Molecular Networking in Natural Product Research: A Modern Strategy for Dereplication and Bioactive Compound Discovery

R. Kamaleeshwari[1], Dr. V. Ganesh[2], P. Nandhini[3], G. Abinaya[4], G. Sivaranjani[5], C. Ramarpandi[6]

[1]Department of Pharmacognosy, The Tamil Nadu Dr.M.G.R Medical University(Chennai), Sasurie College of Pharmacy, Tiruppur.

[2,4,5]Department of Pharmaceutics, The Tamil Nadu Dr.M.G.R Medical University(Chennai), Sasurie College of Pharmacy, Tiruppur.

[3]Department of Pharmacology, The Tamil Nadu Dr.M.G.R Medical University(Chennai), Sasurie College of Pharmacy, Tiruppur.

[6]Department of Pharmaceutical Analysis, The Tamil Nadu Dr.M.G.R Medical University(Chennai), Sasurie College of Pharmacy, Tiruppur.

AUTHOR OF CORRESPONDANCE:

R. Kamaleeshwari

Department of Pharmacognosy,

The Tamil Nadu Dr.M.G.R Medical University(Chennai),

Sasurie College of Pharmacy, Tiruppur.

Email: kamalirajendiran075@gmail.com

ABSTRACT

Natural products obtained from medicinal plants, microorganisms, marine organisms, and other biological sources remain important resources for the discovery of bioactive molecules and therapeutic agents. However, the chemical complexity of natural-product extracts, the presence of numerous structurally related metabolites, and the repeated isolation of previously reported compounds present major challenges in natural-product research. Molecular networking has emerged as a modern computational strategy that enables the organization and visualization of metabolites based on similarities in their tandem mass spectrometry (MS/MS) fragmentation patterns. This review discusses the principles, workflow, analytical techniques, databases, and major applications of molecular networking in natural-product research. Particular emphasis is given to liquid chromatography–mass spectrometry/mass spectrometry (LC-MS/MS), high-resolution mass spectrometry (HRMS), Global Natural Products Social Molecular Networking (GNPS), feature-based molecular networking (FBMN), and spectral libraries used for metabolite annotation and dereplication. Molecular networking facilitates the rapid recognition of known compounds, characterization of related chemical families, and prioritization of unexplored molecular clusters, thereby reducing the problem of compound rediscovery. Its applications in pharmacognosy include phytochemical profiling, chemotaxonomy, botanical authentication, quality control, and herbal-drug standardization. Integration with metabolomics, in-silico fragmentation, chemometrics, and artificial intelligence further expands its potential for systematic natural-product investigation and bioactive compound discovery. Although limitations such as incomplete spectral databases, spectral similarity constraints, and the need for experimental confirmation remain, molecular networking provides a powerful complementary approach to conventional isolation and structure-elucidation methods. Overall, molecular networking represents an important advancement toward faster, data-driven, and comprehensive exploration of natural-product chemical diversity.

Keywords: *Molecular networking; Natural products; Dereplication; Bioactive compound discovery; LC-MS/MS; HRMS; GNPS; Feature-based molecular networking; Metabolomics; Pharmacognosy.*

1) INTRODUCTION

1.1 Natural Products and Drug Discovery

Natural products obtained from medicinal plants, microorganisms, marine organisms, and other biological sources have played an important role in the development of therapeutic agents. Their structural diversity, chemical complexity, and wide range of biological activities make them valuable sources of lead molecules for drug discovery. Several clinically important drugs and drug leads have originated directly from natural products or have been developed

from natural-product scaffolds. Natural products have particularly contributed to the discovery of therapeutic agents for infectious diseases, cancer, cardiovascular disorders, and other chronic diseases. [1,2]

The chemical diversity of natural products provides access to molecular structures that may be difficult to obtain through conventional synthetic approaches. Advances in analytical technologies, genomics, metabolomics, and bioinformatics have further expanded the ability to investigate this chemical diversity. Modern natural-product research therefore combines traditional approaches such as extraction, isolation, pharmacological screening, and structural elucidation with advanced analytical and computational techniques. [1,3]

Despite these advantages, the discovery of new bioactive natural products remains challenging because biological samples contain complex mixtures of primary and secondary metabolites. Consequently, efficient approaches are required to rapidly identify known compounds and prioritize potentially novel molecules for further investigation.

1.2 Challenges in Natural Product Research

Natural-product research involves several challenges from sample collection to the identification of biologically active compounds. Crude extracts may contain hundreds of metabolites with different chemical properties, often occurring at highly variable concentrations. The isolation and structural characterization of individual compounds from such complex mixtures can therefore be time-consuming and technically demanding. [1,4]

Another major challenge is the repeated isolation of compounds that have already been reported. This problem, commonly known as compound rediscovery, can consume considerable time and resources without providing genuinely new chemical entities. In addition, some bioactive metabolites may occur in very low concentrations, making their detection and isolation difficult. Structural similarities among related metabolites can further complicate compound identification and characterization. [4,5]

Traditional bioactivity-guided fractionation generally involves repeated separation, purification, and biological testing of fractions. Although this approach remains useful, it can require extensive experimental work before the chemical identity of an active constituent is established. These limitations have encouraged the development of rapid analytical and computational strategies for profiling complex natural-product mixtures and prioritizing promising metabolites. [1,3]

Figure:1 Natural products in drug design and discovery: Historical perspective.

