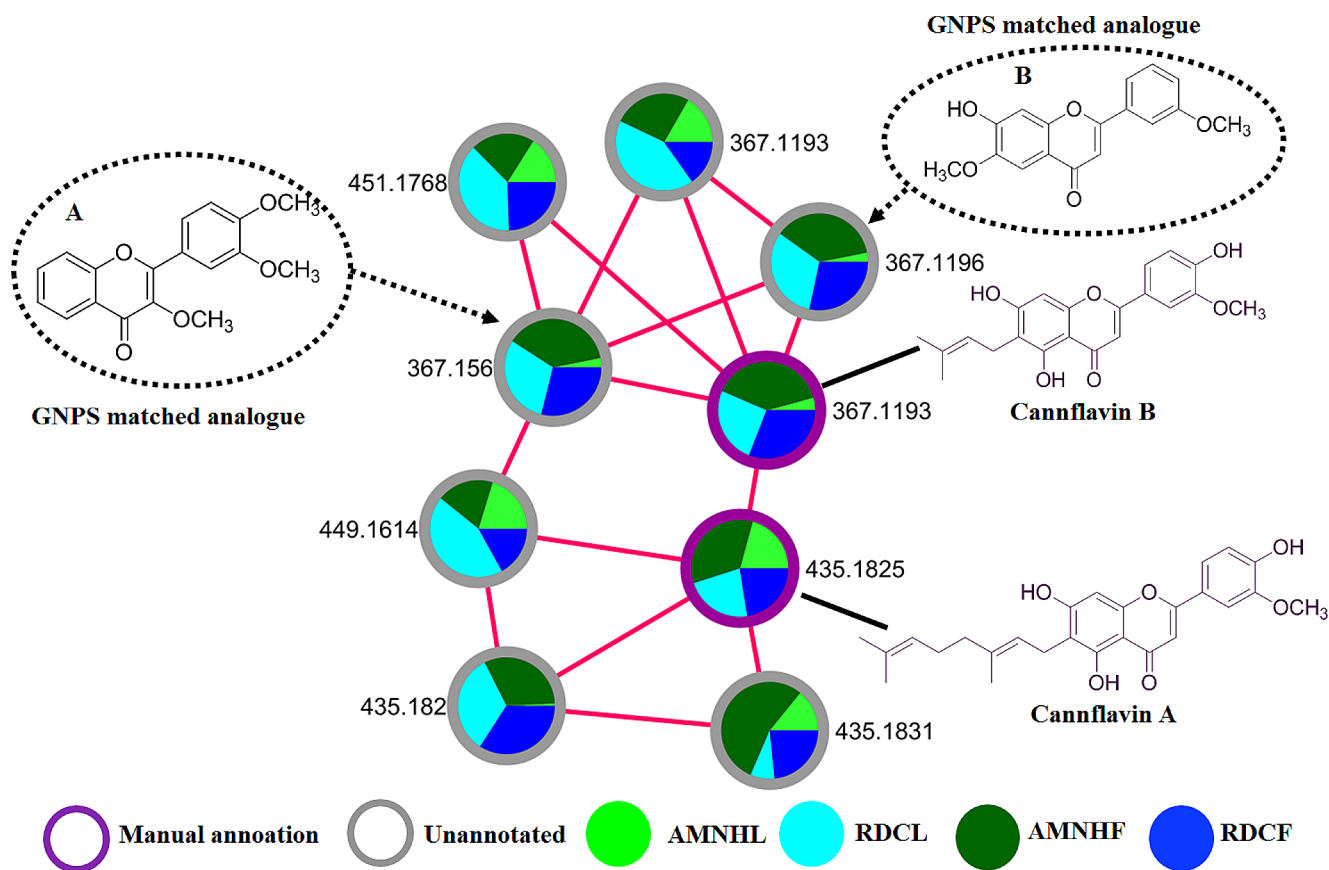


Fig. 4 Cannflavin cluster. A representation of the distribution of the detected cannflavins (cannflavin A and B) and some of the GNPS matched analogues acquired in ESI negative. *Key:* Amnesia haze leaves (AMNHL), Amnesia haze flowers (AMNHF), Royal dutch cheese leaves (RDCL) and Royal dutch cheese flowers (RDCF). **(A)** Unknown compound (m/z 367.156) **(B)** unknown compound (m/z 367.1193) matched through GNPS analogue library as afformosin and '3,3',4'-trimethoxyflavone respectively.

unknowns) such as m/z 609.1476 identified as luteolin-glucoside-arabinoside. When evaluating the chemical makeup of the glucosylated flavonoids, Fig. 5A illustrates that m/z 609.1476 shares a common substructure (unknown motif_61) with the manually annotated metabolites luteolin-glucoside and luteolin-rutinoside. This resulted in the structural elucidation of luteolin-glucoside-arabinoside (m/z 609.1476) and led to the shared unknown motif_61 being annotated as a substructure that is related to luteolin (m/z 285). Additionally, this also suggests that motif_107, motif_48, and motif_45 are all related to some sugar loss, respectively.

Moreover, it can be postulated that motif_76 and motif_42 are both related to the apigenin scaffold as they both form the core structure of the putative annotations of vitexin (m/z 431.0996) and vicenin II (m/z 593.1529) where motif_76 is found in compounds related to vitexin while motif_42 is found in compounds related to vicenin II. This implies that the mass fragmental pattern captured in motif_76 is related to apigenin with mass features specific for vitexin, whereas the pattern captured in motif_42 is

related to apigenin with mass features specific for vicenin. Furthermore, MS2LDA also highlighted some of the chemical moieties that make up cannabinoids as shown in Fig. 5B. It is worth noting that the fragmentation patterns (in mass spectrometry) of C-linked flavonoids, such as vitexin display key structural information. The latter can aid in differentiating flavonoids by examining first-order ion spectra or low-energy CID spectra. However, a detailed account on this (C-linked and O-linked glucosides) is beyond the scope of this study, and the reader is referred to the literature (Ma et al., 2000; Cuyckens & Claeys, 2004; Cao et al., 2014; Geng et al., 2016). It suffices here to mention that O-glycosides, linked through an oxygen atom, frequently undergo glycosidic cleavages, producing a variety of sugar moiety fragments and prominent Y ions formed by rearrangement reactions at the interglycosidic bonds. On the other hand, C-glycosides (such as vitexin), with a more stable carbon-carbon linkage, mostly molecular ions are observed together with cross-ring cleavages of the saccharidic residue and the loss of molecules of water, i.e., mostly X ions, in the

