

Review

High-Throughput Mining of Novel Compounds from Known Microbes: A Boost to Natural Product Screening

Surya Nandan Meena ¹, Anna Wajs-Bonikowska ^{2,*}, Savita Girawale ¹, Md Imran ³, Preethi Poduval ⁴ and Kisan M. Kodam ¹

¹ Department of Chemistry, Savitribai Phule Pune University, Pune 411007, India; suryanandanmeena6@gmail.com (S.N.M.); kisan.kodam@unipune.ac.in (K.M.K.)

² Institute of Natural Products and Cosmetics, Faculty of Biotechnology and Food Sciences, Łódź University of Technology, Stefanowskiego Street 2/22, 90-537 Łódź, Poland

³ Department of Botany, University of Delhi, Delhi 110007, India

⁴ Department of Biotechnology, Dhempe College of Arts and Science, Miramar, Goa 403001, India; biotech.preethi@unigoa.ac.in

* Correspondence: anna.wajs-bonikowska@p.lodz.pl

Abstract: Advanced techniques can accelerate the pace of natural product discovery from microbes, which has been lagging behind the drug discovery era. Therefore, the present review article discusses the various interdisciplinary and cutting-edge techniques to present a concrete strategy that enables the high-throughput screening of novel natural compounds (NCs) from known microbes. Recent bioinformatics methods revealed that the microbial genome contains a huge untapped reservoir of silent biosynthetic gene clusters (BGC). This article describes several methods to identify the microbial strains with hidden mines of silent BGCs. Moreover, antiSMASH 5.0 is a free, accurate, and highly reliable bioinformatics tool discussed in detail to identify silent BGCs in the microbial genome. Further, the latest microbial culture technique, HiTES (high-throughput elicitor screening), has been detailed for the expression of silent BGCs using 500–1000 different growth conditions at a time. Following the expression of silent BGCs, the latest mass spectrometry methods are highlighted to identify the NCs. The recently emerged LAESI-IMS (laser ablation electrospray ionization-imaging mass spectrometry) technique, which enables the rapid identification of novel NCs directly from microtiter plates, is presented in detail. Finally, various trending ‘dereplication’ strategies are emphasized to increase the effectiveness of NC screening.

Keywords: natural product; microbes; biosynthetic gene clusters (BGCs); genome mining; antiSMASH; HiTES; LAESI-IMS; dereplication



Citation: Meena, S.N.; Wajs-Bonikowska, A.; Girawale, S.; Imran, M.; Poduval, P.; Kodam, K.M. High-Throughput Mining of Novel Compounds from Known Microbes: A Boost to Natural Product Screening. *Molecules* **2024**, *29*, 3237. <https://doi.org/10.3390/molecules29133237>

Academic Editor: José Rubén Tormo

Received: 3 June 2024

Revised: 2 July 2024

Accepted: 3 July 2024

Published: 8 July 2024



Copyright: © 2024 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

The period between 1940 and 1980 was recognized as the golden age of drug research, as several drugs were discovered in this era, like antibacterial, antifungal, antiviral, anti-cancer, antidiabetic, hypercholesterolemia, immunosuppressants, etc. [1–3]. In addition to their clinical utility, natural products are invaluable sources for nutrition [4], pigments as natural colors [5], and cosmetics [6]. The advancement of basic biological research to understand health and various diseases [7–9] led to the discovery of druggable natural compounds (NCs) from microbes. During the drug discovery era, around 200–300 NCs/year were found, and in the late 1970s, the rate was enhanced to ≈500 NCs/year. After the 1980s, the rate of drug discovery slowed down as it became harder to identify novel and useful NCs from the microbes (Figure 1).

The high rate of re-discovery of NCs discouraged the natural product community [10,11]. Increased use of combinatorial chemistry approaches to develop an analogue of NCs and high-throughput screening of the synthetic small compound libraries for their bioactivity was also a significant reason [12,13]. The other major non-scientific reason for the sluggish pace of drug

discovery was the lack of interest among large drug-producing companies [14]. This is due to higher regulatory hurdles related to long-term pre-clinical and clinical studies prior to sale. Another reason is the short time of sale of these drugs until the expiry of their patent and even the brief duration of usage by customers until the expiry date of the drugs. The approximate expenses of producing a new drug are around \$1 billion and require 14–15 years [15–17]. Even though the growth of drug discovery from natural sources was greatly affected by technical and non-scientific factors, such factors are not in the context of this review article.

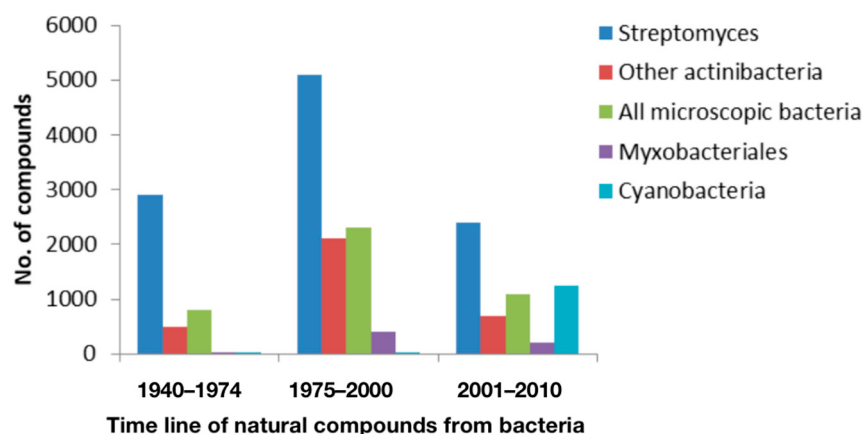


Figure 1. Approximate number of natural compounds discovered from bacteria during the period from 1940 to 2010.

1.1. Why Novel Drugs Are Required?

Despite several limitations, natural product research should be continued due to unmet needs. Since the 1980s, several new diseases have emerged (e.g., AIDS, Hanta, SARS-CoV, H1N1/09 virus, MERS-CoV, Ebola, Zika virus, yellow fever, etc.), and the currently running pandemic COVID-19 poses a global threat to the human race [18,19]. Increased pathogen resistance to available antibiotics and the toxic effects of certain medicines have triggered the need for novel medicines [20,21]. Instead of discovering novel drugs from natural sources, large drug manufacturers have focused on using combinatorial chemistry during 1990–2005 to generate a library of small compounds [22]. Further efforts were made to screen these compounds for their bioactivity in a high-throughput manner [23]. Unfortunately, success was not achieved due to several reasons, such as difficulty in synthesizing chemically diverse, high-quality libraries of molecules with desirable features of natural products, such as diverse functionality, significant skeletal diversity, and good pharmacokinetic properties, from readily available and inexpensive building blocks [24,25]. By adopting the combinatorial chemistry screening approach, researchers could not succeed in creating diverse and pharmacologically active compounds (Table 1) [26]. Therefore, once again, the researchers have realized the importance of small bioactive compounds from natural sources in drug discovery. Although several drugs were discovered during the slow growth phase of drug discovery, the NCs discovered during this phase were not sufficient to meet the demands of global healthcare challenges [27–29]. Therefore, the untapped sources of the novel NCs need to be explored through the use of cutting-edge techniques.

Table 1. Comparative analysis of natural compound discovery from various sources [26].

