

Review

# MS/MS-Based Molecular Networking: An Efficient Approach for Natural Products Dereplication

Guo-Fei Qin <sup>1,\*</sup>, Xiao Zhang <sup>2,†</sup>, Feng Zhu <sup>1</sup>, Zong-Qing Huo <sup>1</sup>, Qing-Qiang Yao <sup>3</sup>, Qun Feng <sup>1</sup>, Zhong Liu <sup>1</sup>, Gui-Min Zhang <sup>1,2</sup>, Jing-Chun Yao <sup>1,\*</sup> and Hong-Bao Liang <sup>1,2,\*</sup>

<sup>1</sup> State Key Laboratory of Generic Manufacture Technology of Chinese Traditional Medicine, Lunan Pharmaceutical Group Co., Ltd., Linyi 273400, China

<sup>2</sup> College of Pharmacy, Shandong University of Traditional Chinese Medicine, Jinan 250355, China

<sup>3</sup> Jining Medical University, Jining 272067, China

\* Correspondence: qinguoifei@126.com (G.-F.Q.); yaojingchun@lunan.cn (J.-C.Y.);

lianghongbao1985@163.com (H.-B.L.); Tel.: +86-539-503-0319 (G.-F.Q.)

† These authors contributed equally to this work.

**Abstract:** Natural products (NPs) have historically played a primary role in the discovery of small-molecule drugs. However, due to the advent of other methodologies and the drawbacks of NPs, the pharmaceutical industry has largely declined in interest regarding the screening of new drugs from NPs since 2000. There are many technical bottlenecks to quickly obtaining new bioactive NPs on a large scale, which has made NP-based drug discovery very time-consuming, and the first thorny problem faced by researchers is how to dereplicate NPs from crude extracts. Remarkably, with the rapid development of omics, analytical instrumentation, and artificial intelligence technology, in 2012, an efficient approach, known as tandem mass spectrometry (MS/MS)-based molecular networking (MN) analysis, was developed to avoid the rediscovery of known compounds from the complex natural mixtures. Then, in the past decade, based on the classical MN (CLMN), feature-based MN (FBMN), ion identity MN (IIMN), building blocks-based molecular network (BBMN), substructure-based MN (MS2LDA), and bioactivity-based MN (BMN) methods have been presented. In this paper, we review the basic principles, general workflow, and application examples of the methods mentioned above, to further the research and applications of these methods.

**Keywords:** MS/MS-based molecular networking; natural products dereplication; classical MN (CLMN); feature-based molecular networking (FBMN); ion identity molecular networking (IIMN); building blocks-based molecular network (BBMN); substructure-based MN (MS2LDA); bioactivity-based molecular networking (BMN)



**Citation:** Qin, G.-F.; Zhang, X.; Zhu, F.; Huo, Z.-Q.; Yao, Q.-Q.; Feng, Q.; Liu, Z.; Zhang, G.-M.; Yao, J.-C.; Liang, H.-B. MS/MS-Based Molecular Networking: An Efficient Approach for Natural Products Dereplication. *Molecules* **2023**, *28*, 157. <https://doi.org/10.3390/molecules28010157>

Academic Editor: Ana Estévez-Braun

Received: 8 November 2022

Revised: 13 December 2022

Accepted: 20 December 2022

Published: 24 December 2022



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As the result of millions of years of evolutionary optimization, natural products (NPs) have been endowed with privileged pharmacological functions, and historically became the most important source for drug discovery [1–3]. Among the 1394 small-molecule drugs approved by the United States Food and Drugs Administration (FDA) from 1981 to 2019, 31.6% were botanical drugs, unaltered NPs, and NP derivatives, and 30.4% were synthetic drugs with NP pharmacophores or the mimicry of NPs [4], which means that close to 2/3 of the small-molecule medicines of this period were associated with NPs. Meanwhile, in the top 200 pharmaceuticals by retail sales in 2021, NP-derived medicines were successful in the areas of antibiotics and antifungal, anticancer, cholesterol-lowering, immunosuppression, and antihypertensive properties [5,6]. Despite such tremendous success, it is noticeable that many pharmaceutical companies have terminated their programs to screen new chemical entities from NPs since 2000 [7–9]. The reasons given were the rapid advances of biopharmaceuticals [10], kinase-based drugs [11], antibody-drug conjugates (ADC) [12], proteolysis-targeting chimeras (PROTAC) [13], and other methodologies [14,15]. However,

no fundamental breakthroughs to overcome the drawbacks of NPs have been made for some time [5,7–9], especially in terms of rapidly screening new and bioactive NPs from complex extracts; economically obtaining sufficient quantities of pure target compounds was the less widely advertised reason [8,9,16].

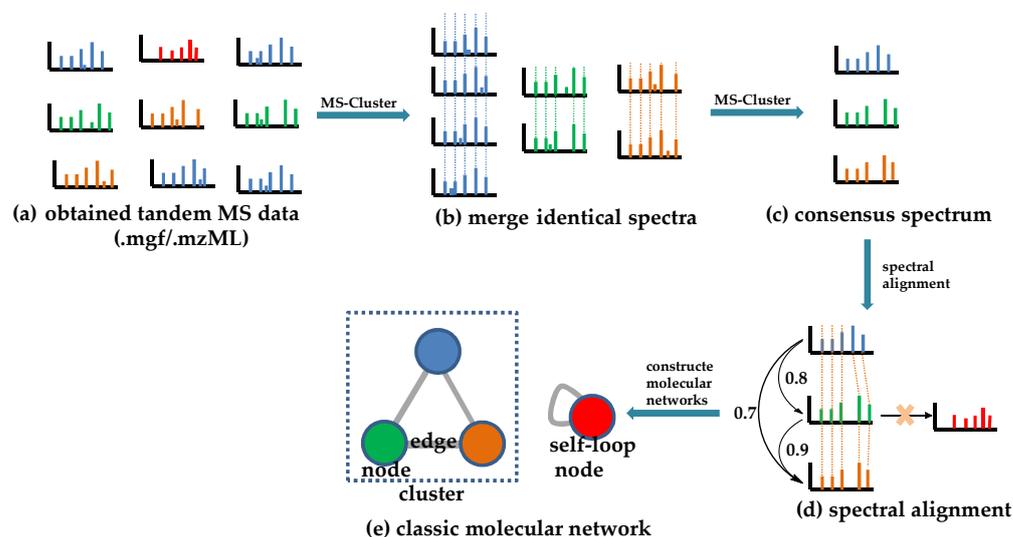
Indeed, with the large and increasing number of NPs (estimated at 600,000), rediscovery was commonplace in natural product research [16–19], and consequently, the problem of how to rapidly identify new NPs from complex mixtures has become a thorny challenge that needs to be resolved [18,19]. To circumvent this issue, a number of early prioritization strategies were achieved by manually comparing characteristics such as the ultraviolet-visible spectra (UV/Vis spectra), nuclear magnetic resonance (NMR), or mass spectra (MS) with various databases [20–22], or by tracking biological activity and other methods [23–25]. In practice, these methods were also accompanied by laborious, time-consuming procedures and high rediscovery rates [16]. With the recent rapid advances in analytical instrumentation and artificial intelligence technology, proteomics [26–28], genomics [29], metabolomics [30], and transcriptomics [31,32] have enabled tremendous achievements that greatly promoted and influenced the development of life sciences. In the past decade, the research method and technology of metabolomics and proteomics were also borrowed to prioritize the targeted isolates of NPs [33–35]. Since a major bottleneck in the omics pipeline is the annotation and identification of the spectral data, many spectral interpretation methods, such as MS- and/or NMR-based approaches, were developed [36–39]. Among them, tandem mass spectrometry (MS/MS)-based molecular networking (MN) has become an increasingly popular and attractive NPs research tool that integrated the advantages of sensitiveness, high throughput, and the robustness of MS/MS with the ability of MN to organize and visualize large MS/MS datasets [37,38].