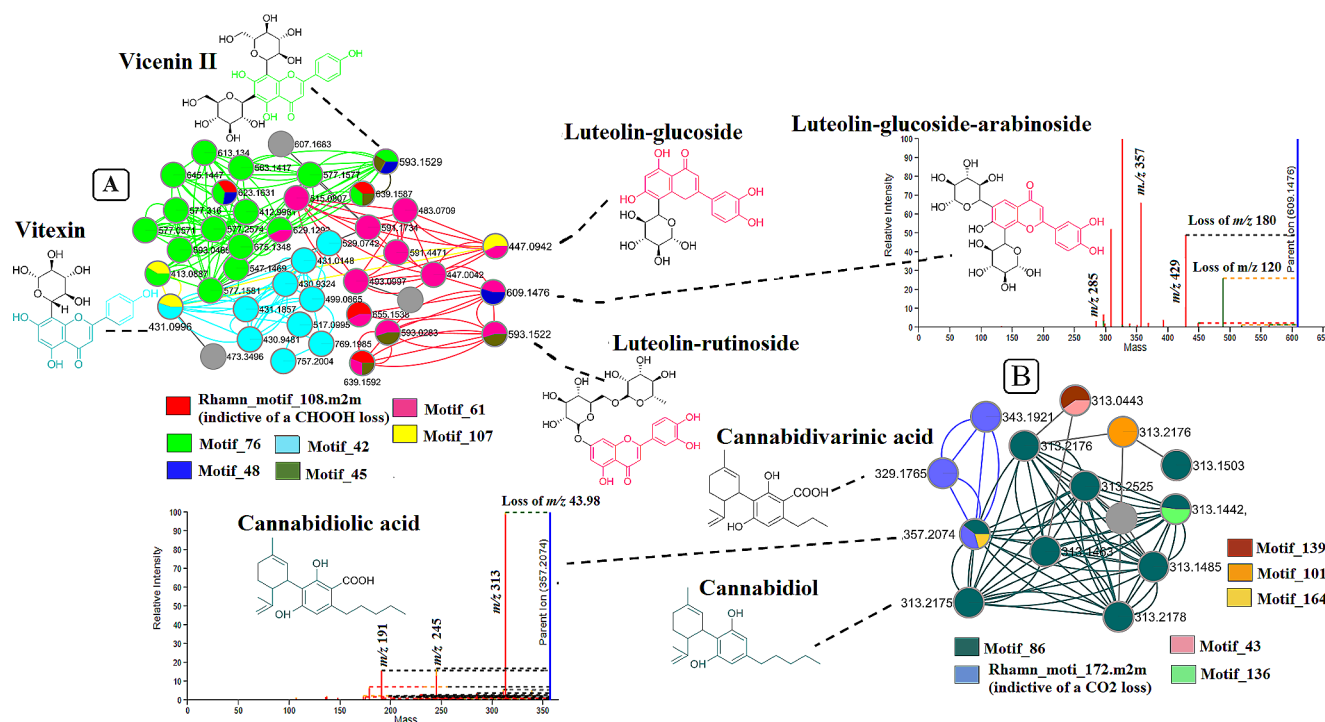


Fig. 5 MS2LDA output. Molecular network clusters extracted from an MS2LDA substructure exploration analysis of *Cannabis* cultivars AMNH and RDC in ESI negative ionization mode. (A) and (B) illustrate the MS2LDA-driven annotations of flavonoids (A) and cannabinoids (B) in the leaves of the cultivars. The coloured nodes represent some of the recognized substructures.

nomenclature model by Domon and Costello (Cuyckens & Claeys, 2004).

Furthermore, in addition to flavonoids, compounds such as cannabidiol (CBD) were annotated through GNPS spectral library matching and with the use of FBMN, it was shown to be structurally similar to cannabidiolic acid (CBDA - m/z 357.2074) and cannabidivarinic acid (CBDVA - m/z 329.1765) (Fig. 2). The MS2LDA discovered substructures of these clustered metabolite features clearly describe the chemical relationship between the compounds through rhamn_motif_172.m2m which describes the carboxylic acid that is attached to CBDA and CBDVA (Fig. 5B). Furthermore, this led to the shared unknown motif_86 being annotated as a substructure that is related to cannabidiol (m/z 313.2175). Combining Mass2Motif annotations and library matches, we can observe how the known compounds and unknown cannabidiol analogues are associated with the cannabidiol substructure. This makes sense as the cannabidiol derivatives would have a cannabidiol core structure. Altogether, the here highlighted Mass2Motifs are indicative of flavonoids and cannabinoids, respectively, and their decorations (i.e., glycosylations). Consequently, this confirms and gives confidence to the FBMN annotations such as those previously made on the highlighted molecular family (Fig. 2) and listed in Table S1.

In addition to MS2LDA, Network Annotation Propagation (NAP) was also applied to facilitate metabolite feature annotation and to serve as an input for chemical compound classification performed by MolNetEnhancer. NAP performs in silico fragmentation-based metabolite predictions or annotations of candidate structures through searching several databases including GNPS, ChEBI (Chemical Entities of Biological Interest), SUPER NATURAL (II) and DNP. Therefore, using these two scoring methods, NAP is a beneficial tool when experimental spectra are matched to a few spectral libraries matches as it also allows the propagation of annotations in the absence and presence of library spectral matches (da Silva et al., 2018; Nepthali et al., 2022; Kim et al., 2023). The outputs of NAP obtained in this study are reported in the supplementary material. In brief, when looking at the annotation of individual features, we could verify that NAP was able to correctly predict 18 (out of 997 spectral features) metabolites for the leaves data and 25 (out of 998) for the flower data. In both leave and flower extracts, the NAP predicted metabolites include, cannabidiol, vitexin, orientin, and luteolin. These metabolite annotations were verified through manual inspection of fragmentation patterns and literature. Whilst other NAP annotations could not be verified as being exact matches, many are likely representing structurally related molecules present in compound databases, and all

structural annotations were used by MolNetEnhancer to infer chemical compound classes of the molecular families in the generated networks.

To get an improved chemical compound class annotation and a comprehensive chemical overview of the measured spectral data, the outputs from FBMN, MS2LDA, NAP and DEREPLICATOR + were combined in an enhanced molecular network workflow, MolNetEnhancer, which integrates also the automated chemical classification of molecular families with ClassyFire (Ernst et al., 2019). MolNetEnhancer enabled the chemical annotation, visualization, and discovery of the subtle substructural diversity within molecular families (Fig. 6).

The generated networks indicated the varying distribution of major superclasses such as phenylpropanoids & polyketides, lipids and lipid-like molecules, organic oxygen compounds, benzonoids and organoheterocyclic compounds across the leaves (Fig. 6A) and flowers (Fig. 6B). A major difference observed between the plant-tissues of the cultivars (Fig. 6) was the abundance of phenylpropanoids and polyketides superclass, which was the second most dominating superclass in the leaves of the cultivars (Fig. 6A). Phenylpropanoids and polyketides include flavonoids, lignins, phenolic acids, stilbenes and coumarins. A considerable number of flavonoids, a chemical class whose biosynthetic pathway was identified as one of the significant metabolic pathways in the two cultivars (Table S2 and Figure S3), were identified in the leaves. These

metabolites, in addition to their plant functions, exhibit biological activities such anti-inflammatory and antioxidant properties.

Other polyphenolic compounds such as feruloyl quinic acid and caffeoylquinic acid reported in Table S1 were also detected in the leaves. These phenolic acids which form part of chlorogenic acids are known to possess anticarcinogenic properties due to their antioxidant activities (Makhafola et al., 2016; Bouyahya et al., 2022). When evaluating the superclasses present in the flowers of the two cultivars, the Phenylpropanoids and polyketide superclass (comprised of polyphenolic compounds) was less dominant. Contradictory to this observation, flowers are regarded as important reproductive organs thus high levels of polyphenols are expected (Piccolella et al., 2020). However, the clear lack of polyphenols in the flowers of the studied cultivars (Fig. 6B) poses questions on the presence of intact biosynthetic routes of these compounds or the effects that the breeding process could have had on the chemistry of these cultivars. Moreover, literature regarding polyphenols in the flowers of *Cannabis* cultivars remains scarce and limited (Izzo et al., 2020). It has been shown that *Cannabis* cultivars lose genetic variation due to domestication and excessive breeding for selective traits (Clarke & Merlin, 2016). Therefore, such losses could contribute to the observed differences in the distribution of the phenylpropanoids and polyketides (polyphenols) between the leaves and flowers of the studied cultivars.