	Number	%	Drugs	Success%
Synthetic compounds	8–10 M	93–94%	2000–2500	0.005%
All natural compounds (plants + animals + microbes)	≈500,000	4.7–5.8%	1200–1300	0.6%
Microbial compounds	≈70,000	0.66–0.82%	450–500	1.6%

1.2. Conventional and Current Strategies for Drug Discovery

The conventional way of discovering drugs from microbes is time-consuming, expensive, and laborious. The separation and purification of microbial secondary metabolites is a critical step and needs expertise. Further, accurate identification of the purified compounds is also manual and requires several analytical techniques, like nuclear magnetic resonance spectroscopy (NMR), infrared spectroscopy (IR), liquid chromatography-mass spectrometry (LC-MS), etc., which eventually slows down the pace of drug discovery. Presently, system biology techniques such as genomics and bioinformatics tools have the potential to enable researchers to understand the arrangement and regulation of secondary metabolite-encoding genes and their network in the microbial genome [30]. Further, the fusion of natural product screening with high-throughput analysis, genomics, chemistry, metabolomics, and microbial biodiversity exploration may offer great opportunities for drug discovery [31,32]. High-throughput screening techniques allow the quick identification of hundreds of samples, while genomics and metabolomics approaches help in understanding the detailed information about the genetic and metabolic potential of the microbes.

Further advancements in the upcoming technologies, such as next-generation sequencing (NGS), artificial intelligence (AI), and machine learning (ML), could further revolutionize the field of microbial drug discovery. These technologies could potentially automate and streamline many of the laborious and time-consuming steps in the drug discovery process, which will eventually make the entire process more efficient and cost-effective. Traditional methods may overlook novel drug candidates, but modern techniques could enhance their discovery. Thus, the future of microbial drug discovery looks promising, with the potential to yield a wealth of new therapeutics for a variety of diseases.

1.3. Novel Sources: Novel Drug Approach

The large-scale identification of novel bacterial species in the microbial world, followed by biochemical screening to identify NCs, has always been an opportunity for researchers worldwide. But, exploring novel microbial species for specific NCs or detecting novel NCs from microorganisms has always been a challenging mission. As an alternative, combinatorial libraries comprised of synthetic compounds depict a diverse range of scaffolds with minimal efforts for drug screening. But, in the laboratory, high-throughput screening of the combinatorial libraries gives a lower hit rate than natural compound libraries [33,34]. It has been realized that the role of combinatorial chemistry in drug discovery has remained just to assist the strategies involved in natural product discovery, not to replace them [35]. Considering past research experiences, it has been clearly evidenced that novel sources were always explored for new drugs, but this approach could not sustain the pace of current drug discovery strategies and the global need for drugs. Therefore, the past consequences in drug discovery trigger the researchers to rethink exploring the available microbial sources with hidden cryptic genes for novel drugs.

Earlier, drug discovery was based on the screening of NCs from source microorganisms, e.g., *Streptomyces* sp., whereas after the discovery of biosynthetic genes in the late 1970s, researchers came to know the genetic basis of the synthesis of NCs [36]. A microbial strain generally produces many NCs; e.g., strains of *Micromonospora* sp., *Streptomyces* sp., *Myxococcus xanthus*, *Aspergillus ochraceus*, and *Sphaeropsideles* sp. were found to produce 50, 12, 38, 16, and 19 compounds, respectively [37–39]. The bacterial genes responsible for producing NCs are present in the form of clusters known as biosynthetic gene clusters (BGCs). Profiling of genomic sequences facilitated the expansion of several gene clusters [40,41] and potential drug targets [42,43], which gave hope that genomics could resolve the pharmaceutical productivity crisis. Several bacterial species have more BGCs, like *Streptomyces* sp., while *Ktedonobacteria* sp. consists of 30 and 104 BGCs, respectively [44]. In another study, induced expression of BGCs in *Aspergillus nidulans* using the transcriptional regulator *LaeA* resulted in the expression of several novel NCs (e.g., 14 non-ribosomal peptides, two indole alkaloids, 27 polyketides, and one terpene), which were absent in the control species of *A. nidulans*, underscoring the importance of BGCs in the production of diverse NCs [45].

Using the recent bioinformatics tools, more detailed analysis of the BGCs revealed that the bacterial genome contains a trove of silent or cryptic BGCs [46–49], which are not expressed during normal growth conditions in the laboratory [46]. Indeed, the expression of a large part of microbial BGCs varies with specific environmental or cultural conditions [50]. Recent studies in genome mining have disclosed that microbial genomes comprise a large number of obscure silent BGCs [44,45]. This shows that the bacterial genome contains a trove of silent BGCs, and that could be a gold mine for novel drug discovery. Therefore, in order to identify novel NCs from the silent BGCs, this review paper discusses a strategic workflow comprised of the latest strategies and techniques that enable the high-throughput fashion screening of novel NCs from bacterial species.

2. Microbial Genome Mining for Cryptic BGCs

Genome mining is an appropriate method to evaluate the secondary metabolite potential of microorganisms. Genome mining is carried out with the help of bioinformatics tools for the identification and characterization of gene clusters associated with the biosynthesis of NCs. The availability of a large number of bacterial genome sequences publicly is a great alternative to finding the large number of BGCs encoded for the diverse and novel NCs [51]. The strains of unculturable bacterial species are the most untainted source for the genome mining and identification of novel NCs [52].

Within the BGCs, a significant portion of the microbial genome comprised genes of different enzymes such as polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS), or mixed PKS-NRPS. Gene clusters of PKS/NRPS are particularly found within bacterial species of Proteobacteria, Actinobacteria, Firmicutes, and Cyanobacteria. However, actinomycetes stand out as prime examples of a ‘productive genome’, given their potential to produce a variety of bioactive compounds [53,54]. Polyketides (PKs) and non-ribosomal peptides (NRPs) are a diverse group of bioactive secondary metabolites that are synthesized by PKS and NRPS, respectively. To date, more than 23,000 NCs derived from PKs and NRPs have been identified [55] with a huge spectrum of biological activities such as anticancer (doxorubicin and mithramycin) [56], antibiotics (tetracyclines, erythromycin, and rapamycin, daptomycin), [57,58] immunosuppressants (FK506) [59], and the NRP/PK hybrid compound epothilone is an antitumor agent [60].

Nowadays, several bioinformatics tools are available that are helpful in genome mining for silent BGCs from potent microbes [61]. To perform the bacterial genome mining, we first need the desired genome sequence of the interested microorganism. Bacterial genome sequences can be downloaded from authentic online centers, e.g., the national center for biotechnology information (NCBI) database (<http://www.ncbi.nlm.nih.gov>) and the Joint Genome Institute (JGI) database (<http://jgi.doe.gov>), in the form of GenBank, EMBL, or FASTA files. Once the genome sequence is obtained, the gene clusters associated with the biosynthesis of a secondary metabolite can be identified and characterized by using different bioinformatics tools [62,63].

2.1. Bioinformatics Tools for Genome Mining

To date, several bioinformatics tools have been developed for searching metabolic gene clusters, such as antiSMASH (Antibiotics and Secondary Metabolite Analysis Shell) [64], SMURF (Secondary Metabolite Unique Regions Finder) (<http://www.jcvi.org/smurf>) [65], NP.searcher (Natural Product Searcher) (<https://dna.sherman.lsi.umich.edu/>) [66], CLUSEAN (Cluster Sequence Analyzer) (<https://bitbucket.org/tilmweber/clusean/src/master/>) [67], ClustScan (Cluster Scanner) [68], MIDDAS-M (Motif Independent De Novo Detection Algorithm for Secondary Metabolite Gene Clusters) [69], CASSIS (Cluster Assignment by Island of Sites) [70], SMIPS (Secondary Metabolites by Inter ProScan) (<https://www.ebi.ac.uk/interpro/search/sequence/>) [70], software, and C-Hunter (http://fcg.tamu.edu/C_Hunter/) [71]. The numerous novel NCs expressed by BGCs are currently in use as antibiotics (tetracyclines, erythromycin, rapamycin, and daptomycin), immunotherapy (FK506), anticancer (doxorubicin and mithramycin) agents, etc. [72–74].