1.3 Need for Dereplication

Dereplication is the early identification of known compounds present in natural-product extracts before extensive isolation and purification are undertaken. Its primary purpose is to distinguish previously characterized metabolites from potentially novel compounds and thereby reduce the unnecessary re-isolation of known constituents. [5,6]

Conventional dereplication approaches use information obtained from chromatographic, spectroscopic, and spectrometric techniques, including HPLC, UV–visible spectroscopy, mass spectrometry, tandem mass spectrometry, and NMR spectroscopy. These experimental data can be compared with reference databases and spectral libraries to obtain putative compound annotations. The development of large natural-product databases and improved analytical instrumentation has significantly increased the speed and efficiency of dereplication. [6]

Effective dereplication enables researchers to focus isolation and characterization efforts on chemically interesting or potentially novel metabolites. It therefore acts as an important decision-making step between crude extract profiling and targeted compound isolation. However, conventional dereplication may become difficult when extracts contain numerous structurally related metabolites. This limitation created a need for approaches capable of examining relationships among multiple metabolites simultaneously.

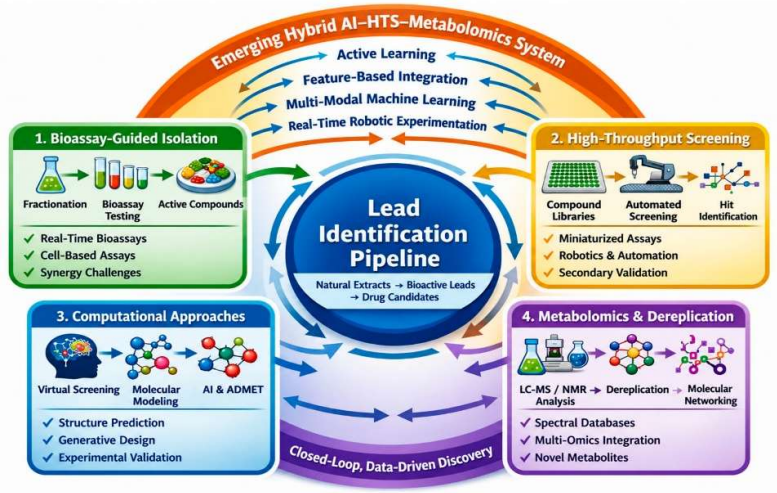

Figure:2 Strategies for lead compounds identification from natural products.

1.4 Emergence of Molecular Networking

The emergence of tandem mass spectrometry (MS/MS)-based molecular networking provided a new approach for investigating complex natural-product mixtures. Molecular networking organizes MS/MS spectra according to their spectral similarity, allowing structurally related or chemically similar molecules to be visualized as connected nodes within a network. In this way, compounds belonging to the same or related chemical families can be recognized even when their structures have not been completely elucidated. [7,8]

The development of the Global Natural Products Social Molecular Networking (GNPS) platform further facilitated the application of molecular networking by providing an open computational environment for processing, visualizing, sharing, and comparing MS/MS data. Molecular networking can connect experimental spectra with reference spectral libraries and can also reveal relationships among unknown metabolites. [8,9]

Molecular networking has consequently become an important strategy for modern natural-product dereplication and chemical discovery. It can assist researchers in recognizing known metabolites, grouping related analogues, prioritizing unexplored chemical families, and directing isolation toward metabolites with potential biological importance. Recent developments have expanded the approach beyond classical molecular networking to include feature-based molecular networking, ion-identity molecular networking, and other computational strategies. [9,10]

Thus, molecular networking represents an important transition from conventional one-compound-at-a-time analysis toward a network-based approach for understanding the chemical diversity of natural-product extracts.

2) Principles of Molecular Networking

2.1 Concept and Working Principle

Molecular networking is a computational approach that uses tandem mass spectrometry (MS/MS) data to visualize relationships among molecules present in complex biological samples. Instead of analyzing each metabolite independently, molecular networking compares MS/MS fragmentation spectra and groups chemically related compounds into a network. In the resulting network, each node generally represents an MS/MS spectrum or molecular feature, while the connections between nodes represent the degree of spectral similarity between them.

The basic workflow begins with the extraction of metabolites from a natural-product sample, followed by chromatographic separation and MS/MS analysis. The resulting spectra are processed computationally and compared with one another. Spectra showing sufficient similarity are connected to form clusters or molecular families. When a spectrum matches a reference spectrum in a spectral library, the corresponding compound can be annotated. Unknown spectra that are closely connected to known compounds can subsequently be prioritized for further investigation.

Molecular networking therefore provides a network-based representation of the chemical composition of an extract, allowing researchers to rapidly recognize known compounds, related analogues, and potentially unexplored metabolites. This is particularly useful for natural-product samples because they often contain highly complex mixtures of structurally related secondary metabolites.