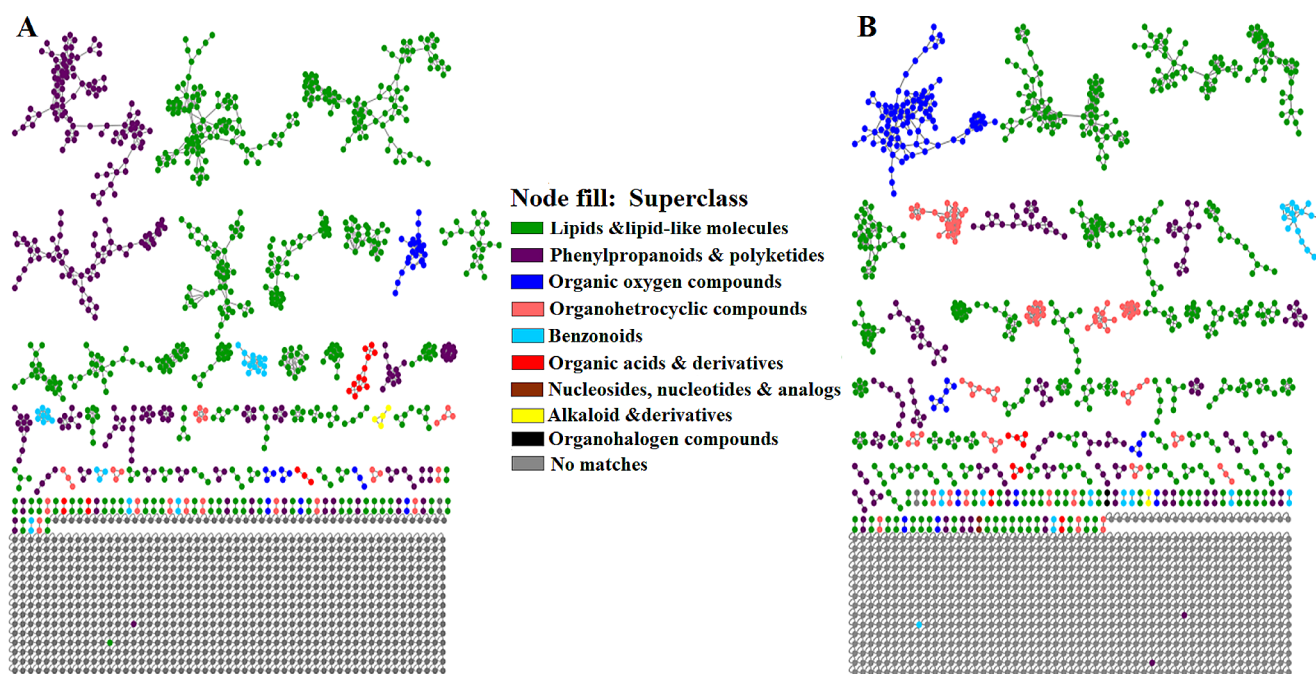


Fig. 6 Enhanced molecular networking. MolNetEnhancer mass spectral FBMN of (A) leaves and (B) flowers of both AMNH and RDC *Cannabis* cultivars acquired in ESI negative. The MolNetEnhancer FBMNs highlight the chemical superclasses present in leaves and flowers of two cultivars. The networks were computed from the combined outputs of FBMN putative annotations, substructure annotations (MS2LDA), network annotation propagation (NAP), and DEREPLICATOR.

Additionally, *Cannabis* genomics data lacks plant-tissue-specific data (Hussain et al., 2021) but the metabolite data generated in this study could lay the foundation for further exploration of the regulatory genetic networks of chemical classes in plant tissues of *Cannabis* cultivars. For example, when zooming in on these identified superclasses and detailing the relative abundance of the identified metabolites, Figure S2 highlights the cultivar-specific and plant-tissue-specific metabolite differences where lipid-like molecules such as Glc-Glc-octadecatrienoyl-sn-glycerol (isomer 2) are dominant in the tissues of RDC while flavonoids (phenylpropanoids and polyketide superclass) such as isoquercitrin are more dominant in the tissues of AMNH. These metabolic differences could be related to changes in regulatory networks underlying the specialized metabolome of these cultivars.

The chemical profiling achieved through the GNPS-molecular networking tools revealed a complex and diverse phytochemical space of the cultivars based on the leaf and flower metabolite profiles. The leaves and flowers were characterized by the presence of chemical classes such as flavonoids, cannabinoids and phospholipids which have various medicinal properties as discussed above. Possible anti-cancer properties are one of the medicinal properties that were postulated from the revealed metabolomes of the two cultivars. To further explore this postulation, in silico molecular docking studies were performed to computationally predict the anti-cancer or anti-proliferative properties

of the revealed cultivar chemistries supported by biological assay studies.

3.3 Anti-proliferative properties of AMNH and RDC: in silico molecular docking and biological assay studies

The phytochemistry of both cultivars illuminated various chemical classes including cannabinoids and flavonoids—with some described to have anti-cancer properties amongst other biological activities. Selected cannabinoid and flavonoid ligands, based on the elucidated chemical profiles of the two cultivars (Tables S1 and S3), were docked against various cancer targets. Each ligand was docked against its known target since the targets have been shown or suggested to play vital roles in cancer pathways. The interactions of the ligands with the receptor/protein targets were investigated through computer-assisted modelling to reveal their effective binding capacities (Table S3). Illustratively, the most significant receptor and ligand interactive binding affinity or docking score (-10.3 kcal/mol) was observed between cannabinoid receptor type 2 (CB2) and Δ^9 -THC (Fig. 7). The amino acid residues involved in the interaction were shown to be SER285, PHE281, VAL113 and SER90.

This was followed by the interaction between cannabinoid receptor 2 (CB2) and CBD where the binding affinity of -9.5 kcal/mol. CB1 and CB2 are endocannabinoid cannabinoid receptors or membrane proteins with seven

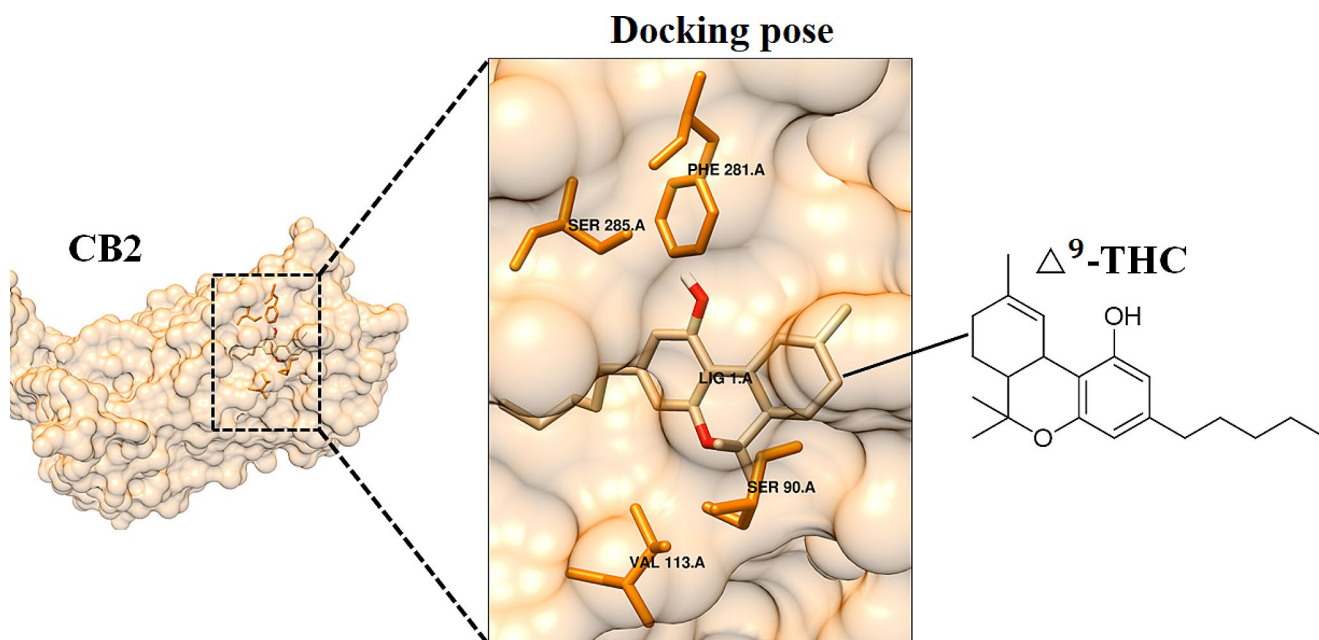


Fig. 7 Molecular docking to predict the activity of selected *Cannabis* metabolites. Structure of CB2 active site in complex with Δ^9 -THC. The docking pose indicates the receptor interactive amino acid residues SER285, PHE281, VAL113 and SER90 that position Δ^9 -THC in the binding pocket of CB2.

transmembrane helices and belong to the rhodopsin-like G-protein-coupled receptor superfamily (Sharif et al., 2016). CB1 is predominately found in the central nervous system (CNS) and in addition to its ability to bind to a wide range of cannabinoids, CB1 actively binds and mediates the psychoactive effects of Δ^9 -THC (Kendall & Yudowski, 2017). On the other hand, CB2, which had the highest docking score or interacting with Δ^9 -THC, is known to play a significant role in the regulation of immune responses, inflammation, pain, and other metabolic processes (Yeliseev & Gawrisch, 2017). In the many types of cancers, the activation of both CB1 and CB2 triggers several pathways. In lung cancer, it has been reported that when bound to the CB receptors, Δ^9 -THC (identified only in the flowers of the cultivars-Table S1) upregulates Tribbles homolog 3 (TRB3) which propagates autophagy mediated apoptosis (Salazar et al., 2009; Fu et al., 2023). Moreover, studies conducted on pancreatic Mia PaCa2 cell line showed that, Δ^9 -THC induced caspase-3 activation (characteristic of apoptotic cell death as illustrated above) and stimulated the *de novo* synthesis of ceramide which increased the cell apoptotic rate through the up-regulation of stress-regulated protein p8 (Carracedo et al., 2006; Laezza et al., 2020).