Since 2011, antiSMASH has been helping scientists with their genome mining projects. Currently, antiSMASH is the most trusted and widely used bioinformatics tool, with the additional key features incorporated by CLUSEAN, the NRPS predictor, and Cluster Finder, making it useful for identification and characterization of gene clusters responsible for the biosynthesis of secondary metabolites [75,76]. The current version, antiSMASH 5.0, contains 6200 full annotated and 18,576 drafts of the bacterial genome [77]. The antiSMASH is capable of identifying the BCGs, PKS, NRPS terpenes, siderophores, nucleosides, bacteriocins, beta-lactams, butyrolactones, melanin, antibiotics, and metabolites belonging to other classes [64].

In the current version of antiSMASH 5.0, new detection rules for gene clusters have been added. They improved the prediction of PKS, annotation of resistance genes, gene ontology annotation, and 'new region' concept. The antiSMASH 5.0 version encompasses new features for the detection of gene clusters encoded for different secondary products such as acyl-amino acids, β -lactones, C-nucleosides, polybrominated diphenyl ethers, and lipolanthins [78]. Villebro et al. examined the microbial genome using the antiSMASH 5.0 advance module for the identification of biosynthetic gene clusters, specifically for the PKS gene [79]. The study involved the detection and classification of particular genes or proteins in PKS that are responsible for polyketide synthesis. The module is capable of predicting the presence of aromatic polyketide, including possible starter units. A number of elongated malonyl moieties during PKS synthesis will also give clues about the class and molecular weight of the product. Moreover, the module will predict the cyclization patterns present in the product with high accuracy [79]. By using antiSMASH, BGCs encoded for the synthesis of new antibiotics such as pseudopyronine A and B were investigated in the genome of *Pseudomonas putida* BW11M1 [80]. Thus, the cluster-finder algorithm of antiSMASH 5.0 helps in BGC prediction and is capable of finding the signature biosynthetic genes encoded for specific enzymes involved in secondary metabolite biosynthesis. In addition, antiSMASH provides knowledge about PKS/NRPS domain analysis and annotation, substrate specificity, prediction of the core chemical structure of PKS/NRPS, secondary metabolism, protein family analysis, and gene cluster comparative analysis [81].

Workflow of antiSMASH

antiSMASH can be operated by using a web server (<https://antismash-db.secondarymetabolites.org/>) or it can be run as an independent version on a quality-grade desktop computer.

For the identification of microbial silent BGCs or gene clusters encoding the biosynthesis of secondary metabolites, a standardized workflow of antiSMASH is shown in Figure 2. For the genome sequence analysis, antiSMASH accepts multiple data formats, such as GenBank, FASTA, or EMBL, as an input file. This tool also retrieves the data from NCBI when the accession number is known. The precision of the antiSMASH data analysis is strongly dependent on the genome sequence. This tool cannot predict the gene clusters in the sequences that involve a large number of small contigs. After the submission of the job at the homepage (<https://antismash.secondarymetabolites.org/>), gene cluster prediction starts, and under the normal server condition, the entire analysis is completed in 0.5–2 h. For the typical bacterial genome, gene cluster prediction analysis may take several days. After completion of sequence analysis, results are available to download from the antiSMASH server [82].

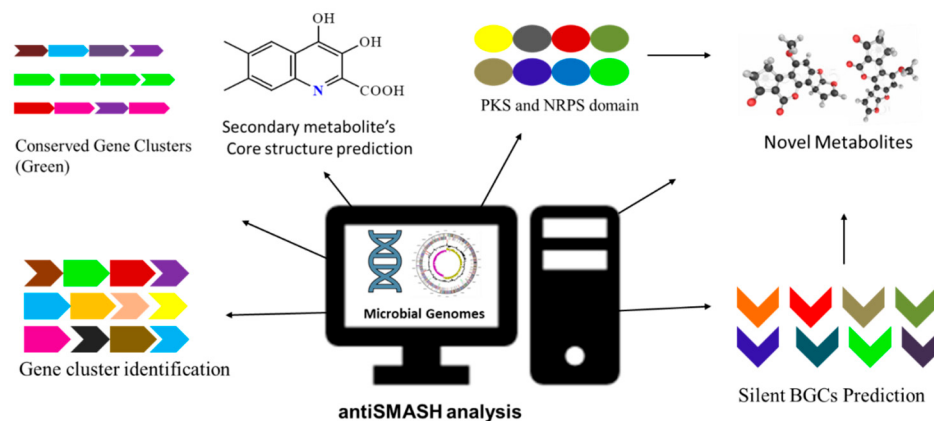


Figure 2. Flowchart of bacterial genome mining using antiSMASH for detection of biosynthesis gene clusters.

3. Strategies to Activate the Expression of Silent BGCs

Silent or cryptic BGCs can be identified bioinformatically, but it is difficult to induce their expression under normal growth conditions in the laboratory. The production of secondary metabolites in microorganisms is accomplished by the expression of BGCs in a controlled manner [83–86]. The silent or cryptic BGCs in bacteria are not expressed in normal growth conditions [46]. Indeed, the expression of a large part of microbial BGCs varies with specific environmental or cultural conditions [50]. The scientific community has recognized the importance of silent BGCs, and hence various strategies have been established to induce the expression of silent gene clusters.

The major strategies that include the expression of BGCs are ribosome engineering [87], co-culture screening [88], heterologous host [89], reporter-based selection of mutants [90], overexpression of regulatory proteins [91], and insertion of constitutive/inducible promoters [92–94]. These methods were effectively applied in order to induce the expression of silent BGCs in many species. The involvement of complex genetic manipulations during ribosome engineering and the insertion of constitutive/inducible promoter methods [95] are the challenges encountered. The typical culture conditions required for a certain type of bacterial group during co-culture screening affect the expression of BGCs, making this technique unsuitable for natural product research [96]. Complex molecular studies are involved in reported guided mutant selection and heterologous expression. Consequently, these techniques are unfit for the throughput screening of NCs (they are not in the context of this article) [97]. Here, we will focus on a new strategy called HiTES (high-throughput elicitor screening), which enables the expression of the cryptic metabolites in a high-throughput fashion [98–101].

3.1. High-Throughput Expression of Silent BGCs

Elicitors are small compounds available in the library format used for activating the expression of silent BGCs. The HiTES technique is an elicitor screening method that identifies specific small chemical compounds from the library necessary for eliciting the cryptic BGCs. The HiTES method mainly includes the following two steps: the activation of the BGCs by elicitor screening, followed by the detection of newly expressed cryptic metabolites using a genetic readout reporter assay. In this approach, a reporter gene such as LacZ [95] or eGFP [102] is inserted within the BGCs, and the subsequent strain is cultured in a 96-well plate. The microbial culture plate is incubated with a small compound library comprised of different compounds (elicitors) in each well.

After incubation, if the elicitor promotes the growth of the strain, the observed color or fluorescence in the culture plate will highlight the importance of high-throughput screening of small compound elicitors in enhancing the activity of the silent BGCs (Figure 3). HiTES enables the expression of the cryptic gene clusters in different cultural conditions at a time (500–1000). Millions of small compounds in the library are available to screen as an

elicitor for inducing the expression of microbial cryptic BGCs. The commercially available small compound libraries are ChemDiv (<https://www.chemdiv.com/complete-list/>), MolPort (<https://www.molport.com/shop/libraries-collections>), and Screen-Well® (<https://www.enzolifesciences.com/BML-2865/screen-well-natural-product-library/>). The HiTES technique enables the researchers to induce the expression of silent BGCs in a rapid manner and also provides insight into the regulation of these gene clusters.