2.2 MS/MS-Based Molecular Similarity

The fundamental principle of molecular networking is the comparison of MS/MS fragmentation patterns. In tandem mass spectrometry, a precursor ion is selected and fragmented to produce characteristic product ions. The resulting MS/MS spectrum provides information related to the structural features of the precursor molecule. Structurally similar molecules often produce partially similar fragmentation patterns because they contain common structural subunits or functional groups[11].

In molecular networking, the MS/MS spectrum of one molecule is compared with spectra obtained from other molecules. A spectral similarity score is calculated using the matching of fragment ions and their relative intensities. When the similarity exceeds a defined threshold, the corresponding nodes are connected by an edge. The strength of the relationship can be represented by the edge score, with stronger spectral similarity generally producing stronger connections.

This approach allows chemically related compounds to be recognized even when their exact structures are not yet known. For example, a known flavonoid and several structurally related flavonoid analogues may form a connected cluster because they share characteristic fragmentation features. Such relationships can guide the investigation of related compounds that may otherwise be difficult to identify individually[12].

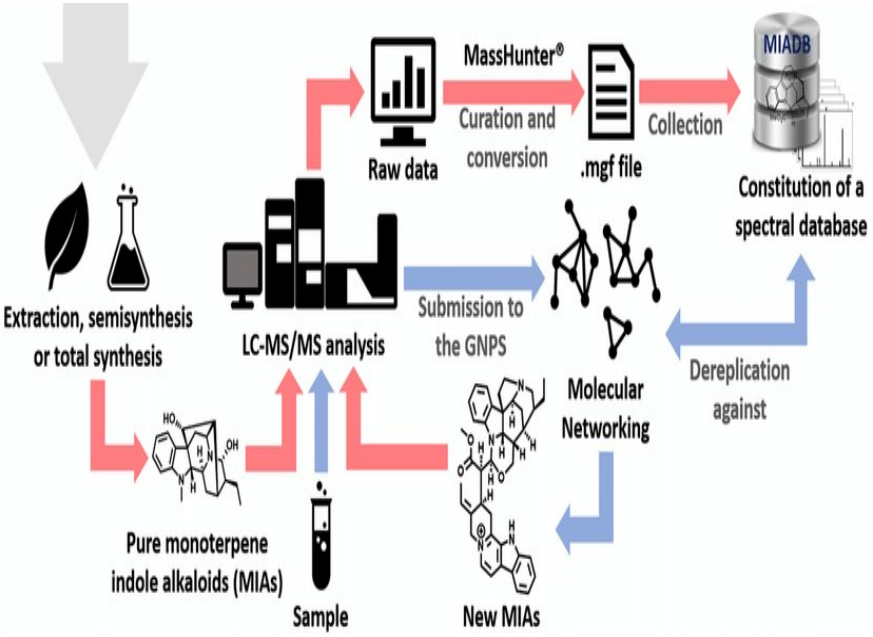

Figure:3 Construction of the MIADB (red arrows) and application in a molecular networking-based dereplication workflow (blue arrows).

However, MS/MS similarity should not be considered definitive proof of structural identity. Isomeric compounds can generate similar spectra, and different structures may occasionally produce related fragmentation patterns. Therefore, molecular-network-based annotations are generally considered putative unless confirmed using authentic standards, orthogonal analytical techniques, or complete spectroscopic characterization.

2.3 Molecular Networks and Chemical Families

The collection of connected MS/MS spectra forms a molecular network, which provides a visual representation of the chemical relationships within a sample. In this network, individual nodes correspond to detected molecular features or MS/MS spectra, whereas edges indicate spectral similarity. Groups of interconnected nodes are referred to as molecular families or molecular clusters[13,14].

Molecular families are particularly valuable in natural-product research because biosynthetic pathways frequently generate groups of structurally related metabolites. A network may therefore reveal a cluster corresponding to a particular chemical class, such as flavonoids, alkaloids, terpenoids, or glycosides. If one member of the cluster is identified through a spectral library, the remaining connected nodes can provide clues about related compounds and can be prioritized for further characterization[15].

The visualization of molecular families also facilitates dereplication and compound prioritization. Known compounds can be recognized rapidly, while clusters containing unidentified or poorly characterized metabolites can be selected for targeted isolation and structural elucidation. Molecular networking therefore changes the traditional approach from

investigating individual peaks sequentially to examining groups of chemically related metabolites as interconnected chemical families.

3) Analytical Techniques and Molecular Networking Workflow

3.1 LC-MS/MS and HRMS

LC-MS/MS is a powerful analytical platform for the investigation of complex natural-product mixtures. Liquid chromatography separates individual metabolites according to their physicochemical properties, while mass spectrometry detects compounds based on their mass-to-charge ratio (*m/z*). In tandem mass spectrometry, precursor ions are fragmented to generate product-ion spectra that provide structural information about the detected metabolites.

High-resolution mass spectrometry (HRMS) provides accurate mass measurements and high resolving power, allowing improved determination of molecular formulas and discrimination between closely related metabolites. LC-HRMS/MS is therefore particularly valuable for untargeted metabolite profiling, dereplication and characterization of structurally diverse natural products. However, accurate mass and MS/MS data alone may not always establish the complete structure of a compound; confirmation may require authentic standards and complementary spectroscopic techniques such as NMR. [16,17]

3.2 Sample Preparation and Data Acquisition

Sample preparation is a critical step in natural-product molecular-networking studies because extraction conditions can strongly influence the chemical profile obtained. Plant or biological materials are generally dried, powdered and extracted using suitable solvents or solvent mixtures according to the polarity and stability of the target metabolites. The extracts are subsequently filtered and appropriately diluted before instrumental analysis.