The classical MS/MS-based MN (CLMN) was first reported by Dorrestein's group in 2012 to investigate the metabolic profiles of living microbial colonies [40]. In 2013, the Global Natural Product Social Molecular Networking (GNPS, <http://gnps.ucsd.edu>) group, a sharing and community-based web platform used to store, analyze, share, and compare MS/MS data, as well as perform the generation of MN [41,42], further promoted the development of MN in different research groups. Currently, this technique is widely used, such as in the study of forensics [43], food chemistry [44], environmental science [45], plant science [46], and others [47,48]. In natural product research, on the basis of CLMN, the feature-based MN (FBMN) [42], ion-identifying MN (IIMN) [49], building blocks-based molecular network (BBMN) [50], substructure-based MN (MS2LDA) [51], and bioactive MN (BMN) [52] were presented in the form of different interpreting methods of the obtained data. In this paper, we review the basic principles, general workflow, and application examples of the above-mentioned methods, aiming to promote further research and applications.

## 2. Classical Molecular Networking (CLMN)

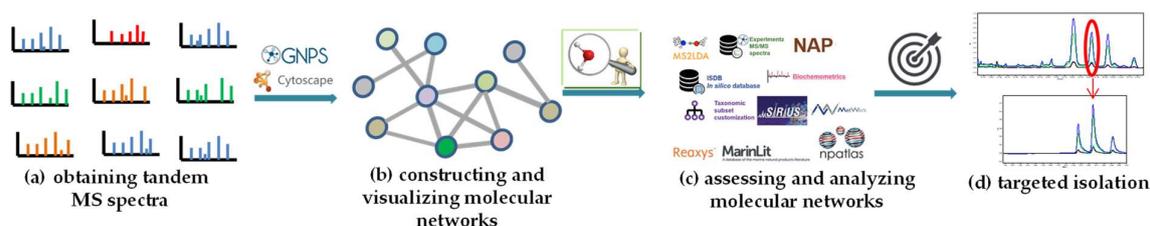
The theoretical rationale of CLMN is that molecules with similar structures will exhibit considerable similarities in their MS/MS spectra, and vice versa. Thus, similar molecules in complex mixtures can be clustered to form “molecular families” by the mass spectral similarities of molecules. The spectral similarities can be calculated with a vector-based modified “cosine score” (ranging from 0 to 1; the higher the score is, the more similar the result will be), which takes into account the number of matching fragment ions, the relative intensities of the peaks, and the parent mass accuracy [53]. As shown in Figure 1, the obtained tandem MS spectra (Figure 1a) are first processed to give a consensus spectrum (Figure 1c) by merging identical spectra (Figure 1b), using the MS-Cluster algorithm to avoid identical spectra appearing more than once [25,53]. Then, a modified algorithm is used to calculate the spectral similarity score (Figure 1d). Peaks from one consensus spectrum are compared with peaks from the other, either at identical  $m/z$  values or with their  $\Delta m/z$ , considering that a  $\Delta m/z$  change to the precursor ion may lead to shifting a

subset of fragment peaks by  $\Delta m/z$  [36,53]. Finally, a molecular network is constructed on the basis of the calculated spectral similarity score (Figure 1e). In the network, the “molecular families” and the “molecular only similar with itself” variables are represented by “cluster” and “self-loop node”, respectively. In the cluster, similar molecules (“node”) are connected by lines (“edge”), and the thickness of the edges showcases the level of their similarity [53].



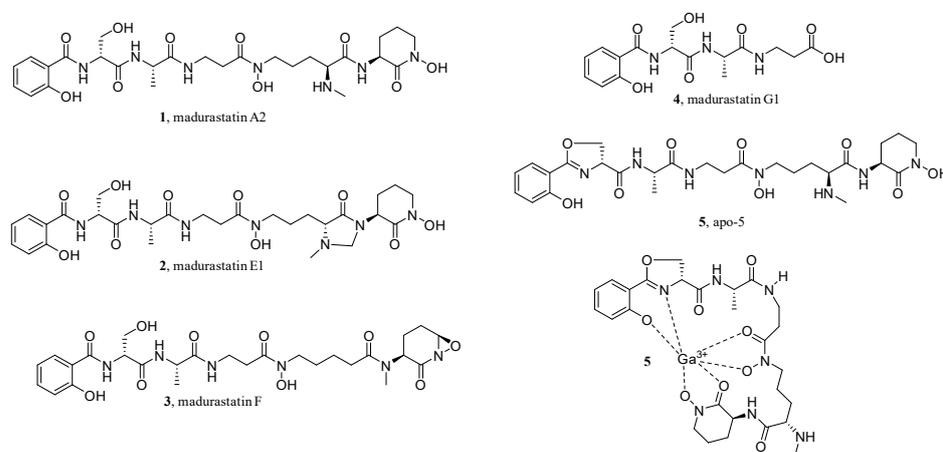
**Figure 1.** Schematic representation of the principle of CLMN. (a) The obtained tandem MS data. (b) The merging of identical spectra. (c) The consensus spectrum. (d) Spectral alignment. (e) The classic molecular network.

A schematic workflow for a CLMN dereplication pipeline is presented in Figure 2. There are four main steps: (Figure 2a) obtaining the tandem MS spectra; (Figure 2b) constructing and visualizing the molecular networks; (Figure 2c) assessing and analyzing the molecular networks; (Figure 2d) targeted isolation (Figure 2) [25]. As the tandem mass spectrometry experiments for data acquisition represent one of the most important factors affecting molecular networks, all samples should be prepared and analyzed in the same way [24]. After uploading the obtained tandem MS spectra to the GNPS platform, the completed job can be visualized either in the platform [41] or in Cytoscape [54]. A detailed protocol from the tandem mass spectrometry experiments, via a publishable and reproducible molecular network in the GNPS platform, has been provided by Dorrestein et al. in *Nature Protocols* [53], and we can refer to this protocol here. Another tool to generate and visualize molecular networks is the MetGem software (<https://metgem.github.io>, accessed on 7 November 2022) [55], which was developed based on the t-distributed stochastic neighbor embedding (t-SNE) algorithm in 2018, a well-known visualization technique used for high-dimensional data [56]. The t-SNE-based MetGem allows clustering spectra, relying on local details within the entire data space rather than individual links between spectra, and can thereby avoid having too many “self-loop nodes” or fusing “molecular families” in the molecular networking when a similarity cutoff is set [36]. However, the t-SNE-based method could not offer information about the relationships between “nodes”, and it is complementary to the cosine similarity-based classic GNPS-style MN [36]. More recently, deep learning mass spectral similarity scoring methods have also been developed, such as Spec2Vec and MS2DeepScore, which derive abstract spectral embeddings to assess spectral similarity by learning the fragment relationships among large amounts of spectral data [57,58]. Of course, mismatches between the calculated mass spectral similarity scores and the true structural similarities are very common, and the comprehensive use of multiple methods can reduce those mismatches.