The cannabinoid CBD, which is known to interact with CB1 and CB2 in cancer pathways, was also shown to interact with the G-coupled protein receptor 55 (GPR55) where the docking score was measured to be -7.8 kcal/mol (Table S3), slightly lower than the docking scores on CB1 and CB2. GPR55 is one of the newly discovered endogenous cannabinoid receptors that now form part of the extended endocannabinoid system (Ramer et al., 2019). The endocannabinoid receptor GPR55 is abundant in the brain, skeletal muscle, gastrointestinal (GI) tract, white adipose tissue, the islets of Langerhans (β and α cells) and the pancreas. In cancer studies, GPR55 has been investigated for its vital role in cancer-promoting activities (Falasca & Falasca, 2022). The normal stimulation of GPR55 by its endogenous ligand (lysophosphatidylinositol (LPI) activates the pro-tumorigenic Akt and extracellular receptor kinase (ERK) pathways which promote cell proliferation. However, when CBD binds to GPR55, it acts as an antagonist and thus promotes antiproliferation effects in cancer cell lines (Laezza et al., 2020).

When evaluating the flavonoid content of cultivars, literature highlights that more than 20 flavonoids have been identified in *Cannabis* of which the most abundant are flavone (luteolin and apigenin) and flavonol (quercetin and kaempferol) aglycones and glycosides (Bautista et al., 2021; Cásedas et al., 2022). Thus, flavonoids identified and reported in Table S1 such as apigenin, quercetin and kaempferol, in their non-glycosylated forms, were also investigated for their binding affinities to cancer targets (Table S3). Moreover, the anti-cancer properties of

cannflavin A, a flavone unique to the *Cannabis* plant, were also investigated. Among all the screened flavonoids, cannflavin A had the most significant docking score of -9.3 kcal/mol through its interaction with amino acids SER144, THR143, GLU196, TRY200, PRO206, TRY202, ARG167 on caspase 3, a cysteine-aspartic acid protease that is predominantly found in the cytoplasm of cells and exists as an inactive pro-enzyme (pro-caspase 3) (Luo et al., 2010; Ponder & Boise, 2019). Studies done on cannflavin A as a ligand to caspase-3 have shown that the interaction causes slight cleavage on caspase-3 which results in cytotoxic effects to a variety of cancer cells such as human bladder transitional carcinoma cells (Tomko et al., 2022). The stimulation of caspase-3 results in apoptosis, often described as a “point of no return” for a cell and is characterized by apoptotic nuclear changes such as DNA fragmentation, chromatin condensation and nuclear disruption (Luo et al., 2010; Boudreau et al., 2019).

Overall, all the flavonoids and cannabinoids exhibited good binding affinity with the various cancer targets which may be responsible for the anti-proliferative properties of *Cannabis* flower extracts observed in Figures S4 and Fig. 8. When considering *Cannabis* in natural consumer products and cancer research, several studies have been done to support the efficacy of *Cannabis* varieties for various cancer treatment-related symptoms. Some studies have also explored the efficacy of *Cannabis* in inducing apoptosis in cancer cells. However, most of these studies mainly focus on the effect of isolated major cannabinoids (e.g., CBD, CBG, THC etc.) and focus less on other chemical classes or their combined synergistic effect. Since *Cannabis* flowers are mostly studied in drug discovery and bioactive compound exploration, herein, the flowers of the two cultivars (AMNHF and RDCF) previously profiled for their chemical content (through computational metabolomics) were investigated for anti-proliferative effects through in vitro cell viability assays complemented with in silico molecular docking studies (discussed above) to determine if they elicited similar cytotoxic effects. The in vitro cell viability assays such as the alamar blue assay (Longhin et al., 2022) showed that both AMNHF and RDCF extracts were cytotoxic to the cancerous MIA PaCa-2 cell line which indicated anti-proliferative properties (Figure S4). However, the effect of AMNHF displayed a concentration-dependent pattern on the MIA PaCa-2 cells where lower cell viability (high cell death) was achieved through lower extract concentrations which can be seen as counterintuitive. This points out the instability of crude extracts versus isolated compounds on cytotoxic screening since different compounds could be present at different concentrations thus the observed effect. Additionally, the observed effect can be further investigated through more reliable and reproducible studies such as xCELLigence cell

index analysis to show real-time cytotoxic activities of the extracts over time (Kho et al., 2015; Cerignoli et al., 2018). Nonetheless, the observed anti-proliferative properties were further supported and highlighted by the caspase activities of the MIA PaCa-2 cells upon being treated with AMNHF and RDCF extracts respectively.

Various stimuli initiate or induce the energy-dependent apoptosis, an anti-proliferative process, where the activation of enzymes called caspases (cysteine-aspartic proteases) occurs (Saraste & Pulkki, 2000; Elmore, 2007). Caspases are proteolytic enzymes that have well-defined roles in cell death mediated by apoptosis, necroptosis, and autophagy. Apoptotic caspases (involved in both intrinsic and extrinsic apoptotic pathways) are expressed as inactive pro-caspases that are activated to their active forms- caspase 2,3,7,8,9 and 10. Their activation propagates a cascade of signalling events that result in the controlled demolition of cellular components (McIlwain et al., 2013; Shalini et al., 2015).

Cell death caused by AMNHF and RDCF on MIA PaCa-2 treated cells was clearly indicated in Figure S4, and since caspases are over-expressed during cell death caused by apoptosis, caspase activity was then investigated. An evident increase in caspase activity was seen for the RDCF and AMNHF treated MIA PaCa-2 cells (Fig. 8). The untreated cells were observed to have a reading of 3378.3 RLU while RDCF, following the chemotherapeutic drug (etoposide), presented a significantly high average luminescence reading of 7093.3 RLU- which was a 210% increase in caspase activity when compared to the untreated cells. Moreover, the AMNHF treated cells presented an average luminescence reading of 4603.7 RLU, a 136.3% increase in caspase activity which was lower than that of RDCF. However, these evident increases in caspase activity indicate the

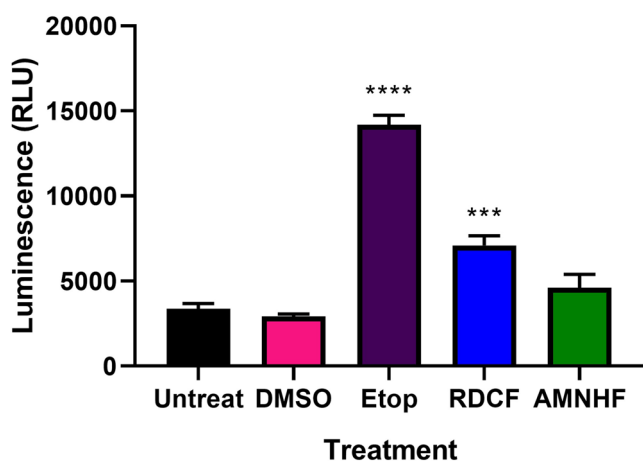


Fig. 8 Caspase activity. The average luminescence readings (in relative light units) for the 24-hour treatment of the MIA PaCa 2 cell line showing the untreated cells, the cells treated with 0.1% DMSO, 100 μ M etoposide and the IC50s of RDCF (65.7 μ g/ml) and AMNHF (2 μ g/ml). The asterisk (*) represents (*** P < 0.001, **** P < 0.0001) as calculated by a T-test between untreated and treated samples.

possible up-regulation of apoptotic caspases (caspase 3 and 7) and thus point to apoptotic events due to the treatment of the cells with RDCF and AMNHF. These results also correlate to the Alamar blue cytotoxic activities observed on the treated MIA PaCa-2 cells.