During LC-MS/MS analysis, chromatographic separation is followed by acquisition of precursor-ion and MS/MS fragmentation data. Different acquisition strategies, including data-dependent acquisition (DDA) and data-independent acquisition (DIA), can be employed depending on the analytical objective. Appropriate blanks, quality-control samples and replicate analyses are important for monitoring background signals and analytical reproducibility. [18,19]

The acquired raw data are subsequently converted into standardized formats and processed using appropriate computational tools. Careful data acquisition and quality control are essential because incomplete or low-quality MS/MS spectra can affect spectral matching and subsequent molecular-network construction.

3.3 Data Processing and Network Generation

Raw LC-MS/MS data require computational processing before molecular-network construction. The initial processing involves peak detection, feature extraction, chromatographic alignment, noise reduction and removal of background signals. In feature-based molecular networking, information such as precursor *m/z*, retention time and feature abundance is retained along with the corresponding MS/MS spectra.

The processed spectra are subjected to spectral comparison based on the similarity of their fragmentation patterns. Spectra with sufficient similarity are connected to form a network. Individual nodes represent detected molecular features, while edges represent relationships between spectra. Groups of interconnected nodes can subsequently be visualized as molecular families, allowing structurally related metabolites to be examined collectively. [19,20]

Computational tools such as MZmine, XCMS and MS-DIAL can be used for feature detection and preprocessing, while molecular-networking platforms can subsequently be used for network construction and visualization. The generated network can be compared with reference spectral libraries to support compound annotation and dereplication. Unknown clusters can then be prioritized for further structural characterization and isolation. [21]

4) Molecular Networking Platforms and Databases

4.1 Global Natural Products Social Molecular Networking (GNPS)

The Global Natural Products Social Molecular Networking (GNPS) is an open-access web-based platform developed for the analysis, visualization, sharing, and annotation of tandem mass spectrometry (MS/MS) data. It has become an important computational resource in natural-product research because it combines molecular-networking analysis with spectral-library searching and community-based data sharing. GNPS allows researchers to upload MS/MS datasets, generate molecular networks, compare experimental spectra with reference spectra, and annotate metabolites within complex biological samples.

One of the major advantages of GNPS is its ability to place experimental MS/MS spectra into a broader chemical context. Spectra with similar fragmentation patterns can be connected within a molecular network, while matches to reference spectra can assist in the annotation of known compounds. This allows researchers to distinguish known metabolites from potentially unexplored molecular families and can support the prioritization of compounds for further investigation[22].

GNPS also promotes community curation and data sharing, allowing researchers to contribute MS/MS spectra and associated metadata to publicly accessible resources. Consequently, datasets deposited by different research groups can contribute to the expansion of spectral knowledge and facilitate the re-analysis of previously generated data.

For natural-product research, GNPS has applications in dereplication, metabolite annotation, chemical profiling, molecular-family characterization, and bioactive compound discovery. The platform has also expanded beyond classical molecular networking by supporting several specialized workflows and annotation tools[23].

4.2 Feature-Based Molecular Networking (FBMN)

Feature-Based Molecular Networking (FBMN) is an advanced molecular-networking approach that incorporates chromatographic feature information into the molecular network. Unlike classical molecular networking, which primarily compares MS/MS spectra, FBMN combines spectral similarity with information such as precursor *m/z*, retention time, feature intensity, and chromatographic peak information.

In FBMN, molecular features are first detected and aligned using suitable data-processing software. The corresponding MS/MS spectra are then used to establish relationships between chemically related features. This approach provides a more detailed representation of the chemical composition of a sample and allows quantitative information to be incorporated into the molecular network[24].

An important advantage of FBMN is its ability to improve the visualization of isomeric and closely related compounds by retaining chromatographic information. It also allows researchers to compare the abundance of molecular features across different samples or experimental conditions. Therefore, FBMN is particularly useful for metabolomics studies, comparative phytochemical profiling, and natural-product dereplication.

FBMN has increasingly become a standard approach for modern molecular-networking studies. Recent reviews also identify FBMN as one of the major developments beyond classical molecular networking, alongside approaches such as ion-identity molecular networking and bioactivity-based molecular networking[25].