**Figure 2.** Schematic workflow for CLMN in NP dereplication. (a) Obtaining the tandem MS spectra. (b) Constructing and visualizing molecular networks. (c) Assessing and analyzing molecular networks. (d) Targeted isolation.

As CLMN can visualize and map the chemical space in the extracts of organisms, it is widely used to determine the preferred species [59,60], the culture conditions of microorganisms [60], isolation workflow [61], etc. For example, in searching for siderophores from *Actinomadura* sp. RB99, activity assays and MS/MS-based CLMN were used as the dereplication strategies [59]. First, after co-culturing with *Pseudoxylaria* sp. X802, the extracts obtained from the colony of RB99 and the interaction zone of inhibition, as well as RB99 cultures grown on different media, were analyzed by an high-resolution electrospray ionization tandem mass spectrometry (ESI-HRMS<sup>2</sup>)-based MN. The obtained GNPS network suggested chemical diversity and dereplicated clusters of phosphoethanolamines, phosphocholines, oligosaccharides, pseudoxylallemycins, and cytochalasins, together with an interesting small GNPS cluster. Further analysis of the proposed molecular formulas of the interesting small cluster indicated structural changes of -O and -CH<sub>2</sub> and a peptidic backbone, with a putative *N,O*-ratio characteristic for siderophores. Then, based on these findings and the optimized cultivation conditions, the up-scaled refermentation of RB99 led to the isolation of five new madurastatin derivatives (1–5), including a siderophore-metal complex (5) (Figure 3).

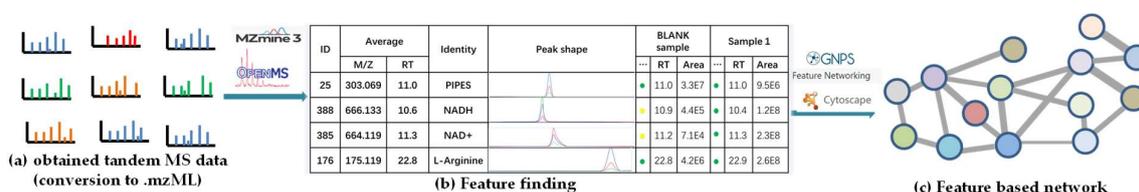


**Figure 3.** Structures of new madurastatin derivatives.

### 3. Feature-Based Molecular Networking (FBMN)

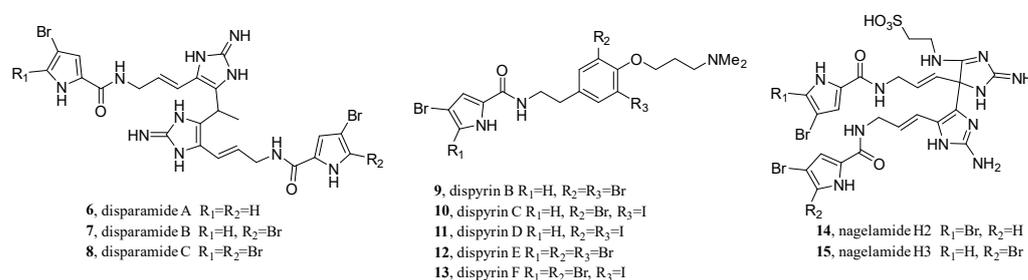
Although CLMN is very convenient for the rapid processing of large-scale MS/MS datasets, it cannot differentiate positional isomers or stereoisomers, or provide accurate relative quantitative information, due to the limitations of the MS-Cluster algorithm. To address this issue, Dorrestein's group developed FBMN by integrating comparative metabolomics with MN in 2017 [42]. In this method, not only the fragmentation data but also the isotope patterns, the retention times, and the ion mobility spectrometry can be compared. Compared with CLMN, there are two main different steps in the workflow of FBMN (Figure 4). First, the obtained tandem MS spectra (Figure 4a) should be pre-processed using MZmine [62], OpenMS [63], or other feature detection and alignment tools [42] to detect, group, and align those features (Figure 4b). Second, the exported feature lists ((.cvs, feature

quantification table) and (.mgf, MS<sup>2</sup> spectral summary file)) are uploaded to perform the dedicated feature networking workflow on the GNPS platform, to generate a feature-based network (Figure 4c) [42].



**Figure 4.** Schematic representation of the principles of FBMN. (a) The obtained tandem MS data. (b) Feature-finding. (c) The feature-based network.

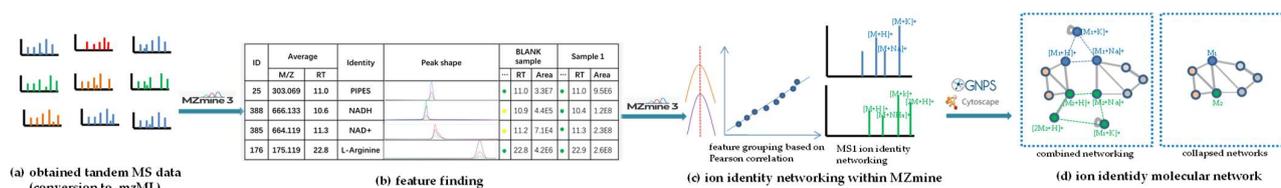
Limited by the chromatographic feature-finding tools and different experimental conditions, FBMN is especially suitable for one or a few samples and has become the second most commonly used tool in GNPS [42]. In revisiting the bromopyrrole alkaloids of the extensively investigated marine sponge, *Agelas dispar*, FBMN was used by Berlinck's team as the dereplication strategy [64]. After separation by extraction and C<sub>8</sub> RP column chromatography, the defatted EtOH/MeOH extract of *A. dispar* was divided into five fractions. Then, three fractions with brominated compounds were subjected to Sephadex LH-20 to yield 63 fractions, which were further analyzed by quadrupole time of flight (QToF)-MS/MS to generate a feature-based molecular network. Finally, after dereplication with the in-house and in silico database (ISDB), clusters of undescribed compounds were selected for study; this resulted in the isolation of disparamides A–C (6–8, with a novel carbon skeleton) and seven other new compounds (9–15) (Figure 5).



**Figure 5.** Structures of disparamides A–C and seven other new compounds.

#### 4. Ion Identity Molecular Networking (IIMN)

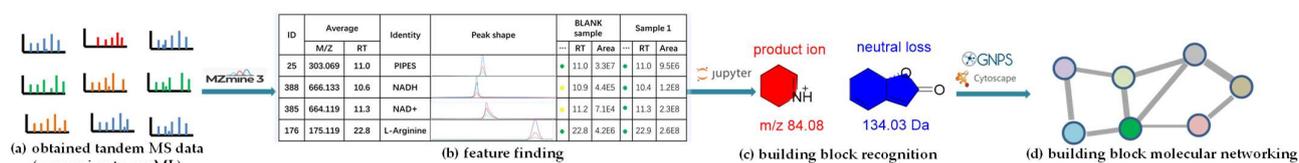
In 2021, Dorrestein's group further employed IIMN to overcome the disconnected sub-networks of "molecular families", which was caused by limitations inherent to the different fragmentation behavior of the multiple ion species of a given compound [49]. For example, the two ion species of the same molecule, [M + H]<sup>+</sup> and [M + Na]<sup>+</sup>, typically stay unconnected in a molecular network. As all ions from the same molecule can be connected, based on the known mass differences of ion species, an additional MS<sup>1</sup> ion identity networking layer was fused to FBMN to create IIMN networks (Figure 6). Compared with FBMN, in the progress of detecting and aligning features with MZmine, one more feature list (a .csv file of ion identity networking (Figure 6c)) should also be exported. Although those three feature files are also uploaded to the FBMN, the resulting data of the IIMN are different from those of the FBMN; two .graphml networking files containing three networks (FBMN networks, IIMN networks, and collapsed IIMN networks (Figure 6d)) will be generated. Recently, this method, combined with activity profiling, was used for the prioritization of compounds inhibiting AK strain-transforming kinase (AKT) in human melanoma cells [65].