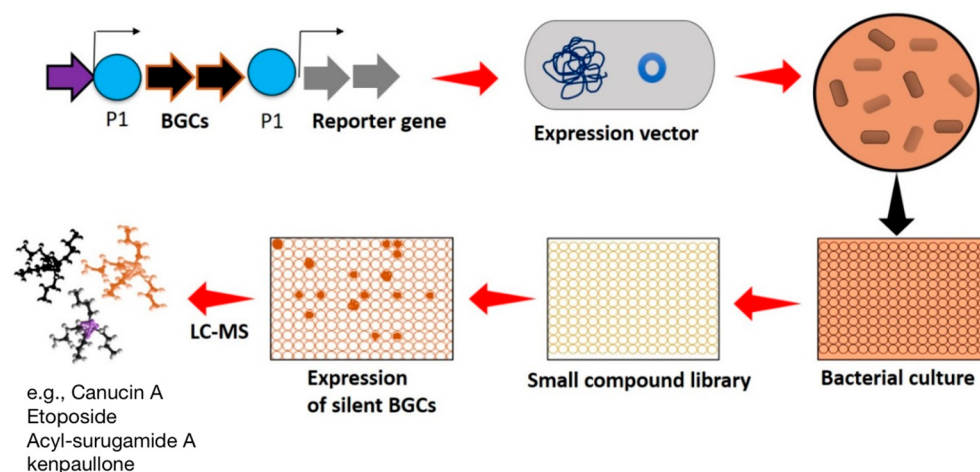


Figure 3. High-throughput elicitor screening (HiTES) technique to activate the silent BGCs. Genetic constructs contain reporter genes used to indicate the expression of cryptic BGCs in the microbes.

The major drawback of the HiTES technique is relying on the genetic construct to readout the expression of NCs. The methodology does not include the structural recognition of expressed cryptic metabolites; therefore, users cannot predict novel or already known compounds. Thus, the HiTES technique does not link the bioactivity of the cryptic metabolites; therefore, activity assays need to be conducted upon the discovery of cryptic compounds. The aforementioned drawbacks of the HiTES limit the use of this technique in the high-throughput screening of natural products.

3.1.1. Imaging Mass Spectrometry in High-Throughput Screening of NCs

Recently, Xu et al. amended HiTES as a genetic-free technique by removing the genetic construct used to read the expressed cryptic metabolites in the microbial culture [103]. In the genetic-free HiTES technique, expression of silent BGCs is triggered by deploying the bacterial culture in hundreds of different (500–1000) cultural conditions (small compound elicitors) at a time.

The HiTES technique coupled with laser ablation electrospray ionization-imaging mass spectrometry (LAESI-IMS) is referred to as 'HiTES-IMS'. Subsequent analysis of the expressed metabolites is carried out in a high-throughput manner using laser ablation electrospray ionization coupled with imaging mass spectrometry (LAESI-IMS). The latest advances in the IMS technique have provided a flexible way for rapid analysis of biological samples for both known and unknown compounds [104]. The IMS analysis requires a very small quantity of samples and provides the facility to analyze specific compounds. There are currently several IMS techniques available to identify NCs [104], but the LAESI coupled with mass spectroscopy (MS/MS) is a relevant technique to detect NCs directly from biological samples in a high-throughput fashion [50].

The LAESI-IMS is a newly emerged technique in which the biological sample absorbs a mid-infrared laser ($\lambda = 2.94 \mu\text{m}$), creating an ablation plume of compounds that are ionized by electrospray and inserted into the mass spectrometer (Figure 4) [105–108]. This technique is capable of identifying a broad range of molecules, including lipids, peptides, alkaloids, phenolics, and several other types of compounds [109]. The LAESI-IMS method has been validated effectively on several biological samples belonging to bacteria, plants,

and fungi [110,111]. This technique uses an ambient ionization method; therefore, samples can be analyzed directly without any prior preparation at atmospheric pressure. Unlike other IMS methods, LAESI often has certain disadvantages, as it is not ideal for analyzing dried samples and cannot differentiate the isobaric ions. However, these problems can be resolved by dissolving the sample in water before imaging [110] and analyzing it using LAESI coupled with MS/MS [111].

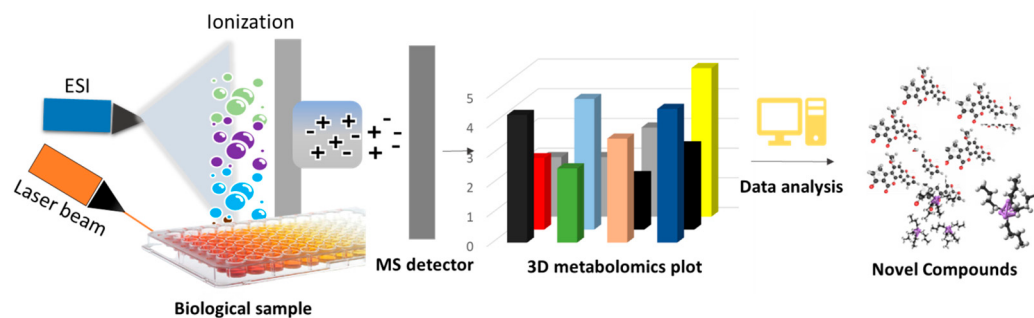


Figure 4. HiTES coupled with laser ablation electrospray ionization-imaging mass spectroscopy (LAESI-IMS) technique for the high-throughput identification of the novel NCs directly from the microbial culture plate.

3.1.2. HiTES Coupled with the IMS Technique in High-Throughput Screening of NCs

The HiTES-IMS technique is an endogenous monoculture strategy for eliciting silent BGCs and detecting the cryptic NCs of a given bacterial strain in a rapid and untargeted fashion. Using HiTES-IMS, screening of novel NCs from known or diverse bacteria has become automated and easier. The HiTES-IMS technique has been effectively validated by inducing the expression of silent BGCs and subsequent identification of cryptic NCs in different bacterial strains belonging to Gram-positive, Gram-negative, and distinct actinomycetes [97]. Regardless of the bacterial species and the genome sequence details, HiTES-IMS can express and identify several cryptic metabolites under different cultural conditions at a time. Compared with conventional genetic-based methods that typically take a couple of weeks to months to express and identify novel cryptic metabolites from a microbial source, HiTES-IMS can do the same job with ultra-accuracy in a few hours or a couple of days. Furthermore, this technique, along with enabling the expression of silent BGCs, may also be effective in enhancing the production of already-recognized bioactive compounds and recognizing the respective elicitor compounds in the library. Li et al. used the LAESI technique for creating 3D images of metabolites by ‘depth profiling’ a bacterial sample [107]. Recently, LAESI-IMS was used to identify novel NCs by screening a library of small compounds (1000 elicitors) to activate the silent BGCs in a throughput manner. They found Canucin A as a novel compound, utilizing kenpaullone as an elicitor [97]. Furthermore, Tamm et al. (2019) described several novel compounds that were identified through the HiTES-IMS technique [98].

Workflow of HiTES-IMS

The selected bacterial cultures in 96-well plates were subjected to HiTES with a 502-member natural-product library [112]. After proper incubation time, the expression of NCs is analyzed using LAESI-IMS directly from the sample plate. The LAESI-IMS generates a raw data file of induced or enhanced metabolites, and the file format varies with the individual manufacturer’s software used. Furthermore, each company’s software generates its own raw data file types that do not lend themselves to open-source software platforms. To address this problem, an open and standard format for IMS datasets has been implemented in the imzML format [113]. The data analysis software of LAESI-IMS extracts the signals of each well above a set cut-off value and depicts the data in both 2D and 3D plots. The *x*-axis of the plot shows the *m/z* value, and the *y*-axis depicts the intensity of individual NCs expressed in the presence of respective elicitors. This approach allows for the detection

of individual compounds, whether their expression is enhanced or triggered by basic visualization. In the presence or absence of elicitors, the metabolomics data produced in individual wells of the 96-well plate can be compared to recognize the induced novel compounds. The same approach can be used to identify the enhanced yield of specific secondary metabolites. Furthermore, 'dereplication' is also an accurate and rapid method for distinguishing the known and unknown (novel) compounds within the sample results [114]. The data on the detected metabolites are extracted by MatLab 3D ver. R2024a software and can be exported to an Excel file. The hit compounds can be eventually validated using flask culture followed by HPLC-MS/MS technique.