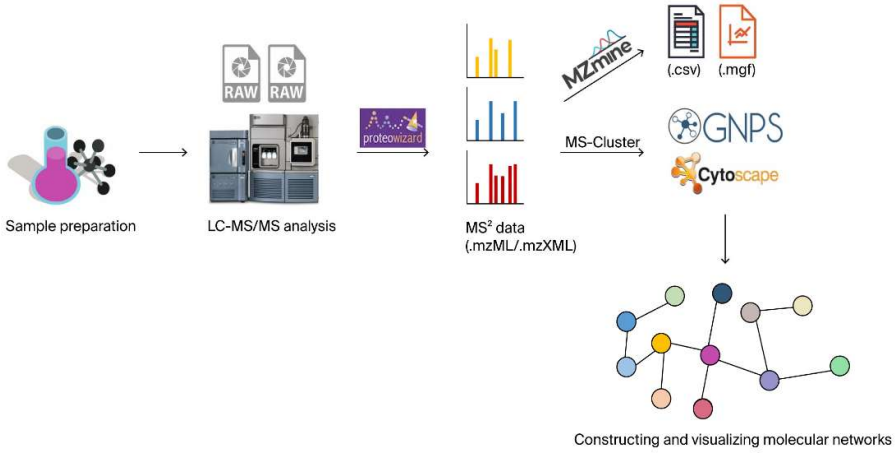

Figure:4 MS/MS-based molecular networking workflow. Colored bar charts represent MS^2 data, recording the fragment ion information of individual molecular features from mass spectrometric analysis. The bar height corresponds to the relative abundance (or peak intensity) of the corresponding fragment ions, with distinct colors distinguishing between sample sources or detection batches. The spherical component represents the constructed molecular network, where each ball denotes a molecular feature node—ball color indicates different compound categories (e.g., alkaloids, flavonoids), and ball size reflects node connectivity (or matching confidence between the molecular feature and reference database entries).

4.3 Spectral Libraries and Databases

Spectral libraries are essential components of molecular networking because they provide reference MS/MS spectra against which experimentally acquired spectra can be compared. A close match between an experimental spectrum and a reference spectrum can support the annotation of a known metabolite and thereby facilitate dereplication.

Several public spectral resources are relevant to natural-product research. MassBank provides experimentally measured mass spectra with associated chemical information, while MassBank of North America (MoNA) is a large, metadata-oriented repository containing experimental and in-silico mass spectral records. ReSpect is particularly relevant to plant research because it contains MS/MS information for phytochemicals[26,27].

Other resources include the Human Metabolome Database (HMDB) and the MassBank European and Japanese resources. GNPS incorporates several external spectral libraries, including MassBank, ReSpect, HMDB, and other specialized libraries, thereby increasing the range of spectra available for molecular-network annotation.

The usefulness of spectral libraries depends strongly on the quality, completeness, and metadata associated with reference spectra. Because many natural products remain poorly represented in existing databases, an unknown compound may fail to obtain a direct library match even when it is biologically or chemically important. Molecular networking can overcome part of this limitation by connecting unidentified spectra to related known compounds, thereby providing structural clues and guiding subsequent isolation and characterization[28].

Thus, GNPS provides the computational environment, FBMN provides an advanced network-building strategy, and spectral libraries provide reference information for metabolite annotation and dereplication. Together, these resources form an important technological framework for modern natural-product research.

5) Molecular Networking for Dereplication

5.1 Identification of Known Compounds

Dereplication is an essential step in natural-product research that aims to recognize known compounds at an early stage, before extensive isolation and structural elucidation are performed. Molecular networking facilitates this process by comparing MS/MS fragmentation spectra obtained from natural-product extracts with reference spectra available in spectral databases. When an experimental spectrum shows a high degree of similarity to a reference spectrum, the corresponding metabolite can be assigned a putative identity. This allows known constituents to be recognized directly within complex mixtures.

In a molecular network, an identified compound can serve as an anchor node for related unidentified metabolites. Structurally related compounds may produce similar fragmentation patterns and therefore appear close to the identified compound within the network. This provides useful structural information even when the related metabolites do not have exact matches in spectral libraries. However, network similarity alone does not establish a definitive structure; final confirmation generally requires authentic standards and/or complementary spectroscopic evidence[29].

Modern dereplication workflows are increasingly combining molecular networking with in-silico MS/MS fragmentation and multiple spectral-similarity algorithms. Such

approaches can expand annotation beyond compounds having experimentally measured reference spectra and improve the ability to recognize structurally related natural products.

5.2 Chemical Family Characterization

One of the important advantages of molecular networking is its ability to organize metabolites into chemical families according to similarities in their MS/MS fragmentation patterns. Compounds sharing common structural features tend to form connected clusters within the network. These clusters can represent groups of related metabolites, such as flavonoids, alkaloids, terpenoids, peptides, or glycosides.

Identification of one or more compounds within a cluster can provide structural clues about the remaining members of the same chemical family. For example, when a known flavonoid glycoside is annotated within a cluster, neighboring nodes may represent related glycosides differing in sugar composition, hydroxylation, methylation, or other structural modifications. Molecular networking therefore allows researchers to study families of metabolites rather than individual compounds in isolation[30,31].

This family-based organization is particularly valuable for complex plant extracts, where numerous analogues may occur simultaneously. It can help researchers recognize patterns of chemical diversity and prioritize molecular clusters containing potentially unexplored metabolites. Recent developments have further expanded molecular networking through approaches such as feature-based molecular networking, ion-identity molecular networking, substructure-based approaches, and alternative spectral-similarity methods.

5.3 Avoiding Rediscovery of Known Metabolites

The repeated isolation of previously reported natural products is a major limitation in conventional natural-product discovery. Extensive purification and structural elucidation of a known metabolite can consume considerable time, resources, and biological material without generating a genuinely new compound. Dereplication therefore provides an important strategy for reducing this rediscovery problem.