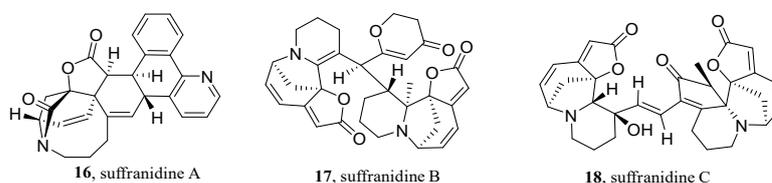
**Figure 6.** Schematic representation of the principle of IIMN. (a) The obtained tandem MS data. (b) Feature-finding. (c) Ion identity networking within MZmine. (d) The ion identity molecular network.

## 5. Building Blocks-Based Molecular Network (BBMN)

Natural products are usually formed by biosynthetic pathways from specific building blocks. In 2021, Ye's group first presented a BBMN to facilitate the efficient discovery of novel *securinega* alkaloids from *Flueggea suffruticosa* [50]. This method combined the neutral loss/product ion-scanning strategy [66] and MN, which can selectively identify biogenetically relevant secondary metabolites and optimize the bulky MS<sup>2</sup> data by further filtering the features files in the pre-processing of the MS dataset (Figure 7). In this research, the EtOH extract from the twigs and leaves of *F. suffruticosa* was acidified with 10% HCl, extracted using CH<sub>2</sub>Cl<sub>2</sub>, basified using NH<sub>4</sub>OH, and re-extracted with CH<sub>2</sub>Cl<sub>2</sub> to yield a total alkaloid. The acquired and deconvoluted ultra high performance liquid chromatography (UHPLC)-MS<sup>2</sup> data (7a) of the total alkaloid comprised features extracted with MZmine (Figure 7b). Then, *m/z* at 84.08 and 134.03 Da were selected as the product ion and the neutral loss, to detect the building blocks of *securinega* alkaloids from the obtained features (Figure 7c), and were then filtered using a Python script, respectively. Subsequently, the BBMNs were constructed by GNPS and MetGem (Figure 7d) and annotated by an in-house library. Then, three unknown and unclustered nodes with *m/z* values over 400 were selected as the target compounds. Finally, LC-MS-guided isolation led to the isolation of three novel *securinega* alkaloids (**16–18**) (Figure 8).



**Figure 7.** Schematic representation of the principle of BBMN. (a) The obtained tandem MS data. (b) Feature-finding. (c) Building-block recognition. (d) Building-block molecular networking.

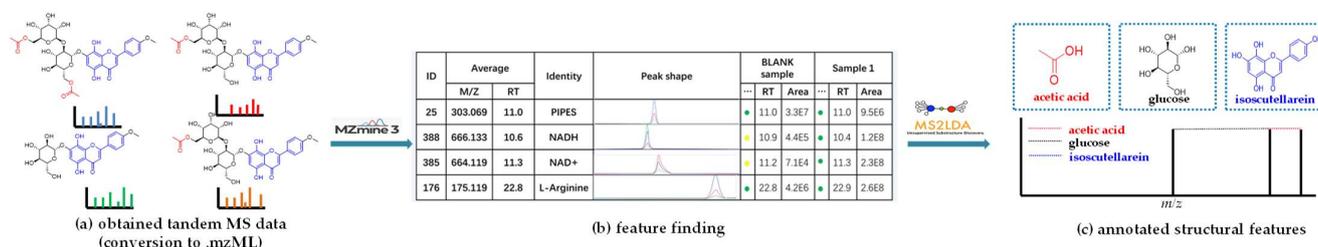


**Figure 8.** Structures of the novel *securinega* alkaloids.

## 6. Structure-Based Molecular Networking (MS2LDA)

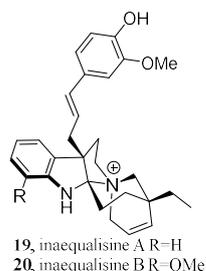
Apart from the efficient use of MS<sup>2</sup> datasets, the spectral annotation of molecular networks is another algorithmic bottleneck. It has been noted that many strategies, such as the In Silico MS/MS DataBase (ISDB) [67], network annotation propagation (NAP) [68], Sirius [69], MetWork [70], and MS2LDA [51,71] have been proposed. In 2016, inspired by the text-mining algorithm for latent Dirichlet allocation (LDA) [72], Hooft and colleagues presented an MS2LDA that can find and match co-occurring molecular fragments and neutral spectra (“Mass2Motifs”) from fragmentation spectra (Figure 9a) in an unsupervised manner [36,51] (Figure 9). In turn, the obtained Mass2Motifs provide information on the functional groups, building blocks, or even scaffolds of a compound (Figure 9c) that

can be used for the de novo annotation of unknown molecules without reference spectra. Other substructure discovery-based tools, including the metabolite substructure auto-recommender (MESSAR) [73], and CSI (compound structure identification):FingerID [74] were also developed. Among them, MESSAR is a complementary approach to MS2LDA that can provide an automated structural annotation of Mass2Motifs.



**Figure 9.** Schematic representation of the principles of MS2LDA. (a) The obtained tandem MS data. (b) Feature-finding. (c) The annotated structural features.

For example, in the search for monoterpene indole alkaloids (MIAs) from the roots of *Callichilia inaequalis* Stapf (Apocynaceae), MS2LDA was employed to fine-tune the isolation workflow [75]. The obtained MS<sup>2</sup> data of the seven alkaloid fractions were processed by FBMN. After MIADB [76] spectral library annotation, some clusters of the molecular network remained unannotated; thus, MS2LDA was utilized to further unearth more information from these clusters. As a result, molecular family A, a nine-parent mass shared cluster, exhibited an intriguing mass loss of 162.075 Da that did not match any MotifDB. A further literature survey indicated that this Mass2Motif may be a hexose unit, but this annotation was inconsistent with both the molecular formulas and the elemental composition. Hence, molecular family A was selected for targeted isolation, then liquid chromatography (LC)-diode array detector (DAD)-MS-evaporative light scattering detector (ELSD)-guided isolation led to the discovery of two novel hybrid alkylated phenylpropane MIAs (19–20) (Figure 10).