No data has been reported on the displayed anti-proliferative activities of *Cannabis* cultivars AMNHF and RDCF. Therefore, we may hypothesize that the observed activities stem from the distinct chemical profiles of the two cultivars. Consequently, AMNHF and RDCF present promising subjects for upcoming investigations, including isolating compounds of various chemical classes and assessing their potential for anti-cancer properties. It is noted that these compounds usually undergo natural metabolism upon introduction into the human body. Therefore, future research focusing on the catabolism of these metabolites following *Cannabis* consumption and their route within the body will be important to assess the full health-related properties of various *Cannabis* cultivars.

4 Conclusion

The numerous *Cannabis* cultivars that are made accessible for medicinal purposes, are classified into several categories based only on the plant's cannabinoid content. Nonetheless, the metabolomics computational strategies used herein such as molecular networking approaches, helped in visualizing, and elucidating the chemical diversity of the studied cultivars beyond their cannabinoid content. The revealed plant-tissue based chemical profiles of the studied cultivars showed variation in the distribution of phytocannabinoids (major and minor) as well as other phytochemicals. Moreover, the phytochemical space of the leaves (which are underused in clinical studies) was highlighted as an alternative source of compounds with possible medicinal value. Such revelations, including the varying distribution of chemical classes across the studied cultivars, can assist in understanding the metabolome of *Cannabis* and gives an opportunity for the discovery of novel compounds in the elucidated chemical space.

Since the therapeutic outcomes of *Cannabis* are non-generic, the knowledge generated from the chemotyping done herein can also inform further experiments such as the testing of the biological activities of the plant-tissues as they may be used to address different ailments based on their non-generic chemical composition. When considering the flower anti-proliferative activities of the extracts, these findings put emphasis on the importance of metabolic profiling *Cannabis* cultivars as they can inform their use as pharmacological agents, in this case as possible cancer suppressors. However, since this study was based on crude

extracts, the anti-proliferative activities observed could be a result of the synergistic effect of multiple compounds. Thus, further studies such as fractionation and compound isolation and evaluation of some of the identified metabolites in the studied *Cannabis* extracts on cancer cell lines are still needed. Moreover, based on the observed metabolite differences between the plant tissues of the studied cultivars, future studies could include integrating genomics and transcriptomics studies for in-depth analysis of the biosynthetic routes and regulation of the chemical compound classes illuminated in the cultivars.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11306-024-02125-y>.

Acknowledgements University of Venda, Biochemistry Department are gratefully thanked for access to the SHIMADZU LCMS-9030. The South African national research fund (NRF) is highly thanked for bursary support to A.M.

Author contributions A.M. performed experiments, data analysis, data curation, wrote the original draft. M.C. supervised the study, performed data curation and analysis. A.P.K. supervised the molecular docking work, edited the manuscript text. N.E.M. supervised the LCMS analyses, data analysis, reviewed and edited the manuscript text. J.J.J.vd.H. supervised the data analysis and curation, reviewed and edited the manuscript text. F.T. conceptualized the study, supervised the study, data analysis and curation, reviewed and edited the manuscript text, and attended to the project administration. All authors read and approved the manuscript.

Data availability The spectral data presented in this study is made available on the supplementary files as job links and can also be found in an online repository at <https://massive.ucsd.edu/MSV000093886>.

Declarations

Conflicts of interest JJJvdH is member of the Scientific Advisory Board of NAICONS Srl., Milano, Italy and consults for Corteva Agriscience, Indianapolis, IN, USA. All other authors declare to have no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aizpurua-Olaizola, O., Soydaner, U., Öztürk, E., Schibano, D., Simsir, Y., Navarro, P., Etxebarria, N., & Usobiaga, A. (2016). Evolution of the cannabinoid and terpene content during the growth of *Cannabis sativa* plants from different chemotypes. *Journal of Natural Products*, 79(2), 324–331. <https://doi.org/10.1021/acs.jnatprod.5b00949>.
- Aliferis, K. A., & Bernard-Perron, D. (2020). Cannabinomics: Application of Metabolomics in Cannabis (*Cannabis sativa* L.) Research and Development. *Frontiers in Plant Science*, 11(5), 1–20. <https://doi.org/10.3389/fpls.2020.00554>.
- Aron, A. T., Gentry, E. C., McPhail, K. L., Nothias, L. F., Nothias-Esposito, M., Bouslimani, A., Petras, D., Gauglitz, J. M., Sikora, N., Vargas, F., van der Hooft, J. J. J., Ernst, M., Kang, K., Bin, Aceves, C. M., Caraballo-Rodríguez, A. M., Koester, I., Weldon, K. C., Bertrand, S., Roullier, C., & Dorrestein, P. C. (2020). Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nature Protocols*, 15(6), 1954–1991. <https://doi.org/10.1038/s41596-020-0317-5>.
- Balant, M., Gras, A., Ruz, M., Vallès, J., Vitales, D., & Garnatje, T. (2021). Traditional uses of Cannabis: An analysis of the CANNUSE database. *Journal of Ethnopharmacology*, 279, 1–11. <https://doi.org/10.1016/j.jep.2021.114362>.
- Bautista, J. L., Yu, S., & Tian, L. (2021). Flavonoids in Cannabis sativa: Biosynthesis, Bioactivities, and Biotechnology. *ACS Omega*, 6(8), 5119–5123. <https://doi.org/10.1021/acsomega.1c00318>.
- Benididir, M. A., Kang, K., bin, Genta-Jouve, G., Huber, F., Rogers, S., & van der Hooft, J. J. J. (2021). Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches. *Natural Product Reports*, 38(11), 1967–1993. <https://doi.org/10.1039/d1np00023c>.
- Boudreau, M. W., Peh, J., & Hergenrother, P. J. (2019). Procaspase-3 overexpression in Cancer: A Paradoxical Observation with therapeutic potential. *ACS Chemical Biology*, 14(1), 2335–2348. <https://doi.org/10.1021/acscchembio.9b00338>.
- Bouyahya, A., Omari, N. E., Bakrim, S., Hachlafi, N. E., Balahbib, A., Wilairatana, P., & Mubarak, M. S. (2022). Advances in Dietary Phenolic compounds to improve Chemosensitivity of Anticancer drugs. *Cancers*, 14(4573), 1–20. <https://doi.org/10.3390/cancers14194573>.
- Cao, J., Yin, C., Qin, Y., Cheng, Z., & Chen, D. (2014). Approach to the study of flavone di-C-glycosides by high performance liquid chromatography-tandem ion trap mass spectrometry and its application to characterization of flavonoid composition in *Viola yedoensis*. *Journal of Mass Spectrometry*, 49, 1010–1024. <https://doi.org/10.1002/jms.3413>.
- Carracedo, A., Gironella, M., Lorente, M., Garcia, S., Guzmán, M., Velasco, G., & Iovanna, J. L. (2006). Cannabinoids induce apoptosis of pancreatic tumor cells via endoplasmic reticulum stress-related genes. *Cancer Research*, 66(13), 6748–6755. <https://doi.org/10.1158/0008-5472.CAN-06-0169>.
- Cásedas, G., Moliner, C., Maggi, F., Mazzara, E., & López, V. (2022). Evaluation of two different Cannabis sativa L. extracts as antioxidant and neuroprotective agents. *Frontiers in Pharmacology*, 13, 1–11. <https://doi.org/10.3389/fphar.2022.1009868>.
- Cerignoli, F., Abassi, Y. A., Lamarche, B. J., Guenther, G., Ana, D. S., Guimet, D., Zhang, W., Zhang, J., & Xi, B. (2018). In vitro immunotherapy potency assays using real-time cell analysis. *Plos One*, 13(3), 1–21. <https://doi.org/10.1371/journal.pone.0193498>.