Molecular networking can identify known compounds and their analogues at an early stage of the investigation. Once a known molecular family has been recognized, researchers can make an informed decision about whether further purification of that chemical family is necessary. Greater attention can instead be directed toward unannotated clusters, unusual molecular features, or chemically distinct metabolites that may have a higher potential for novelty[32,33,34].

Recent integrated approaches have further improved this strategy. For example, the Integrated Molecular Networking Workflow for Natural Product Dereplication (IMN4NPD) combines multiple computational tools and spectral-similarity methods to improve the annotation of natural products, including small clusters and individual nodes that may be overlooked by conventional workflows. Such approaches demonstrate the continuing development of molecular networking from a simple visualization method toward a more comprehensive dereplication platform.

6) Applications in Pharmacognosy and Herbal Drug Research

6.1 Phytochemical Profiling

Phytochemical profiling is an important component of pharmacognostic evaluation because medicinal plants contain complex mixtures of primary and secondary metabolites. Conventional phytochemical analysis generally focuses on individual constituents or selected marker compounds, whereas LC-MS/MS-based molecular networking provides a broader representation of the chemical composition of a plant extract. It enables the simultaneous visualization of numerous metabolites and their structural relationships within a complex sample.

In molecular networking, metabolites with similar MS/MS fragmentation patterns are grouped into molecular families. This allows related phytochemicals such as flavonoids, alkaloids, terpenoids, phenolic compounds and glycosides to be recognized as interconnected groups. Annotation of one or more compounds within a molecular family can provide structural information about other related metabolites that may not have direct spectral-library matches[35].

Molecular networking can also be used to compare the phytochemical profiles of different plant species, plant parts, geographical samples, developmental stages or extracts. Differences in the distribution and abundance of molecular families can provide characteristic chemical fingerprints for individual botanical materials. Such comprehensive profiling can therefore complement conventional pharmacognostic and phytochemical investigations.

A recent 2026 study demonstrated the potential of integrating GNPS molecular networking with SIRIUS-based structural annotation to compare medicinal plants and identify characteristic metabolite fingerprints with potential chemotaxonomic relevance.

6.2 Chemotaxonomy

Chemotaxonomy uses the distribution of characteristic chemical constituents to support the classification and identification of plants. Secondary metabolites such as alkaloids, flavonoids, terpenoids, phenolics and other specialized metabolites can serve as chemical characteristics for distinguishing closely related taxa.

Molecular networking can strengthen chemotaxonomic studies by providing a systematic comparison of metabolite patterns between species or genera. Instead of relying on a limited number of marker compounds, molecular networking enables researchers to examine multiple chemical families simultaneously. Species-specific or genus-specific molecular features can therefore be identified and used as potential chemotaxonomic markers.

Comparative molecular networking can also reveal shared and unique metabolites between closely related medicinal plants. Shared molecular families may reflect common biosynthetic pathways, whereas species-specific metabolites may provide useful chemical characteristics for taxonomic discrimination. Integration of molecular networking with in-silico structural annotation can further improve the identification of metabolites with potential chemotaxonomic significance[36].

Thus, molecular networking provides a modern approach to chemotaxonomy by connecting phytochemical diversity with botanical classification, thereby complementing morphological, anatomical and genetic methods traditionally used in pharmacognosy.

6.3 Authentication and Quality Control

Authentication is essential to ensure that the correct medicinal plant species or botanical ingredient is used in herbal medicines. Substitution, adulteration, geographical variation and differences in processing can alter the chemical composition of herbal materials. Conventional authentication methods based on morphology and microscopy may become difficult when plant materials are powdered, extracted or incorporated into finished formulations.

LC-MS/MS-based molecular networking can provide a chemical fingerprint of botanical materials, allowing the metabolite composition of an unknown sample to be compared with that of an authenticated reference material. Characteristic molecular families and metabolite patterns can help distinguish genuine materials from closely related species or adulterants. Studies have demonstrated the use of LC-MS/MS molecular networking for both botanical authentication and characterization of commercial botanical supplements[37,38].

Molecular networking can also complement DNA-based authentication, chromatographic fingerprinting and chemometric analysis. While DNA methods primarily provide information about biological identity, molecular networking provides information about the chemical composition of the material. Combining these approaches can therefore provide complementary evidence for botanical authentication.

For quality control, molecular networking can help detect differences between batches, geographical sources, processing conditions and storage conditions by comparing their metabolite profiles. Such approaches are particularly useful for complex herbal products in which quality cannot always be adequately represented by a single chemical marker[39].

6.4 Herbal Drug Standardization

Standardization aims to ensure consistent identity, quality, composition and safety of herbal drugs and their preparations. Conventional standardization often relies on pharmacopoeial parameters, physicochemical tests, chromatographic fingerprints and selected chemical markers. However, because herbal medicines contain numerous constituents that may contribute collectively to their biological effects, analysis based on a single marker may not always provide a comprehensive assessment of quality.