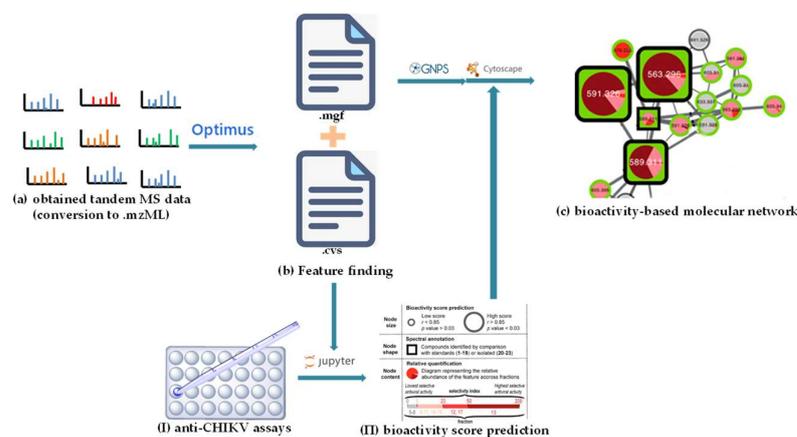


**Figure 10.** Structures of two novel hybrid alkylated phenylpropane MIAs.

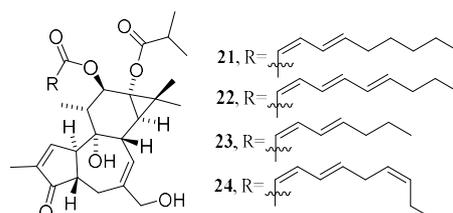
## 7. Bioactivity-Based Molecular Networking (BMN)

In addition, other NP prioritization strategies were also developed, based on the combination of other layers of information, such as biochemometrics [77], genomics [78], and taxonomy [79]. In the search for bioactive compounds against the chikungunya virus (CHIKV) from *Euphorbia dendroides*, Dorrestein's group presented BMN by combining chemometrics with MN in 2018 [52]. Chemometrics can distinguish the active and inactive compounds in mixtures but cannot provide structural information regarding them [80]. In this study, after chromatographic separation, the obtained 18 fractions of the *E. dendroides* latex extract were detected to obtain their LC-MS/MS data (Figure 11a) and anti-CHIKV activities (Figure 11I). The MS<sup>2</sup> data were pre-processed by Optimus [81] to generate the spectral features files (Figure 11b) (the .mgf file and the .cvs file, which were used to generate CLMN and bioactivity scores, respectively) (see Figure 11). The bioactivity scores for each ion were calculated using a Pearson correlation that correlated the relative abundance of an ion and the selectivity index of fractions (Figure 11II). Finally, combined with the predicted

bioactivity scores, the bioactive molecular network was generated using GNPS (Figure 11c), in which a large node size indicated high activity, and relative abundance in a certain fraction was shown by a pie chart. A detailed analysis of the molecular networks indicated that the compounds in molecular network 2 were most likely to be promising bioactive candidates. Then, four new  $4\beta$ -deoxyphorbol esters (21–24) (Figure 12) with anti-CHIKV activities were obtained by HPLC-guided isolation.



**Figure 11.** Schematic representation of the principle of BMN. (a) The obtained tandem MS data. (b) Feature-finding. (I) Anti-CHIKV assays. (II) Bioactivity score prediction. (c) The bioactivity-based molecular network.



**Figure 12.** Structures of four new  $4\beta$ -deoxyphorbol esters.

## 8. Conclusions

Needless to say, NPs have played an important role in the discovery of small-molecule drugs in the past. According to Linington's analysis [18], the future for NPs is very bright, and the chemical space is large. The launch of ADC [12] and the industrialization of Eribulin (Halaven<sup>TM</sup>) [82] showed us the infinite possibilities for new drugs from NPs. Nevertheless, aggressive innovation is needed to tackle bottlenecks such as the rapid discovery and large-scale availability of NPs. In the past decade, the advent of MS/MS-based MN has greatly promoted the development of NP dereplication methods. In this review, we introduced CLMN, FBMN, IIMN, BBMN, MS2LDA, and BMN, along with their basic principles, general workflow, and application examples, hoping to further the research and applications of these methods. Although there are numerous studies of MN, such as those on datasets, algorithms, data pre-processing, annotation, and the different types of mass spectrometers or hybrids with other methods, this review covers the key concepts and steps in the molecular network construction pipeline, which is helpful for beginners hoping to learn this methodology.

It is worth noting that MS-based analysis is a biased detection method, depending on how well the compounds fragment, and some NPs do not ionize as well as others. NMR-based approaches can make up for this shortcoming, and dedicated 2D-NMR-based prioritization strategies have been developed rapidly [83–85]. As the two techniques can complement each other, new hybrid MS/NMR approaches are emerging for NPs prioritization [39]. On the other hand, although many annotation methods have been developed, the exploitation of molecular networks still requires manual intervention and

expertise. We believe that open access to data, including both the MS and NMR spectra, and advancements in big data approaches will improve the efficiency of NPs dereplication.

**Author Contributions:** Conceptualization, G.-F.Q., J.-C.Y., and H.-B.L.; investigation, G.-F.Q., X.Z., F.Z., and Z.-Q.H.; writing—original draft, G.-F.Q., X.Z., F.Z., and Z.-Q.H.; writing—review, editing, and revision, Q.-Q.Y., Q.F., Z.L., G.-M.Z., J.-C.Y., and H.-B.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Shandong Provincial Key Research and Development Program (Major Technological Innovation Project) (2021CXGC010508), Taishan Industrial Leading Talents Program (the grant number is not available), Shandong Provincial Natural Science Foundation (ZR2021QH289), and the Independent Innovation Major Project of Linyi City, Shandong Province (2019ZDZX001).

**Acknowledgments:** Special thanks are extended to the Taishan Industrial Leading Talents Program of Shandong Province.

**Conflicts of Interest:** The Company Lunan Pharmaceutical Group Co., Ltd. had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. The company Lunan Pharmaceutical Group Co., Ltd. has been informed about the results and has no objections to having the results published in *Molecules*.