During the high-throughput screening of natural products, a large amount of mass spectral data is produced by mass spectrometry. This situation becomes challenging, particularly in the high-throughput screening experiments using the HiTES-IMS technique, which directly penetrate hundreds of biological samples at a time and produce an enormous amount of MS/MS raw data. Such data contain both known and unknown MS/MS spectra. Therefore, an accurate and quick analysis of this data is a daunting task. However, the new dereplication techniques currently available have rendered it possible to rapidly identify known MS/MS spectra and molecular networking, accompanied by the annotation of unknown MS/MS data [115,116].

4. Dereplication of Natural Products

The process in which structural information about a chemical compound is used for its identification among the sample data is called 'dereplication'. The dereplication process has been proven to be an essential step in natural product research. The dereplication method enables the rapid detection of chemical compounds from the crude MS data during the screening of natural products [117]. Thus, dereplication enhances the performance of high-throughput screening of natural products [118]. Several databases, such as PubChem (≈ 83 million compounds), ChemSpider (≈ 58 million), ChEMBL (≈ 2.1 million), ChemBank (≈ 1.2 million), ChEBI (440,000), SciFinder (≈ 161 million), and several other libraries, are available with the structural information of millions of chemical compounds [119]. This approach eventually eliminates the re-isolation and structural identification processes [117].

In recent years, the development of several useful bioinformatics tools has been assisted by the process of dereplication through the identification of known compounds from the huge crude MS/MS data [120,121]. These tools identify the NCs by matching the MS data generated from the experiment to the MS data repositories.

To assist the rapid screening of NCs, dereplication procedures usually consider the combination of various separation strategies, spectroscopy, and database searching methods [119]. The dereplication of NCs can be achieved best with the help of a computer-based strategy called molecular networking (MN). The MN-based dereplication comprised the visualization and interpretation of complex MS/MS data for the rapid identification of the known compounds and molecular networking of the unknown mass spectra in crude MS/MS data [122]. The MN consists of two crucial steps, i.e., (1) organization and visualization of MS/MS datasets and (2) automation in database search. The key strategies for MN-based dereplication include the following: (1) integration of MS data; (2) harnessing mass shift differences; and (3) integration of functional annotation. The integration and visualization of tandem MS data are being achieved based on the similarity of the spectral map. For structurally similar compounds, nodes of similar fragmentation spectra would cluster together and generate the cluster of analogues [123]. The MN utilizes the harnessing of mass shift differences and exploits mass-to-charge (m/z) ratios between related molecules [122]. This strategy facilitates the profiling of meta-mass shift chemicals that allow the identification of known chemical groups and reveal the specific biochemical transformation [124].

More recently, researchers have used the algorithm tool DEREPLICATOR+ (<http://mohimanilab.cbd.cmu.edu/software/>), which utilizes spectral networks for the high-throughput identification of variants of known natural products. Among the sev-

eral available tools, to date, the DEREPLICATOR+ tool has gained wide acceptance in the high-throughput screening of natural products [117]. It has been reportedly improved and facilitates the identification of polyketides, terpenes, benzenoids, alkaloids, flavonoids, etc.

Workflow to Generate MN

The workflow that involves MN-based dereplication is schematically represented in Figure 5. MN is a computer-based approach that extracts and arranges the experimental MS/MS data according to their spectral similarities [125]. Global Natural Product Social (GNPS) is currently the most widely used and accepted online platform for MN (<https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp>). The basic process for producing MN utilizing raw data involves data collection, data file conversion, GNPS-MassIVE upload, MN generation commands, and network analysis accompanied by visualization. For developing a molecular network, the first step is the collection of MS/MS spectra. The file format of the raw data needs to be converted to an open format such as mzXML, mzML, or MGF. These open file formats are accessible on the GNPS-MassIVE platform. After submission of the MS/MS datasets, the GNPS program identifies the known compounds and treats the unknown compounds to generate a molecular network using the spectral similarity of the datasets. The generation of a molecular network can be achieved based on the cosine scores of MS/MS spectra [126]. Basically, cosine scores measure the relatedness in MS/MS spectra. Furthermore, the MS/MS spectral data needs to be converted to mzXML format, which is a text-based format of MS/MS data [123,127]. Finally, MN can be analyzed and visualized using GNPS, but Cytoscape, an open-source platform, can also be used [124]. The generated text file can be imported into Cytoscape to visualize the molecular network. Using Cytoscape, identical molecules in the analogue or compound family can be explored visually.

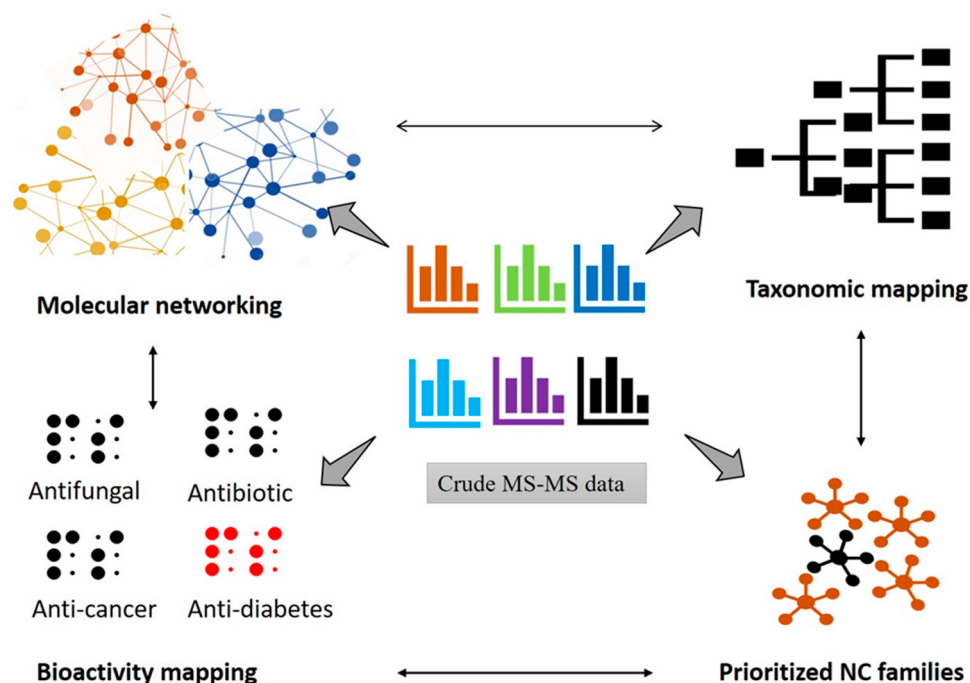


Figure 5. Various applications of crude MS-MS data in GNPS for the generation of molecular networking (MN).

5. Conclusions

In concurrent efforts to bring the antibiotic era back, success in natural product research does not end with the identification of novel NCs from natural sources. The researchers need to modify their routine approaches to screen NCs in a high-throughput fashion. Recently, microbial silent BGCs have emerged as an opportunistic trove for the mining of novel NCs. The interdisciplinary techniques discussed here, such as antiSMASH, HiTES,

and LAESI-IMS, are the latest and validated for their usage in natural product research. The strategic workflow described here begins with the identification of the potential microbial strain containing a large reservoir of silent BGCs, activation followed by the expression of silent BGCs, and finally the identification of novel NCs in a high-throughput fashion. Further, dereplication strategies enhance the rapid identification of known compounds and the generation of a molecular network of unknown MS/MS spectra on the GNPS platform. The sequential application of these strategies will eventually enable the rapid screening of novel NCs.