Molecular networking offers a holistic chemical profiling approach that can complement conventional standardization. Molecular features and molecular families detected in authenticated herbal materials can be used to establish characteristic chemical fingerprints. These fingerprints may subsequently be compared between batches to evaluate chemical consistency and identify abnormal or unexpected metabolite patterns[40,41].

The integration of molecular networking with metabolomics and multivariate data analysis can further support the evaluation of batch-to-batch variation, geographical origin, processing effects, adulteration and chemical markers. Metabolomics-based quality-control

approaches have increasingly been investigated because they consider multiple metabolites rather than relying exclusively on one or two indicator compounds.

7) Conclusion

Molecular networking has emerged as a valuable modern strategy for addressing several major challenges associated with natural-product research. By organizing MS/MS spectra according to their structural and spectral similarities, it enables researchers to examine complex metabolite mixtures as interconnected molecular families rather than as individual compounds. This network-based approach facilitates rapid dereplication, recognition of known metabolites, characterization of related chemical families, and prioritization of unexplored molecular clusters for further investigation.

The integration of LC-MS/MS and HRMS with computational platforms such as GNPS and feature-based molecular networking has significantly improved the efficiency of phytochemical profiling and metabolite annotation. In pharmacognosy and herbal-drug research, molecular networking can complement conventional approaches by supporting phytochemical profiling, chemotaxonomy, botanical authentication, quality control, and herbal-drug standardization. Its ability to compare chemical fingerprints across species, plant parts, geographical sources, batches, and processing conditions provides a broader perspective on the chemical consistency of natural materials.

Despite these advantages, molecular networking should not be considered a replacement for conventional isolation, purification, and structural elucidation. Spectral similarity provides valuable structural clues, but definitive identification may require authentic standards and complementary techniques such as NMR and other orthogonal analytical methods. In addition, incomplete spectral libraries and limited representation of rare natural products remain important challenges.

Future integration of molecular networking with metabolomics, in-silico fragmentation, chemometrics, artificial intelligence, and expanded spectral databases is expected to further improve metabolite annotation and prioritization. Overall, molecular networking provides a powerful complementary framework for systematic exploration of natural-product chemical diversity and can contribute substantially to efficient dereplication and bioactive compound discovery.

ACKNOWLEDGEMENT:

I’m very thankful to Department of Pharmacognosy. The Tamil Nadu Dr.M.G.R Medical University, Sasurie College of Pharmacy. I would also like to thank the Management, for providing the necessary facilities to carry out this work.

CONFLICT OF INTEREST:

We declare that we have no conflict of interest.

BIBLIOGRAPHY

1) Atanasov AG, Zotchev SB, Dirsch VM, Supuran CT, International Natural Product Sciences Taskforce. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov. 2021;20:200-216.
2) Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. J Nat Prod. 2020;83(3):770-803.
3) Harvey AL, Edrada-Ebel R, Quinn RJ. The re-emergence of natural products for drug discovery in the genomics era. Nat Rev Drug Discov. 2015;14:111-129.
4) Wolfender JL, Nuzillard JM, van der Hooft JJJ, Renault JH, Bertrand S. Accelerating metabolite identification in natural product research: toward an ideal combination of liquid chromatography-high-resolution tandem mass spectrometry profiling and automated structure annotation. Nat Prod Rep. 2019;36:855-868.
5) Demarque DP, Dusi RG, de Sousa FDM, Grossi SM, Silvério MRS, Lopes NP, et al. Mass spectrometry-based dereplication strategies for the rapid identification of bioactive compounds from natural products. Nat Prod Rep. 2020;37
6) Allard PM, Péresse T, Bisson J, Gindro K, Marcourt L, Pham VC, et al. Integration of molecular networking and in silico MS/MS fragmentation for natural product dereplication. Nat Prod Rep. 2016;33
7) Nguyen DD, Wu CH, Moree WJ, Lamsa A, Medema MH, Zhao X, et al. MS/MS networking guided analysis of molecule and biosynthetic diversity. Nat Chem Biol. 2013;9
8) Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol. 2016;34:828-837.
9) Wang Y, Wang Y, Zhang Z, Xu K, Fang Q, Wu X, Ma S. Molecular networking: An efficient tool for discovering and identifying natural products. J Pharm Biomed Anal. 2025;259:116741.
10) Nothias LF, Petras D, Schmid R, Dührkop K, Rainer J, Sarvepalli A, et al. Feature-based molecular networking in the GNPS analysis environment. Nat Methods. 2020;17:905-908.
11) Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology.* 2016;34:828–837.
12) Nguyen DD, Wu CH, Moree WJ, Lamsa A, Medema MH, Zhao X, et al. MS/MS networking guided analysis of molecule and biosynthetic diversity. *Nature Chemical Biology.* 2013;9:300–305.
13) Nothias LF, Nothias-Esposito M, da Silva R, Wang M, Protsyuk I, Zhang Z, et al. Bioactivity-based molecular networking for the discovery of chemical relationships and bioactive natural products. *Nature Chemical Biology.* 2018;14:1098–1104.
14) Nothias LF, Petras D, Schmid R, Dührkop K, Rainer J, Sarvepalli A, et al. Feature-based molecular networking in the GNPS analysis environment. *Nature Methods.* 2020;17:905–908.