- Cicaloni, V., Salvini, L., Vitalini, S., & Garzoli, S. (2022). Chemical profiling and characterization of different cultivars of *Cannabis sativa* L. inflorescences by SPME-GC-MS and UPLC-MS. *Separations*, 9(4). <https://doi.org/10.3390/separations9040090>.
- Clarke, R., & Merlin, M. (2016). *Cannabis: Evolution and ethnobotany*. University of California Press.
- Connecely, L. J., Mauleon, R., Mieog, J., Barkla, B. J., & Kretzschmar, T. (2021). Characterization of the *Cannabis sativa* glandular trichome proteome. *PLoS One*, 16(4). <https://doi.org/10.1371/journal.pone.0242633>.
- Cuyckens, F., & Claeys, M. (2004). Mass spectrometry in the structural analysis of flavonoids. *Journal of Mass Spectrometry*, 39, 1–15. <https://doi.org/10.1002/jms.585>.
- da Silva, R. R., Wang, M., Nothias, L. F., van der Hooft, J. J. J., Caraballo-Rodríguez, A. M., Fox, E., Balunas, M. J., Klassen, J. L., Lopes, N. P., & Dorrestein, P. C. (2018). Propagating annotations of molecular networks using in silico fragmentation. *PLoS Computational Biology*, 14(4). <https://doi.org/10.1371/journal.pcbi.1006089>.
- Duggan, P. J. (2021). The Chemistry of Cannabis and cannabinoids. *Australian Journal of Chemistry*, 74(6), 369–387. <https://doi.org/10.1071/CH21006>.
- Elmore, S. (2007). Apoptosis: A review of programmed cell death. *Toxicologic Pathology*, 25(4), 1–40.
- Ernst, M., Kang, K., Bin, Caraballo-Rodríguez, A. M., Nothias, L. F., Wandy, J., Chen, C., Wang, M., Rogers, S., Medema, M. H., Dorrestein, P. C., & van der Hooft, J. J. J. (2019). Molnetenhancer: Enhanced molecular networks by integrating metabolome mining and annotation tools. *Metabolites*, 9(7). <https://doi.org/10.3390/metabo9070144>.
- Falasca, V., & Falasca, M. (2022). Targeting the Endocannabinoidome in Pancreatic Cancer. *Biomolecules*, 12(2). <https://doi.org/10.3390/biom12020320>.
- Ferber, S. G., Namdar, D., Hen-Shoval, D., Eger, G., Koltai, H., Shoval, G., Shbiro, L., & Weller, A. (2020). The entourage effect: Terpenes coupled with cannabinoids for the treatment of Mood disorders and anxiety disorders. *Current Neuropharmacology*, 18(2), 87–96. <https://doi.org/10.2174/1570159X17666190903103923>.
- Filipiuc, L. E., Ababei, D. C., Alexa-Stratulat, T., Pricope, C. V., Bild, V., Stefanescu, R., Stanciu, G. D., & Tamba, B. I. (2021). Major phytocannabinoids and their related compounds: Should we only search for drugs that act on cannabinoid receptors? *Pharmaceutics*, 13(11). <https://doi.org/10.3390/pharmaceutics13111823>.
- Fu, Z., Zhao, P. Y., Yang, X. P., Li, H., Hu, S. D., Xu, Y. X., & Du, X. H. (2023). Cannabidiol regulates apoptosis and autophagy in inflammation and cancer: A review. *Frontiers in Pharmacology*, 14. <https://doi.org/10.3389/fphar.2023.1094020>.
- Geng, P., Sun, J., Zhang, M., Li, X., Harnly, J. M., & Chen, P. (2016). Comprehensive characterization of C-glycosyl flavones in wheat (*Triticum aestivum*, L.) germ using UPLC-PDA-ESI/HRMSⁿ and mass defect filtering. *Journal of Mass Spectrometry*, 51, 914–930. <https://doi.org/10.1002/jms.3803>.
- Gloss, D. (2015). An overview of products and Bias in Research. *Neurotherapeutics*, 12(4), 731–734. <https://doi.org/10.1007/s13311-015-0370-x>.
- Goerl, B., Watkins, S., Metcalf, C., Smith, M., & Beenhakker, M. (2021). Cannabidiolic acid exhibits entourage-like improvements of anticonvulsant activity in an acute rat model of seizures. *Epilepsy Research*, 169, 1–5. <https://doi.org/10.1016/j.epilepsyres.2020.106525>.
- Hammerle, F., Quirós-Guerrero, L., Rutz, A., Wolfender, J. L., Schöbel, H., Peintner, U., & Siewert, B. (2021). Feature-based molecular Networking—An Exciting Tool to Spot species of the Genus *Cortinarius* with Hidden Photosensitizers. *Metabolites*, 11(11). <https://doi.org/10.3390/metabo11110791>.
- Hussain, T., Jeena, G., Pitakbut, T., Vasilev, N., & Kayser, O. (2021). iScience Cannabis sativa research trends, challenges, and new-age perspectives. *ISCIENCE*, 24(103391). <https://doi.org/10.1016/j.isci.2021.103391>.
- Izzo, L., Castaldo, L., Narváez, A., Graziani, G., Gaspari, A., Rodríguez-Carrasco, Y., & Ritieni, A. (2020). Analysis of Phenolic compounds in Commercial Cannabis sativa L. inflorescences using UHPLC-Q-Orbitrap HRMs. *Molecules*, 25(3). <https://doi.org/10.3390/molecules25030631>.
- Jin, D., Dai, K., Xie, Z., & Chen, J. (2020). Secondary metabolites profiled in Cannabis inflorescences, leaves, stem barks, and roots for Medicinal purposes. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-60172-6>.
- Kang, K., bin, Ernst, M., van der Hooft, J. J. J., da Silva, R. R., Park, J., Medema, M. H., Sung, S. H., & Dorrestein, P. C. (2019). Comprehensive mass spectrometry-guided phenotyping of plant specialized metabolites reveals metabolic diversity in the cosmopolitan plant family Rhamnaceae. *Plant Journal*, 98(6), 1134–1144. <https://doi.org/10.1111/tpj.14292>.
- Kendall, D. A., & Yudowski, G. A. (2017). Cannabinoid receptors in the central nervous system: Their signaling and roles in disease. *Frontiers in Cellular Neuroscience*, 10. <https://doi.org/10.3389/fncel.2016.00294>.
- Kho, D., MacDonald, C., Johnson, R., Unsworth, C. P., O'Carroll, S. J., du Mez, E., Angel, C. E., & Graham, E. S. (2015). Application of xCELLigence RTCA biosensor technology for revealing the profile and window of drug responsiveness in real time. *Biosensors*, 5(2), 199–222. <https://doi.org/10.3390/bios5020199>.
- Kim, C. K., Yu, J., & Lee, M. (2023). Molecular networking-guided isolation of a Phenolic Constituent from *Prunus mume* seed and its antioxidant and anti-inflammatory activities. *Foods*, 12(6), 1146. <https://doi.org/10.3390/foods12061146>.
- Kogan, N. M., & Mechoulam, R. (2007). Cannabinoids in health and disease. *Dialogues Clinical Neuroscience*, 9. www.dialogues-cns.org.
- Ladha, K. S., Ajrawat, P., Yang, Y., & Clarke, H. (2020). Understanding the Medical Chemistry of the Cannabis Plant is critical to Guiding Real World Clinical evidence. *Molecules*, 25(18). <https://doi.org/10.3390/molecules25184042>.
- Laezza, C., Pagano, C., Navarra, G., Pastorino, O., Proto, M. C., Fiore, D., Piscopo, C., Gazzero, P., & Bifulco, M. (2020). The endocannabinoid system: A target for cancer treatment. *International Journal of Molecular Sciences*, 21(3). <https://doi.org/10.3390/ijms21030747>.
- LaVigne, J. E., Hecksel, R., Keresztes, A., & Streicher, J. M. (2021). Cannabis sativa terpenes are cannabimimetic and selectively enhance cannabinoid activity. *Scientific Reports*, 11, 1–15. <https://doi.org/10.1038/s41598-021-87740-8>.
- LaVigne, J. E., Hecksel, R., Keresztes, A., & Streicher, J. M. (2021). Cannabis sativa terpenes are cannabimimetic and selectively enhance cannabinoid activity. *Scientific Reports*, 11, 1–15. <https://doi.org/10.1038/s41598-021-87740-8>.
- Lewis, M. A., Russo, E. B., & Smith, K. M. (2018). Pharmacological foundations of Cannabis chemovars. *Planta Medica*, 84(4), 225–233. <https://doi.org/10.1055/s-0043-122240>.
- Li, C. R., Yang, L. X., Guo, Z. F., Yang, H., Zhang, Y., Wang, Y. M., Zhang, G. Z., Li, P., & Gao, W. (2022). LC-MS-based untargeted metabolomics reveals chemical differences of Cannabis leaves from different regions of China. *Industrial Crops and Products*, 176. <https://doi.org/10.1016/j.indcrop.2021.114411>.
- Longhin, E. M., Yamani, E., Rundén-Pran, N., E., & Dusinska, M. (2022). The alamar blue assay in the context of safety testing of nanomaterials. *Frontiers in Toxicology*, 4. <https://doi.org/10.3389/ftox.2022.981701>.
- Lowe, H., Steele, B., Bryant, J., Toyang, N., & Ngwa, W. (2021). Non-cannabinoid metabolites of Cannabis sativa L. with