Author Contributions: Conceptualization, S.N.M., A.W.-B. and K.M.K.; writing—original draft preparation, S.N.M., S.G., M.I. and P.P.; writing—review and editing, S.N.M., A.W.-B. and K.M.K.; supervision, A.W.-B. and K.M.K.; funding acquisition, A.W.-B. and S.N.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded to S.N.M. through Kothari Postdoctoral Fellowship by the UGC-DAE Consortium for Scientific Research, University Grants Commission India, with grant number F4-2/2006(BSR)/BL/18-19/0416.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be provided upon individual request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Demain, A.L. Importance of Microbial Natural Products and the Need to Revitalize Their Discovery. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 185–201. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fischbach, M.A. Antibiotics from Microbes: Converging to Kill. *Curr. Opin. Microbiol.* **2009**, *12*, 520–527. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rubira, C.; de Oliveira Carvalho, P.E.; Cataneo, A.J.; Carvalho, L.R. Antibiotics for Preventing Infection in People Receiving Chest Drains. *Cochrane Database Syst. Rev.* **2017**, *2017*, CD009165. [\[CrossRef\]](#)
- Kusmann, M.; Abe Cunha, D.H.; Berciano, S. Bioactive Compounds for Human and Planetary Health. *Front. Nutr.* **2023**, *10*, 1193848. [\[CrossRef\]](#)
- Singh, T.; Kumar, V.; Rathore, S.; Vyas, A.; Nagarajan, R.; Panwar, H. Natural Bio-Colorant and Pigments: Sources and Applications in Food Processing. *J. Agric. Food Res.* **2023**, *12*, 100628. [\[CrossRef\]](#)
- Rybczyńska-Tkaczyk, K.; Łopusiewicz, Ł.; Bartkowiak, A.; Horubała, A.; Miazga-Karska, M.; Sołowiej, B. Natural Compounds with Antimicrobial Properties in Cosmetics. *Pathogens* **2023**, *12*, 320. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wang, B.; Yao, M.; Lv, L.; Ling, Z.; Li, L. The Human Microbiota in Health and Disease. *Engineering* **2017**, *3*, 71–82. [\[CrossRef\]](#)
- Wilson, M. *The Human Microbiota in Health and Disease: An Ecological and Community-Based Approach*; Garland Science: New York, NY, USA, 2018.
- Methé, B.A.; Nelson, K.E.; Pop, M.; Creasy, H.H.; Giglio, M.G.; Huttenhower, C.; Gevers, D.; Petrosino, J.F.; Abubucker, S.; Badger, J.H.; et al. A Framework for Human Microbiome Research. *Nature* **2012**, *486*, 215.
- Li, F.; Wang, Y.; Li, D.; Chen, Y.; Dou, Q.P. Are We Seeing a Resurgence in the Use of Natural Products for New Drug Discovery? *Expert Opin. Drug Discov.* **2019**, *14*, 417–420. [\[CrossRef\]](#)
- Rouhi, A.M. Rediscovering Natural Products. *Chem. Eng. News* **2003**, *81*, 77–91. [\[CrossRef\]](#)
- dos Santos Nascimento, I.J.; de Moura, R.O. Ligand and Structure-Based Drug Design (LBDD and SBDD): Promising Approaches to Discover New Drugs. In *Applied Computer-Aided Drug Design: Models and Methods*; Bentham Science Publishers: Sharjah, United Arab Emirates, 2023; p. 1.
- Volochnyuk, D.M.; Ryabukhin, S.V.; Moroz, Y.S.; Savych, O.; Chuprina, A.; Horvath, D.; Zabolotna, Y.; Varnek, A.; Judd, D.B. Evolution of Commercially Available Compounds for HTS. *Drug Discov. Today* **2019**, *24*, 390–402. [\[CrossRef\]](#) [\[PubMed\]](#)
- Siddiqui, A.A.; Iram, F.; Siddiqui, S.; Sahu, K. Role of Natural Products in Drug Discovery Process. *Int. J. Drug Dev. Res.* **2014**, *6*, 172–204.
- DiMasi, J.A.; Hansen, R.W.; Grabowski, H.G. The Price of Innovation: New Estimates of Drug Development Costs. *J. Health Econ.* **2003**, *22*, 151–185. [\[CrossRef\]](#)
- Morgan, S.; Grootendorst, P.; Lexchin, J.; Cunningham, C.; Greyson, D. The Cost of Drug Development: A Systematic Review. *Health Policy* **2001**, *100*, 4–17. [\[CrossRef\]](#) [\[PubMed\]](#)
- Paul, S.M.; Mytelka, D.S.; Dunwiddie, C.T.; Persinger, C.C.; Munos, B.H.; Lindborg, S.R.; Schacht, A.L. How to Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge. *Nat. Rev. Drug Discov.* **2010**, *9*, 203–214.
- Cohen, M.L. Changing Patterns of Infectious Disease. *Nature* **2000**, *406*, 762–767. [\[CrossRef\]](#) [\[PubMed\]](#)
- Smith, K.F.; Goldberg, M.; Rosenthal, S.; Carlson, L.; Chen, J.; Chen, C.; Ramachandran, S. Global Rise in Human Infectious Disease Outbreaks. *J. R. Soc. Interface* **2014**, *11*, 20140950. [\[CrossRef\]](#)