## References

1. Clardy, J.; Walsh, C. Lessons from natural molecules. *Nature* **2004**, *432*, 829–837. [[CrossRef](#)] [[PubMed](#)]
2. Ma, P.; Xu, H.; Li, J.; Lu, F.; Ma, F.; Wang, S.; Xiong, H.; Wang, W.; Buratto, D.; Zonta, F.; et al. Functionality-independent DNA encoding of complex natural products. *Angew. Chem.* **2019**, *131*, 9335–9362. [[CrossRef](#)]
3. Koch, M.A.; Schuffenhauer, A.; Scheck, M.; Wetzel, S.; Casaulta, M.; Odermatt, A.; Ertl, P.; Waldmann, H. Charting biologically relevant chemical space: A structural classification of natural products (SCONP). *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17272–17277. [[CrossRef](#)] [[PubMed](#)]
4. Newman, D.J.; Cragg, G.M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803. [[CrossRef](#)]
5. Najmi, A.; Javed, S.A.; Al Bratty, M.; Alhazmi, H.A. Modern approaches in the discovery and development of plant-based natural products and their analogues as potential therapeutic agents. *Molecules* **2022**, *27*, 349. [[CrossRef](#)]
6. Top 200 Pharmaceuticals by Retailers in 2021. Available online: <https://njardarson.lab.arizona.edu/content/top-pharmaceuticals-poster> (accessed on 12 August 2022).
7. David, B.; Wolfender, J.L.; Dias, D.A. The pharmaceutical industry and natural products: Historical status and new trends. *Phytochem. Rev.* **2014**, *14*, 299–315. [[CrossRef](#)]
8. Sheridan, C. Recasting natural product research. *Nat. Biotechnol.* **2012**, *30*, 385–387. [[CrossRef](#)]
9. McChesney, J.D.; Venkataraman, S.K.; Henri, J.T. Plant natural products: Back to the future or into extinction? *Phytochemistry* **2007**, *68*, 2015–2022. [[CrossRef](#)]
10. Walsh, G. Biopharmaceutical benchmarks 2018. *Nat. Biotechnol.* **2018**, *36*, 1136–1145. [[CrossRef](#)]
11. Cohen, P.; Cross, D.; Jänne, P.A. Kinase drug discovery 20 years after imatinib: Progress and future directions. *Nat. Rev. Drug Discov.* **2021**, *20*, 551–569. [[CrossRef](#)]
12. Drago, J.Z.; Modi, S.; Chandralapaty, S. Unlocking the potential of antibody-drug conjugates for cancer therapy. *Nat. Rev. Clin. Oncol.* **2021**, *18*, 327–344. [[CrossRef](#)] [[PubMed](#)]
13. Békés, M.; Langley, D.R.; Crews, C.M. PROTAC targeted protein degraders: The past is prologue. *Nat. Rev. Drug Discov.* **2022**, *21*, 181–200. [[CrossRef](#)] [[PubMed](#)]
14. Sabe, V.T.; Ntombela, T.; Jhamba, L.A.; Maguire, G.E.M.; Govender, T.; Naicker, T.; Kruger, H.G. Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review. *Eur. J. Med. Chem.* **2021**, *224*, 113705. [[CrossRef](#)] [[PubMed](#)]
15. Cooper, B.M.; Iegre, J.; O'Donovan, D.H.; Halvarsson, M.Ö.; Spring, D.R. Peptides as a platform for targeted therapeutics for cancer: Peptide-drug conjugates (PDCs). *Chem. Soc. Rev.* **2021**, *50*, 1480–1494. [[CrossRef](#)] [[PubMed](#)]
16. Kurita, K.L.; Linington, R.G. Connecting phenotype and chemotype: High-content discovery strategies for natural products research. *J. Nat. Prod.* **2015**, *78*, 587–596. [[CrossRef](#)] [[PubMed](#)]
17. Sorokina, M.; Steinbeck, C. Review on natural products databases: Where to find data in 2020. *J. Cheminform.* **2020**, *12*, 20. [[CrossRef](#)]
18. Pye, C.R.; Bertin, M.J.; Lokey, R.S.; Gerwick, W.H.; Linington, R.G. Retrospective analysis of natural products provides insights for future discovery trends. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 5601–5606. [[CrossRef](#)]
19. Covington, B.C.; McLean, J.A.; Bachmann, B.O. Comparative mass spectrometry-based metabolomics strategies for the investigation of microbial secondary metabolites. *Nat. Prod. Rep.* **2017**, *34*, 6–24. [[CrossRef](#)]

20. Gaudencio, S.P.; Pereira, F. Dereplication: Racing to speed up the natural products discovery process. *Nat. Prod. Rep.* **2015**, *32*, 779–810. [[CrossRef](#)]
21. El-Elimat, T.; Figueroa, M.; Ehrmann, B.M.; Cech, N.B.; Pearce, C.J.; Oberlies, N.H. High-resolution MS, MS/MS, and UV database of fungal secondary metabolites as a dereplication protocol for bioactive natural products. *J. Nat. Prod.* **2013**, *76*, 1709–1716. [[CrossRef](#)]
22. Wohlleben, W.; Mast, Y.; Stegmann, E.; Ziemert, N. Antibiotic drug discovery. *Microb. Biotechnol.* **2016**, *9*, 541–548. [[CrossRef](#)] [[PubMed](#)]
23. López-Pérez, J.L.; Therón, R.; del Olmo, E.; Díaz, D. NAPROC-13: A database for the dereplication of natural product mixtures in bioassay-guided protocols. *Bioinformatics* **2007**, *23*, 3256–3257. [[CrossRef](#)] [[PubMed](#)]
24. Crüsemann, M.; O’Neill, E.C.; Larson, C.B.; Melnik, A.V.; Floros, D.J.; da Silva, R.R.; Jensen, P.R.; Dorrestein, P.C.; Moore, B.S. Prioritizing natural product diversity in a collection of 146 bacterial strains based on growth and extraction protocols. *J. Nat. Prod.* **2017**, *80*, 588–597. [[CrossRef](#)] [[PubMed](#)]
25. Agarwal, G.; Carcache, P.J.B.; Addo, E.M.; Kinghorn, A.D. Current status and contemporary approaches to the discovery of antitumor agents from higher plants. *Biotechnol. Adv.* **2020**, *38*, 107337. [[CrossRef](#)]
26. Alfaro, J.A.; Bohländer, P.; Dai, M.; Filius, M.; Howard, C.J.; van Kooten, X.F.; Ohayon, S.; Pomorski, A.; Schmid, S.; Aksimentiev, A.; et al. The emerging landscape of single-molecule protein sequencing technologies. *Nat. Methods* **2021**, *18*, 604–617. [[CrossRef](#)]
27. Zhang, F.; Ge, W.; Ruan, G.; Cai, X.; Guo, T. Data-independent acquisition mass spectrometry-based proteomics and software tools: A glimpse in 2020. *Proteomics* **2020**, *20*, e1900276. [[CrossRef](#)]
28. Timp, W.; Timp, G. Beyond mass spectrometry, the next step in proteomics. *Sci. Adv.* **2020**, *6*, eaax8978. [[CrossRef](#)]
29. Hajirasouliha, I.; Tilgner, H.U. The tech for the next decade: Promises and challenges in genome biology. *Genome Biol.* **2019**, *20*, 86. [[CrossRef](#)]
30. Miggiels, P.; Wouters, B.; van Westen, G.J.P.; Dubbelman, A.-C.; Hankemeier, T. Novel technologies for metabolomics: More for less. *TrAC Trends Anal. Chem.* **2019**, *120*, 115323. [[CrossRef](#)]
31. Aldridge, S.; Teichmann, S.A. Single cell transcriptomics comes of age. *Nat. Commun.* **2020**, *11*, 4307. [[CrossRef](#)]
32. Asp, M.; Bergensträhle, J.; Lundberg, J. Spatially resolved transcriptomes—next generation tools for tissue exploration. *Bioessays* **2020**, *42*, 1900221. [[CrossRef](#)] [[PubMed](#)]
33. Caesar, L.K.; Montaser, R.; Keller, N.P.; Kelleher, N.L. Metabolomics and genomics in natural products research: Complementary tools for targeting new chemical entities. *Nat. Prod. Rep.* **2021**, *38*, 2041–2065. [[CrossRef](#)] [[PubMed](#)]
34. Sukmarini, L. Recent advances in discovery of lead structures from microbial natural products: Genomics- and metabolomics-guided acceleration. *Molecules* **2021**, *26*, 2542. [[CrossRef](#)]
35. Wolfender, J.-L.; Litaudon, M.; Touboul, D.; Queiroz, E.F. Innovative omics-based approaches for prioritisation and targeted isolation of natural products—New strategies for drug discovery. *Nat. Prod. Rep.* **2019**, *36*, 855–868. [[CrossRef](#)]
36. Beniddir, M.A.; Kang, K.B.; Genta-Jouve, G.; Huber, F.; Rogers, S.; van der Hooft, J.J.J. Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches. *Nat. Prod. Rep.* **2021**, *38*, 1967–1993. [[CrossRef](#)] [[PubMed](#)]
37. Jarmusch, S.A.; van der Hooft, J.J.J.; Dorrestein, P.C.; Jarmusch, A.K. Advancements in capturing and mining mass spectrometry data are transforming natural products research. *Nat. Prod. Rep.* **2021**, *38*, 2066–2082. [[CrossRef](#)] [[PubMed](#)]
38. Ramos, A.E.F.; Evanno, L.; Poupon, E.; Champy, P.; Beniddir, M.A. Natural products targeting strategies involving molecular networking: Different manners, one goal. *Nat. Prod. Rep.* **2019**, *36*, 960–980. [[CrossRef](#)] [[PubMed](#)]
39. Bingol, K.; Brüsweiler, R. Knowns and unknowns in metabolomics identified by multidimensional NMR and hybrid MS/NMR methods. *Curr. Opin. Biotechnol.* **2017**, *43*, 17–24. [[CrossRef](#)] [[PubMed](#)]
40. Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B.S.; Yang, J.Y.; Kersten, R.D.; van der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J.M.; et al. Mass spectral molecular networking of living microbial colonies. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E1743–E1752. [[CrossRef](#)]
41. Wang, M.; Carver, J.J.; Phelan, V.V.; Sanchez, L.M.; Garg, N.; Peng, Y.; Nguyen, D.D.; Watrous, J.; Kaponov, C.A.; Luzzatto-Knaan, T.; et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828–837. [[CrossRef](#)]
42. Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; et al. Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **2020**, *17*, 905–908. [[CrossRef](#)] [[PubMed](#)]
43. Allard, S.; Allard, P.-M.; Morel, I.; Gicquel, T. Application of a molecular networking approach for clinical and forensic toxicology exemplified in three cases involving 3-MeO-PCP, doxylamine, and chlormequat. *Drug Test. Anal.* **2019**, *11*, 669–677. [[CrossRef](#)] [[PubMed](#)]
44. Ge, Y.-W.; Zhu, S.; Yoshimatsu, K.; Komatsu, K. MS/MS similarity networking accelerated target profiling of triterpene saponins in *Eleutherococcus senticosus* leaves. *Food Chem.* **2017**, *227*, 444–452. [[CrossRef](#)] [[PubMed](#)]
45. Teta, R.; Della Sala, G.; Glukhov, E.; Gerwick, L.; Gerwick, W.H.; Mangoni, A.; Costantino, V. Combined LC-MS/MS and molecular networking approach reveals new cyanotoxins from the 2014 cyanobacterial bloom in Green Lake, Seattle. *Environ. Sci. Technol.* **2015**, *49*, 14301–14310. [[CrossRef](#)]