15) Wolfender JL, Nuzillard JM, van der Hooft JJJ, Renault JH, Bertrand S. Accelerating metabolite identification in natural product research: toward an ideal combination of liquid chromatography-high-resolution tandem mass spectrometry profiling and automated structure annotation. *Natural Product Reports.* 2019;36:855–868.
16) Kind T, Fiehn O. Advances in structure elucidation of small molecules using mass spectrometry. *Bioanal Rev.* 2010;2:23–60.
17) Cajka T, Fiehn O. Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics. *Anal Chem.* 2016;88:524–545.
18) Blaženović I, Kind T, Ji J, Fiehn O. Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. *Metabolites.* 2018;8:31.
19) Pluskal T, Castillo S, Villar-Briones A, Orešič M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics.* 2010;11:395.
20) Tsugawa H, Cajka T, Kind T, et al. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods.* 2015;12:523–526.
21) Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem.* 2006;78:779–787.
22) Dorrestein PC, Gentry EC, Gonzalez DJ, et al. Molecular networking: An enabling technology for natural product research. *Natural Product Reports.* 2016;33:1–15.
23) Aron AT, Gentry EC, McPhail KL, et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nature Protocols.* 2020;15:1954–1991.
24) Blaženović I, Kind T, Ji J, Fiehn O. Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. *Metabolites.* 2018;8:31.
25) Sud M, Fahy E, Cotter D, et al. Metabolomics Workbench: an international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training. *Nucleic Acids Research.* 2016;44:D463–D470.
26) Horai H, Arita M, Kanaya S, et al. MassBank: a public repository for sharing mass spectral data for life sciences. *Journal of Mass Spectrometry.* 2010;45:703–714.
27) Hummel J, Strehmel N, Selbig J, Walther D, Kopka J. Decision tree supported substructure prediction of 1H NMR spectra of complex mixtures. *Phytochemistry.* 2010;71:7–15.
28) Matsuda F, Okuzaki D, Nakabayashi R, et al. MS/MS spectral database for phytochemicals: ReSpect. *Plant Journal.* 2016;
29) Yang JY, Sanchez LM, Rath CM, et al. Molecular networking as a dereplication strategy. *J Nat Prod.* 2013;76(9):1686–1699. doi:10.1021/np400413s.
30) Qin GF, Zhang X, Zhu F, et al. MS/MS-Based Molecular Networking: An Efficient Approach for Natural Products Dereplication. *Molecules.* 2023;28(1):157. doi:10.3390/molecules28010157.
31) Allard PM, Péresse T, Bisson J, et al. Integration of molecular networking and in-silico MS/MS fragmentation for natural product dereplication. *Anal Chem.* 2016;88(6):3317–3323. doi:10.1021/acs.analchem.5b04804.

32) Wang Y, Wang Y, Zhang Z, et al. Molecular networking: An efficient tool for discovering and identifying natural products. *J Pharm Biomed Anal.* 2025;259:116741.
33) IMN4NPD: An Integrated Molecular Networking Workflow for Natural Product Dereplication. *Anal Chem.* 2024;96(7):2990–2997. doi:10.1021/acs.analchem.3c04746.
34) Gaudêncio SP, et al. Advanced Methods for Natural Products Discovery: Bioactivity Screening, Dereplication, Metabolomics Profiling, Genomic Sequencing, Databases and Informatic Tools, and Structure Elucidation. *Mar Drugs.* 2023;21(5):308. doi:10.3390/md21050308.
35) Jouaneh TMM, Motta N, Wu C, Coffey C, Via CW, Kirk RD, Bertin MJ. Analysis of botanicals and botanical supplements by LC-MS/MS-based molecular networking: Approaches for annotating plant metabolites and authentication. *Fitoterapia.* 2022;159:105200. doi:10.1016/j.fitote.2022.105200.
36) Li S, Han Q, Qiao C, Song J, Cheng CL, Xu H. Chemical markers for the quality control of herbal medicines: an overview. *Chin Med.* 2008;3:7. doi:10.1186/1749-8546-3-7.
37) Lo-Coco F, et al. Quality control of herbal medicines: From traditional techniques to state-of-the-art approaches. *Planta Med.* 2021. doi:10.1055/a-1529-8339.
38) Ji P, Yang X, Zhao X. Application of metabolomics in quality control of traditional Chinese medicines: a review. *Front Plant Sci.* 2024;15:1463666. doi:10.3389/fpls.2024.1463666.
39) Liu C, Zuo Z, Xu F, Wang Y. Authentication of herbal medicines based on modern analytical technology combined with chemometrics approach: A review. *Crit Rev Anal Chem.* 2023;53(7):1393–1418. doi:10.1080/10408347.2021.2023460.
40) Brinckmann JA, et al. Chemical authentication of botanical ingredients: A review of commercial herbal products. *Front Pharmacol.* 2021.
41) Zhu L, Xue X, Pan J, et al. Comparative phytochemical profiling of three medicinal Polygonaceae plants using integrated molecular networking and SIRIUS-guided annotation: Implications for chemotaxonomy. *J Chromatogr A.* 2026;[article 467402]. doi:10.1016/j.chroma.2026.467402.