- therapeutic potential. *Plants*, 10(2), 1–17. <https://doi.org/10.3390/plants10020400>.
- Lubell, J. D., & Brand, M. H. (2018). Foliar Sprays of Silver Thio-sulfate Produce Male Flowers on female hemp plants. *Hort-Technology Hortte*, 28(6), 743–747. <https://doi.org/10.21273/HORTTECH04188-18>.
- Luchini, A., Fredolini, C., Espina, B., Meani, F., Reeder, A., Rucker, S., Petricoin, E., & Liotta, L. (2010). Nanoparticle Technology: Addressing the Fundamental roadblocks to protein Biomarker Discovery. *Current Molecular Medicine*, 10(2), 133–141. <https://doi.org/10.2174/156652410790963268>.
- Luo, M., Lu, Z., Sun, H., Yuan, K., Zhang, Q., Meng, S., Wang, F., Guo, H., Ju, X., Liu, Y., Ye, T., Lu, Z., & Zhai, Z. (2010). Nuclear entry of active caspase-3 is facilitated by its p3-recognition-based specific cleavage activity. *Cell Research*, 20(2), 211–222. <https://doi.org/10.1038/cr.2010.9>.
- Ma, Y.-L., Vedernikova, I., van den Heuvel, H., & Claeys, M. (2000). Internal glucose residue loss in protonated *O*-diglycosyl flavonoids upon low-energy collision-induced dissociation. *Journal of the American Society for Mass Spectrometry*, 11(2), 136–144. [https://doi.org/10.1016/S1044-0305\(99\)00133-6](https://doi.org/10.1016/S1044-0305(99)00133-6).
- Maayah, Z. H., Takahara, S., Ferdaoussi, M., & Dyck, J. R. B. (2020). The molecular mechanisms that underpin the biological benefits of full-spectrum cannabis extract in the treatment of neuropathic pain and inflammation. In *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1866(7), 1–12. <https://doi.org/10.1016/j.bbadis.2020.165771>.
- Makhafola, T. J., Elgorashi, E. E., McGaw, L. J., Verschaeye, L., & Eloff, J. N. (2016). The correlation between antimutagenic activity and total phenolic content of extracts of 31 plant species with high antioxidant activity. *BMC Complementary and Alternative Medicine*, 16(1). <https://doi.org/10.1186/s12906-016-1437-x>.
- McIlwain, D. R., Berger, T., & Mak, T. W. (2013). Caspase functions in cell death and disease. *Cold Spring Harbor Perspectives in Biology*, 5(4), 1–28. <https://doi.org/10.1101/cshperspect.a008656>.
- Mohimani, H., Gurevich, A., Shlemov, A., Mikheenko, A., Korobeynikov, A., Cao, L., Shcherbin, E., Nothias, L. F., Dorrestein, P. C., & Pevzner, P. A. (2018). Dereplication of microbial metabolites through database search of mass spectra. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-018-06082-8>.
- Moltke, J., & Hindocha, C. (2021). Reasons for cannabidiol use: A cross-sectional study of CBD users, focusing on self-perceived stress, anxiety, and sleep problems. *Journal of Cannabis Research*, 3(1). <https://doi.org/10.1186/s42238-021-00061-5>.
- Namdar, D., Mazuz, M., Ion, A., & Koltai, H. (2018). Variation in the compositions of cannabinoid and terpenoids in Cannabis sativa derived from inflorescence position along the stem and extraction methods. *Industrial Crops and Products*, 113, 376–382. <https://doi.org/10.1016/j.indcrop.2018.01.060>.
- Nephali, L., Steenkamp, P., Burgess, K., Huyser, J., Brand, M., J van der Hooft, J. J., & Tugizimana, F. (2022). Mass Spectral Molecular networking to Profile the Metabolome of Biostimulant Bacillus strains. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.920963>.
- Nothias, L. F., Nothias-Esposito, M., da Silva, R., Wang, M., Protsyuk, I., Zhang, Z., Sarvepalli, A., Leyssen, P., Touboul, D., Costa, J., Paolini, J., Alexandrov, T., Litaudon, M., & Dorrestein, P. C. (2018). Bioactivity-based molecular networking for the Discovery of Drug leads in natural product bioassay-guided fractionation. *Journal of Natural Products*, 81(4), 758–767. <https://doi.org/10.1021/acs.jnatprod.7b00737>.
- Nothias, L. F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., Protsyuk, I., Ernst, M., Tsugawa, H., Fleischauer, M., Aicheler, F., Aksenov, A. A., Alka, O., Allard, P. M., Barsch, A., Cachet, X., Caraballo-Rodríguez, A. M., da Silva, R. R., Dang, T., & Dorrestein, P. C. (2020). Feature-based molecular networking in the GNPS analysis environment. *Nature Methods*, 17(9), 905–908. <https://doi.org/10.1038/s41592-020-0933-6>.
- Odieka, A. E., Obuzor, G. U., Oyediji, O. O., Gondwe, M., Hosu, Y. S., & Oyediji, A. O. (2022). The Medicinal Natural products of Cannabis sativa Linn.: A review. *Molecules*, 27(5). <https://doi.org/10.3390/molecules27051689>.
- Pang, Z., Chong, J., Zhou, G., De Lima Morais, D. A., Chang, L., Barrette, M., Gauthier, C., Jacques, P. É., Li, S., & Xia, J. (2021). MetaboAnalyst 5.0: Narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research*, 49(W1), W388–W396. <https://doi.org/10.1093/nar/gkab382>.
- Pattnaik, F., Nanda, S., Mohanty, S., Dalai, A. K., Kumar, V., Pon-nusamy, S. K., & Naik, S. (2022). Cannabis: Chemistry, extraction and therapeutic applications. *Chemosphere*, 289. <https://doi.org/10.1016/j.chemosphere.2021.133012>.
- Piccolella, S., Crescente, G., Formato, M., & Pacifico, S. (2020). A cup of hemp coffee by moka pot from southern Italy: An uhplc-hrms investigation. *Foods*, 9(8). <https://doi.org/10.3390/foods9081123>.
- Ponder, K. G., & Boise, L. H. (2019). The prodomain of caspase-3 regulates its own removal and caspase activation. *Cell Death Discovery*, 5(1). <https://doi.org/10.1038/s41420-019-0142-1>.
- Procaccia, S., Lewitah, G. M., Feder, L., Shapira, C., Berman, A., P., & Meiri, D. (2022). Cannabis for Medical Use: Versatile Plant Rather Than a Single Drug. In *Frontiers in Pharmacology*, 13. <https://doi.org/10.3389/fphar.2022.894960>.
- Quinn, R. A., Nothias, L. F., Vining, O., Meehan, M., Esquenazi, E., & Dorrestein, P. C. (2017). Molecular networking as a drug Discovery, Drug Metabolism, and Precision Medicine Strategy. *Trends in Pharmacological Sciences*, 38(2), 143–154. <https://doi.org/10.1016/j.tips.2016.10.011>.
- Radwan, M. M., Chandra, S., Gul, S., & Elsohly, M. A. (2021). Cannabinoids, phenolics, terpenes and alkaloids of Cannabis. *Molecules*, 26(9). <https://doi.org/10.3390/molecules26092774>.
- Ramabulana, A. T., Petras, D., Madala, N. E., & Tugizimana, F. (2021). Metabolomics and molecular networking to characterize the chemical space of four momordica plant species. *Metabolites*, 11(11). <https://doi.org/10.3390/metabo11110763>.
- Ramer, R., Schwarz, R., & Hinz, B. (2019). Modulation of the endocannabinoid system as a potential anticancer strategy. *Frontiers in Pharmacology*, 10(5). <https://doi.org/10.3389/fphar.2019.00430>.
- Rea, K. A., Casaretto, J. A., Al-Abdul-Wahid, M. S., Sukumaran, A., Geddes-McAlister, J., Rothstein, S. J., & Akhtar, T. A. (2019). Biosynthesis of cannflavins A and B from Cannabis sativa L. *Phytochemistry*, 164, 162–171. <https://doi.org/10.1016/j.phytochem.2019.05.009>.
- Rogers, S., Ong, C. W., Wandy, J., Ernst, M., Ridder, L., & van der Hooft, J. J. J. (2019). Deciphering complex metabolite mixtures by unsupervised and supervised substructure discovery and semi-automated annotation from MS/MS spectra. *Faraday Discussions*, 218, 284–302. <https://doi.org/10.1039/c8fd00235e>.
- Rupasinghe, H. P. V., Zheljazkov, V. D., Davis, A., Kumar, S. K., & Murray, B. (2020). Industrial Hemp (Cannabis sativa subsp. sativa) as an emerging source for Value-Added Functional Food Ingredients and nutraceuticals. *Molecules*, 25, 1–24. <https://doi.org/10.3390/molecules25184078>.
- Salazar, M., Carracedo, A., Salanueva, Í. J., Hernández-Tiedra, S., Lorente, M., Egia, A., Vázquez, P., Blázquez, C., Torres, S., García, S., Nowak, J., Fimia, G. M., Piacentini, M., Cecconi, F., Pandolfi, P. P., González-Feria, L., Iovanna, J. L., Guzmán, M., Boya, P., & Velasco, G. (2009). Cannabinoid action induces autophagy-mediated cell death through stimulation of ER stress in human glioma cells. *Journal of Clinical Investigation*, 119(5), 1359–1372. <https://doi.org/10.1172/JCI37948>.
- Saraste, A., & Pulkki, K. (2000). Morphologic and biochemical hallmarks of apoptosis. *Cardiovascular Research*, 45. www.elsevier.com/locate/cardioreswww.elsevier.nl/locate/cardiores.