46. Semple, S.J.; Staerk, D.; Buirchell, B.J.; Fowler, R.M.; Gericke, O.; Kjaerulff, L.; Zhao, Y.; Pedersen, H.A.; Petersen, M.J.; Rasmussen, L.F.; et al. Biodiscoveries within the Australian plant genus *Eremophila* based on international and interdisciplinary collaboration: Results and perspectives on outstanding ethical dilemmas. *Plant J.* **2022**, *111*, 936–953. [[CrossRef](#)]
47. Molino, R.; Rellin, K.F.B.; Nellas, R.B.; Junio, H.A. Sustainable Hues: Exploring the molecular palette of biowaste dyes through LC-MS metabolomics. *Molecules* **2021**, *26*, 6645. [[CrossRef](#)]
48. Maniei, F.; Moghaddam, J.A.; Crüsemann, M.; Beemelmans, C.; König, G.M.; Wägele, H. From Persian Gulf to Indonesia: Interrelated phylogeographic distance and chemistry within the genus *Peronia* (Onchidiidae, Gastropoda, Mollusca). *Sci. Rep.* **2020**, *10*, 13048. [[CrossRef](#)]
49. Schmid, R.; Petras, D.; Nothias, L.-F.; Wang, M.; Aron, A.T.; Jagels, A.; Tsugawa, H.; Rainer, J.; Garcia-Aloy, M.; Dührkop, K.; et al. Ion identity molecular networking for mass spectrometry-based metabolomics in the GNPS environment. *Nat. Commun.* **2021**, *12*, 3832. [[CrossRef](#)]
50. He, Q.-F.; Wu, Z.-L.; Li, L.; Sun, W.-Y.; Wang, G.-Y.; Jiang, R.-W.; Hu, L.-J.; Shi, L.; He, R.-R.; Wang, Y.; et al. Discovery of neurotogenic *securinega* alkaloids from *Flueggea suffruticosa* by a building blocks-based molecular network strategy. *Angew. Chem. Int. Ed.* **2021**, *60*, 19609–19613. [[CrossRef](#)]
51. van der Hooft, J.J.J.; Wandy, J.; Barrett, M.P.; Burgess, K.E.V.; Rogers, S. Topic modeling for untargeted substructure exploration in metabolomics. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13738–13743. [[CrossRef](#)]
52. Nothias, L.-F.; Nothias-Esposito, M.; da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; et al. Bioactivity-based molecular networking for the discovery of drug leads in natural product bioassay-guided fractionation. *J. Nat. Prod.* **2018**, *81*, 758–767. [[CrossRef](#)] [[PubMed](#)]
53. 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.; et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat. Protoc.* **2020**, *15*, 1954–1991. [[CrossRef](#)] [[PubMed](#)]
54. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)] [[PubMed](#)]
55. Olivon, F.; Elie, N.; Grelier, G.; Roussi, F.; Litaudon, M.; Touboul, D. MetGem software for the generation of molecular networks based on the t-SNE algorithm. *Anal. Chem.* **2018**, *90*, 13900–13908. [[CrossRef](#)] [[PubMed](#)]
56. Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
57. Huber, F.; Ridder, L.; Verhoeven, S.; Spaaks, J.H.; Diblen, F.; Rogers, S.; van der Hooft, J.J.J. Spec2Vec: Improved mass spectral similarity scoring through learning of structural relationships. *PLoS Comput. Biol.* **2021**, *17*, e1008724. [[CrossRef](#)]
58. Huber, F.; van der Burg, S.; van der Hooft, J.J.J.; Ridder, L. MS2DeepScore: A novel deep learning similarity measure to compare tandem mass spectra. *J. Cheminform.* **2021**, *13*, 84. [[CrossRef](#)]
59. Lee, S.R.; Schalk, F.; Schwitala, J.W.; Guo, H.; Yu, J.S.; Song, M.; Jung, W.H.; de Beer, Z.W.; Beemelmans, C.; Kim, K.H. GNPS-guided discovery of madurastatin siderophores from the termite-associated *Actinomadura* sp. RB99. *Chem. Eur. J.* **2022**, *28*, e202200612. [[CrossRef](#)]
60. Wu, C.; van der Heul, H.U.; Melnik, A.V.; Lüebben, J.; Dorrestein, P.C.; Minnaard, A.J.; Choi, Y.H.; van Wezel, G.P. Lugdunomycin, an angucycline-derived molecule with unprecedented chemical architecture. *Angew. Chem. Int. Ed.* **2019**, *58*, 2809–2814. [[CrossRef](#)]
61. Bonneau, N.; Chen, G.; Lachkar, D.; Boufridi, A.; Gallard, J.-F.; Retailleau, P.; Petek, S.; Debitus, C.; Evanno, L.; Beniddir, M.A.; et al. An unprecedented blue chromophore found in Nature using a "chemistry first" and molecular networking approach: Discovery of dactylocyanines A-H. *Chem. Eur. J.* **2017**, *23*, 14454–14461. [[CrossRef](#)]
62. 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. [[CrossRef](#)]
63. Röst, H.L.; Sachsenberg, T.; Aiche, S.; Bielow, C.; Weissner, H.; Aicheler, F.; Andreotti, S.; Ehrlich, H.C.; Gutenbrunner, P.; Kenar, E.; et al. OpenMS: A flexible open-source software platform for mass spectrometry data analysis. *Nat. Methods* **2016**, *13*, 741–748. [[CrossRef](#)] [[PubMed](#)]
64. Freire, V.F.; Gubiani, J.R.; Spencer, T.M.; Hajdu, E.; Ferreira, A.G.; Ferreira, D.A.S.; de Castro Levatti, E.V.; Burdette, J.E.; Camargo, C.H.; Tempone, A.G.; et al. Feature-based molecular networking discovery of bromopyrrole alkaloids from the marine sponge *Agelas dispar*. *J. Nat. Prod.* **2022**, *85*, 1340–1350. [[CrossRef](#)] [[PubMed](#)]
65. Hell, T.; Rutz, A.; Dürr, L.; Dobrzyński, M.; Reinhardt, J.K.; Lehner, T.; Keller, M.; John, A.; Gupta, M.; Pert, O.; et al. Combining activity profiling with advanced annotation to accelerate the discovery of natural products targeting oncogenic signaling in melanoma. *J. Nat. Prod.* **2022**, *85*, 1540–1554. [[CrossRef](#)] [[PubMed](#)]
66. Yang, W.Z.; Shi, X.J.; Yao, C.L.; Huang, Y.; Hou, J.J.; Han, S.M.; Feng, Z.J.; Wei, W.L.; Wu, W.Y.; Guo, D.A. A novel neutral loss/product ion scan-incorporated integral approach for the untargeted characterization and comparison of the carboxyl-free ginsenosides from *Panax ginseng*, *Panax quinquefolius*, and *Panax notoginseng*. *J. Pharm. Biomed. Anal.* **2020**, *177*, 112813. [[CrossRef](#)]
67. Allard, P.-M.; Péresse, T.; Bisson, J.; Gindro, K.; Marcourt, L.; Pham, V.C.; Roussi, F.; Litaudon, M.; Wolfender, J.-L. Integration of molecular networking and *in-silico* MS/MS fragmentation for natural products dereplication. *Anal. Chem.* **2016**, *88*, 3317–3323. [[CrossRef](#)]