20. Yoneyama, H.; Katsumata, R. Antibiotic Resistance in Bacteria and Its Future for Novel Antibiotic Development. *Biosci. Biotechnol. Biochem.* **2006**, *70*, 1060–1075. [CrossRef] [PubMed]
21. Silver, L.L.; Bostian, K.A. Discovery and Development of New Antibiotics: The Problem of Antibiotic Resistance. *Antimicrob. Agents Chemother.* **1993**, *37*, 377–383. [CrossRef]
22. Seneci, P.; Miertus, S. Combinatorial Chemistry and High-Throughput Screening in Drug Discovery: Different Strategies and Formats. *Mol. Divers.* **2000**, *5*, 75–89. [CrossRef]
23. Appell, K.; Baldwin, J.J.; Egan, W.J. Combinatorial Chemistry and High-Throughput Screening in Drug Discovery and Development. In *Handbook of Modern Pharmaceutical Analysis*; Elsevier: Amsterdam, The Netherlands, 2001; Volume 3, pp. 23–56.
24. Kodadek, T. The Rise, Fall and Reinvention of Combinatorial Chemistry. *Chem. Comm.* **2011**, *47*, 9757–9763. [CrossRef] [PubMed]
25. Djaballah, H. Chemical Space, High Throughput Screening and the World of Blockbuster Drugs. DDW Spring. 2013. Available online: <https://www.ddw-online.com/chemical-space-high-throughput-screening-and-the-world-of-blockbuster-drugs-1528-201304/> (accessed on 1 July 2023).
26. Bérdy, J. Thoughts and Facts about Antibiotics: Where We Are Now and Where We Are Heading. *J. Antibiot.* **2012**, *65*, 385–395. [CrossRef] [PubMed]
27. Powers, J.H. Antimicrobial Drug Development—The Past, the Present, and the Future. *Clin. Microbiol. Infect.* **2004**, *10*, 23–31. [CrossRef]
28. Projan, S.J.; Shlaes, D.M. Antibacterial Drug Discovery: Is It All Downhill from Here? *Clin. Microbiol. Infect.* **2004**, *10*, 18–22. [CrossRef]
29. Gould, K. Antibiotics: From Prehistory to the Present Day. *J. Antimicrob. Chemother.* **2016**, *71*, 572–575. [CrossRef]
30. Rokem, J.S.; Lantz, A.E.; Nielsen, J. Systems Biology of Antibiotic Production by Microorganisms. *Nat. Prod. Rep.* **2007**, *24*, 1262–1287. [CrossRef] [PubMed]
31. Gaudêncio, S.P.; Pereira, F.; Barata, T.; Vasconcelos, C. Advanced Methods for Natural Products Discovery: Bioactivity Screening, Dereplication, Metabolomics Profiling, Genomic Sequencing, Databases and Informatic Tools, and Structure Elucidation. *Mar. Drugs* **2023**, *21*, 308. [CrossRef] [PubMed]
32. Breinbauer, R.; Manger, M.; Scheck, M.; Waldmann, H. Natural Product Guided Compound Library Development. *Curr. Med. Chem.* **2002**, *9*, 2129–2145. [CrossRef]
33. Ayon, N.J. High-Throughput Screening of Natural Product and Synthetic Molecule Libraries for Antibacterial Drug Discovery. *Metabolites* **2023**, *13*, 625. [CrossRef]
34. Breinbauer, R.; Vetter, I.R.; Waldmann, H. From Protein Domains to Drug Candidates—Natural Products as Guiding Principles in the Design and Synthesis of Compound Libraries. *Angew. Chem.* **2002**, *114*, 2968–2990. [CrossRef]
35. Paululat, T.; Tang, Y.Q.; Grabley, S.; Thiericke, R. Combinatorial Chemistry: The Impact of Natural Products. *Chimica Oggi* **1999**, *17*, 52–56.
36. Procópio, R.E.L.; Silva, I.R.; Martins, M.K.; Azevedo, J.L.; Araújo, J.M. Antibiotics Produced by *Streptomyces*. *Braz. J. Infect. Dis.* **2012**, *16*, 466–471. [CrossRef] [PubMed]
37. Bode, H.B.; Bethe, B.; Höfs, R.; Zeeck, A. Big Effects from Small Changes: Possible Ways to Explore Nature’s Chemical Diversity. *ChemBioChem* **2002**, *3*, 619–627. [CrossRef] [PubMed]
38. Schiewe, H.J.; Zeeck, A. Cineromycins, γ -Butyrolactones and Ansamycins by Analysis of the Secondary Metabolite Pattern Created by a Single Strain of *Streptomyces*. *J. Antibiot.* **1999**, *52*, 635–642. [CrossRef] [PubMed]
39. Hardt, I.H.; Steinmetz, H.; Gerth, K.; Sasse, F.; Reichenbach, H.; Höfle, G. New Natural Epothilones from *Sorangium cellulosum*, Strains So ce90/B2 and So ce90/D13: Isolation, Structure Elucidation, and SAR Studies. *J. Nat. Prod.* **2001**, *64*, 847–856. [CrossRef] [PubMed]
40. Wink, M. Genes of Secondary Metabolism: Differential Expression in Plants and in Vitro Cultures and Functional Expression in Genetically Transformed Microorganisms. In *Primary and Secondary Metabolism of Plant Cell Cultures II*; Kurz, W.G.W., Ed.; Springer: Berlin/Heidelberg, Germany, 1989; pp. 239–251.
41. Martin, J.F.; Liras, P. Organization and Expression of Genes Involved in the Biosynthesis of Antibiotics and Other Secondary Metabolites. *Annu. Rev. Microbiol.* **1989**, *43*, 173–206. [CrossRef] [PubMed]
42. Galperin, M.Y.; Koonin, E.V. Searching for Drug Targets in Microbial Genomes. *Curr. Opin. Biotechnol.* **1999**, *10*, 571–578. [CrossRef]
43. Hurley, L.H. DNA and Associated Targets for Drug Design. *J. Med. Chem.* **1989**, *32*, 2027–2033. [CrossRef] [PubMed]
44. Zheng, Y.; Saitou, A.; Wang, C.M.; Toyoda, A.; Minakuchi, Y.; Sekiguchi, Y.; Ueda, K.; Takano, H.; Sakai, Y.; Abe, K.; et al. Genome Features and Secondary Metabolites Biosynthetic Potential of the Class Ktedonobacteria. *Front. Microbiol.* **2019**, *10*, 893. [CrossRef]
45. Bok, J.W.; Hoffmeister, D.; Maggio-Hall, L.A.; Murillo, R.; Glasner, J.D.; Keller, N.P. Genomic Mining for *Aspergillus* Natural Products. *Chem. Biol.* **2006**, *13*, 31–37. [CrossRef]
46. Cao, H. Advances in Mining and Expressing Microbial Biosynthetic Gene Clusters. *Crit. Rev. Microbiol.* **2023**, *49*, 18–37.
47. Mao, D.; Okada, B.K.; Wu, Y.; Xu, F.; Seyedsayamdost, M.R. Recent Advances in Activating Silent Biosynthetic Gene Clusters in Bacteria. *Curr. Opin. Microbiol.* **2018**, *45*, 156. [CrossRef] [PubMed]
48. Reen, F.J.; Romano, S.; Dobson, A.D.; O’Gara, F. The Sound of Silence: Activating Silent Biosynthetic Gene Clusters in Marine Microorganisms. *Mar. Drugs* **2015**, *13*, 4754–4783. [CrossRef]
49. Ren, H.; Wang, B.; Zhao, H. Breaking the Silence: New Strategies for Discovering Novel Natural Products. *Curr. Opin. Biotechnol.* **2017**, *48*, 21–27. [CrossRef] [PubMed]