- Sarma, N. D., Waye, A., Elsohly, M. A., Brown, P. N., Elzinga, S., Johnson, H. E., Marles, R. J., Melanson, J. E., Russo, E., Deyton, L., Hudalla, C., Vrdoljak, G. A., Wurzer, J. H., Khan, I. A., Kim, N. C., & Giancaspro, G. I. (2020). Cannabis Inflorescence for Medical purposes: USP considerations for Quality attributes. *Journal of Natural Products*, 83(4), 1334–1351. <https://doi.org/10.1021/acs.jnatprod.9b01200>.
- Shalini, S., Dorstyn, L., Dawar, S., & Kumar, S. (2015). Old, new and emerging functions of caspases. *Cell Death and Differentiation*, 22(4), 526–539. <https://doi.org/10.1038/cdd.2014.216>. Nature Publishing Group.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), 2498–2504. <https://doi.org/10.1101/gr.1239303>.
- Sharif, A., Nawaz, H., Rehman, R., Mushtaq, A., & Rashid, U. (2016). A review on bioactive potential of Benzoin Resin. *IJCBS*, 10. www.iscientific.org/Journal.html.
- Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., & Ideker, T. (2011). Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics*, 27(3), 431–432. <https://doi.org/10.1093/bioinformatics/btq675>.
- Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W. M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., & Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics*, 3(3), 211–221. <https://doi.org/10.1007/s11306-007-0082-2>.
- Tomko, A. M., Whynot, E. G., & Dupré, D. J. (2022). Anti-cancer properties of cannflavin A and potential synergistic effects with gemcitabine, cisplatin, and cannabinoids in bladder cancer. *Journal of Cannabis Research*, 4(1). <https://doi.org/10.1186/s42238-022-00151-y>.
- van der Hooft, J. J. J., Wandy, J., Barrett, M. P., Burgess, K. E. V., & Rogers, S. (2016). Topic modeling for untargeted substructure exploration in metabolomics. *Proceedings of the National Academy of Sciences of the United States of America*, 113(48), 13738–13743. <https://doi.org/10.1073/pnas.1608041113>.
- van der Hooft, J. J. J., Mohimani, H., Bauermeister, A., Dorrestein, P. C., Duncan, K. R., & Medema, M. H. (2020). Linking genomics and metabolomic diversity. *Royal Society of Chemistry*, 49, 3297–3314. <https://doi.org/10.1039/D0CS00162G>.
- Vásquez-Ocmín, P., Marti, G., Bonhomme, M., Mathis, F., Fournier, S., Bertani, S., & Maciuk, A. (2021). Cannabinoids vs. whole metabolome: Relevance of cannabinomics in analyzing Cannabis varieties. *Analytica Chemical Acta*, 1184, 1–29.
- Vincenti, F., Montesano, C., Di Ottavio, F., Gregori, A., Compagnone, D., Sergi, M., & Dorrestein, P. (2020). Molecular networking: A useful Tool for the identification of New Psychoactive substances in seizures by LC–HRMS. *Frontiers in Chemistry*, 8. <https://doi.org/10.3389/fchem.2020.572952>.
- Wandy, J., Zhu, Y., van der Hooft, J. J. J., Daly, R., Barrett, M. P., & Rogers, S. (2018). Ms2lda.org: Web-based topic modelling for substructure discovery in mass spectrometry. *Bioinformatics*, 34(2), 317–318. <https://doi.org/10.1093/bioinformatics/btx582>.
- Watts, G. (2006). Cannabis confusions. *Bmj*, 322, 1–2. <https://doi.org/10.1136/bmj.332.7534.175>.
- Yeliseev, A., & Gawrisch, K. (2017). Expression and NMR Structural Studies of Isotopically Labeled Cannabinoid Receptor Type II. *Methods in Enzymology*, 563, 387–403. Academic Press Inc. <https://doi.org/10.1016/bs.mie.2017.06.020>.
- Yu, J. S., Nothias, L. F., Wang, M., Kim, D. H., Dorrestein, P. C., Kang, K., Bin, & Yoo, H. H. (2022). Tandem Mass Spectrometry Molecular Networking as a powerful and efficient Tool for Drug Metabolism studies. *Analytical Chemistry*, 94(2), 1456–1464. <https://doi.org/10.1021/acs.analchem.1c04925>.
- Zhang, M., Otsuki, K., & Li, W. (2023). Molecular networking as a natural products discovery strategy. *Acta Materia Medica*, 2(2). <https://doi.org/10.15212/AMM-2023-0007>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.