68. 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. Propagating annotations of molecular networks using *in silico* fragmentation. *PLoS Comput. Biol.* **2018**, *14*, e1006089. [[CrossRef](#)]
69. Böcker, S.; Dührkop, K. Fragmentation trees reloaded. *J. Cheminf.* **2016**, *8*, 5. [[CrossRef](#)]
70. Beauxis, Y.; Genta-Jouve, G. MetWork: A web server for natural products anticipation. *Bioinformatics* **2019**, *35*, 1795–1796. [[CrossRef](#)]
71. van der Hooft, J.J.J.; Wandy, J.; Young, F.; Padmanabhan, S.; Gerasimidis, K.; Burgess, K.E.V.; Barrett, M.P.; Rogers, S. Unsupervised discovery and comparison of structural families across multiple samples in untargeted metabolomics. *Anal. Chem.* **2017**, *89*, 7569–7577. [[CrossRef](#)]
72. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
73. Papin, J.A.; Liu, Y.; Mrzic, A.; Meysman, P.; De Vijlder, T.; Romijn, E.P.; Valkenburg, D.; Bittremieux, W.; Laukens, K. MESSAR: Automated recommendation of metabolite substructures from tandem mass spectra. *PLoS ONE* **2020**, *15*, e0226770.
74. Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. Searching molecular structure databases with tandem mass spectra using CSI: FingerID. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 12580–12585. [[CrossRef](#)] [[PubMed](#)]
75. Cauchie, G.; N’Nang, E.O.; van der Hooft, J.J.J.; Le Pogam, P.; Bernadat, G.; Gallard, J.-F.; Kumulungui, B.; Champy, P.; Poupon, E.; Beniddir, M.A. Phenylpropane as an alternative dearomatizing unit of indoles: Discovery of inaequalisines A and B using substructure-informed molecular networking. *Org. Lett.* **2020**, *22*, 6077–6081. [[CrossRef](#)] [[PubMed](#)]
76. Fox Ramos, A.E.; Le Pogam, P.; Fox Alcover, C.; Otogo N’Nang, E.; Cauchie, G.; Hazni, H.; Awang, K.; Bréard, D.; Echavarrén, A.M.; Frédéricich, M.; et al. Collected mass spectrometry data on monoterpene indole alkaloids from natural product chemistry research. *Sci. Data* **2019**, *6*, 15. [[CrossRef](#)]
77. Caesar, L.K.; Kellogg, J.J.; Kvalheim, O.M.; Cech, R.A.; Cech, N.B. Integration of biochemometrics and molecular networking to identify antimicrobials in *Angelica keiskei*. *Planta Med.* **2018**, *84*, 721–728. [[CrossRef](#)] [[PubMed](#)]
78. Ouchene, R.; Stien, D.; Segret, J.; Kecha, M.; Rodrigues, A.M.S.; Veckerlé, C.; Suzuki, M.T. Integrated metabolomic, molecular networking, and genome mining analyses uncover novel angucyclines from *Streptomyces* sp. RO-S4 strain isolated from Bejaia Bay, Algeria. *Front. Microbiol.* **2022**, *13*, 906161. [[CrossRef](#)]
79. Olivon, F.; Remy, S.; Grelier, G.; Apel, C.; Eydoux, C.; Guillemott, J.-C.; Neyts, J.; Delang, L.; Touboul, D.; Roussi, F.; et al. Antiviral compounds from *Codiaeum peltatum* targeted by a multi-informative molecular networks approach. *J. Nat. Prod.* **2019**, *82*, 330–340. [[CrossRef](#)]
80. Caesar, L.K.; Cech, N.B. Synergy and antagonism in natural product extracts: When 1+1 does not equal 2. *Nat. Prod. Rep.* **2019**, *36*, 869–888. [[CrossRef](#)]
81. Protsyuk, I.; Melnik, A.M.; Nothias, L.-F.; Rappez, L.; Phapale, P.; Aksenov, A.A.; Bouslimani, A.; Ryazanov, S.; Dorrestein, P.C.; Alexandrov, T. 3D molecular cartography using LC–MS facilitated by Optimus and’ili software. *Nat. Protoc.* **2018**, *13*, 134–154. [[CrossRef](#)]
82. Melvin, J.Y.; Zheng, W.; Seletsky, B.M. From micrograms to grams: Scale-up synthesis of eribulin mesylate. *Nat. Prod. Rep.* **2013**, *30*, 1158–1164.
83. Deyrup, S.T.; Eckman, L.E.; McCarthy, P.H.; Smedley, S.R.; Meinwald, J.; Schroeder, F.C. 2D NMR-spectroscopic screening reveals polyketides in ladybugs. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9753–9758. [[CrossRef](#)] [[PubMed](#)]
84. Bingol, K.; Zhang, F.; Bruschweiler-Li, L.; Bruschweiler, R. Carbon backbone topology of the metabolome of a cell. *J. Am. Chem. Soc.* **2012**, *134*, 9006–9011. [[CrossRef](#)] [[PubMed](#)]
85. Reher, R.; Kim, H.W.; Zhang, C.; Mao, H.H.; Wang, M.; Nothias, L.-F.; Caraballo-Rodríguez, A.M.; Glukhov, E.; Teke, B.; Leao, T.; et al. A convolutional neural network-based approach for the rapid annotation of molecularly diverse natural products. *J. Am. Chem. Soc.* **2020**, *142*, 4114–4120. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.