50. Van der Meij, A.; Worsley, S.F.; Hutchings, M.I.; van Wezel, G.P. Chemical Ecology of Antibiotic Production by Actinomycetes. *FEMS Microbiol. Rev.* **2017**, *41*, 392–416. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Ahmad, S.N.M. Biosynthetic Gene Cluster Evaluation-Genome Mining for Natural Product Formation. Ph.D. Dissertation, Technische Universität München, Munich, Germany, 2023.
52. Iqbal, H.A.; Feng, Z.; Brady, S.F. Biocatalysts and Small Molecule Products from Metagenomic Studies. *Curr. Opin. Chem. Biol.* **2012**, *16*, 109–116. [\[CrossRef\]](#)
53. Wang, H.; Fewer, D.P.; Holm, L.; Rouhiainen, L.; Sivonen, K. Atlas of Nonribosomal Peptide and Polyketide Biosynthetic Pathways Reveals Common Occurrence of Nonmodular Enzymes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9259–9264. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Hug, J.J.; Bader, C.D.; Remškar, M.; Cirnski, K.; Müller, R. Concepts and Methods to Access Novel Antibiotics from Actinomycetes. *Antibiotics* **2018**, *7*, 44. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Katz, L.; Baltz, R.H. Natural Product Discovery: Past, Present, and Future. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 155–176. [\[CrossRef\]](#)
56. Zhang, W.; Wang, L.; Kong, L.; Wang, T.; Chu, Y.; Deng, Z.; You, D. Unveiling the Post-PKS Redox Tailoring Steps in Biosynthesis of the Type II Polyketide Antitumor Antibiotic Xantholipin. *Chem. Biol.* **2012**, *19*, 422–432. [\[CrossRef\]](#)
57. Miao, V.; Coeffet-LeGal, M.F.; Brian, P.; Brost, R.; Penn, J.; Whiting, A.; Martin, S.; Ford, R.; Parr, I.; Bouchard, M.; et al. Daptomycin Biosynthesis in *Streptomyces roseosporus*: Cloning and Analysis of the Gene Cluster and Revision of Peptide Stereochemistry. *Microbiology* **2005**, *151*, 1507–1523. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Park, S.R.; Yoo, Y.J.; Ban, Y.H.; Yoon, Y.J. Biosynthesis of Rapamycin and Its Regulation: Past Achievements and Recent Progress. *J. Antibiot.* **2010**, *63*, 434–441. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Staunton, J.; Weissman, K.J. Polyketide Biosynthesis: A Millennium Review. *J. Nat. Prod. Rep.* **2001**, *18*, 380–416. [\[CrossRef\]](#)
60. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79*, 629–661. [\[CrossRef\]](#)
61. Meesil, W.; Sermwittayawong, N.; Intra, B.; Kitani, S.; Thancharoen, A.; Duangrattanaalert, K.; Igarashi, Y. Genome Mining Reveals Novel Biosynthetic Gene Clusters in Entomopathogenic Bacteria. *Sci. Rep.* **2023**, *13*, 20764. [\[CrossRef\]](#)
62. Adamek, M.; Spohn, M.; Stegmann, E.; Ziemert, N. Mining Bacterial Genomes for Secondary Metabolite Gene Clusters. In *Antibiotics*; Sass, P., Ed.; Humana Press: New York, NY, USA, 2017; pp. 23–47.
63. Medema, M.H.; Blin, K.; Cimermanic, P.; De Jager, V.; Zakrzewski, P.; Fischbach, M.A.; Weber, T.; Takano, E.; Breitling, R. antiSMASH: Rapid Identification, Annotation and Analysis of Secondary Metabolite Biosynthesis Gene Clusters in Bacterial and Fungal Genome Sequences. *Nucleic Acids Res.* **2011**, *39* (Suppl. 2), W339–W346. [\[CrossRef\]](#)
64. Li, Z.; Zhu, D.; Shen, Y. Discovery of Novel Bioactive Natural Products Driven by Genome Mining. *Drug Discov. Ther.* **2018**, *12*, 318–328. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Khaldi, N.; Seifuddin, F.T.; Turner, G.; Haft, D.; Nierman, W.C.; Wolfe, K.H.; Fedorova, N.D. SMURF: Genomic Mapping of Fungal Secondary Metabolite Clusters. *Fungal Genet. Biol.* **2010**, *47*, 736–741. [\[CrossRef\]](#)
66. Li, M.H.; Ung, P.M.; Zajkowski, J.; Garneau-Tsodikova, S.; Sherman, D.H. Automated Genome Mining for Natural Products. *BMC Bioinform.* **2009**, *10*, 185. [\[CrossRef\]](#)
67. Weber, T.; Rausch, C.; Lopez, P.; Hoof, I.; Gaykova, V.; Huson, D.H.; Wohlleben, W. CLUSEAN: A Computer-Based Framework for the Automated Analysis of Bacterial Secondary Metabolite Biosynthetic Gene Clusters. *J. Biotechnol.* **2009**, *140*, 13–17. [\[CrossRef\]](#)
68. Starcevic, A.; Zucko, J.; Simunkovic, J.; Long, P.F.; Cullum, J.; Hranueli, D. ClustScan: An Integrated Program Package for the Semi-Automatic Annotation of Modular Biosynthetic Gene Clusters and In Silico Prediction of Novel Chemical Structures. *Nucleic Acids Res.* **2008**, *36*, 6882–6892. [\[CrossRef\]](#)
69. Umemura, M.; Koike, H.; Nagano, N.; Ishii, T.; Kawano, J.; Yamane, N.; Kozono, I.; Horimoto, K.; Shin-ya, K.; Asai, K.; et al. MIDDAS-M: Motif-Independent De Novo Detection of Secondary Metabolite Gene Clusters through the Integration of Genome Sequencing and Transcriptome Data. *PLoS ONE* **2013**, *8*, e84028. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Wolf, T.; Shelest, V.; Nath, N.; Shelest, E. CASSIS and SMIPS: Promoter-Based Prediction of Secondary Metabolite Gene Clusters in Eukaryotic Genomes. *Bioinformatics* **2016**, *32*, 1138–1143. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Yi, G.; Sze, S.H.; Thon, M.R. Identifying Clusters of Functionally Related Genes in Genomes. *Bioinformatics* **2007**, *23*, 1053–1060. [\[CrossRef\]](#) [\[PubMed\]](#)
72. Cragg, G.M.; Newman, D.J. Natural Products: A Continuing Source of Novel Drug Leads. *Biochim. Biophys. Acta Gen. Subj.* **2013**, *1830*, 3670–3695. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **2012**, *75*, 311–335. [\[CrossRef\]](#)
74. Ziemert, N.; Alanjary, M.; Weber, T. The Evolution of Genome Mining in Microbes—A Review. *Nat. Prod. Rep.* **2016**, *33*, 988–1005. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Blin, K.; Chevrette, M.G.; Lu, X.; Schwalen, C.J.; Kautsar, S.A.; Suarez Duran, H.G.; De Los Santos, E.L.; Kim, H.U.; Nave, M.; Dickschat, J.S. antiSMASH 4.0—Improvements in Chemistry Prediction and Gene Cluster Boundary Identification. *Nucleic Acids Res.* **2017**, *45*, W36–W41. [\[CrossRef\]](#)
76. Blin, K.; Pascal Andreu, V.; De Los Santos, E.L.; Del Carratore, F.; Lee, S.Y.; Medema, M.H.; Weber, T. The antiSMASH Database Version 2: A Comprehensive Resource on Secondary Metabolite Biosynthetic Gene Clusters. *Nucleic Acids Res.* **2019**, *47*, D625–D630. [\[CrossRef\]](#)

77. Blin, K.; Medema, M.H.; Kottmann, R.; Lee, S.Y.; Weber, T. The antiSMASH Database, a Comprehensive Database of Microbial Secondary Metabolite Biosynthetic Gene Clusters. *Nucleic Acids Res.* **2016**, *23*, gkw960. [[CrossRef](#)] [[PubMed](#)]
78. Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S.Y.; Medema, M.H.; Weber, T. antiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47*, W81–W87. [[CrossRef](#)] [[PubMed](#)]
79. Villebro, R.; Shaw, S.; Blin, K.; Weber, T. Sequence-Based Classification of Type II Polyketide Synthase Biosynthetic Gene Clusters for antiSMASH. *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 469–475. [[CrossRef](#)]
80. Bauer, J.S.; Ghequire, M.G.; Nett, M.; Josten, M.; Sahl, H.G.; De Mot, R.; Gross, H. Biosynthetic Origin of the Antibiotic Pseudopyronines A and B in *Pseudomonas putida* BW11M1. *Chem. Bio Chem.* **2015**, *16*, 2491–2497. [[CrossRef](#)]
81. Ren, H.; Shi, C.; Zhao, H. Computational Tools for Discovering and Engineering Natural Product Biosynthetic Pathways. *iScience* **2020**, *23*, 100795. [[CrossRef](#)] [[PubMed](#)]
82. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H.U.; Brucoleri, R.; Lee, S.Y.; Fischbach, M.A.; Müller, R.; Wohlleben, W.; et al. antiSMASH 3.0—A Comprehensive Resource for the Genome Mining of Biosynthetic Gene Clusters. *Nucleic Acids Res.* **2015**, *43*, W237–W243. [[CrossRef](#)]
83. Baral, B.; Akhgari, A.; Metsä-Ketelä, M. Activation of Microbial Secondary Metabolic Pathways: Avenues and Challenges. *Synth. Syst. Biotechnol.* **2018**, *3*, 163–178. [[CrossRef](#)]
84. Dubey, M.K.; Meena, M.; Aamir, M.; Zehra, A.; Upadhyay, R.S. Regulation and Role of Metal Ions in Secondary Metabolite Production by Microorganisms. In *New and Future Developments in Microbial Biotechnology and Bioengineering*; Gupta, V.K., Ed.; Elsevier: New York, NY, USA, 2019; pp. 259–277.
85. Bibb, M.J. Regulation of Secondary Metabolism in Streptomycetes. *Curr. Opin. Microbiol.* **2005**, *8*, 208–215. [[CrossRef](#)] [[PubMed](#)]
86. Stutzman-Engwall, K.J.; Otten, S.L.; Hutchinson, C.R. Regulation of Secondary Metabolism in *Streptomyces* spp. and Overproduction of Daunorubicin in *Streptomyces peucetius*. *J. Bacteriol. Res.* **1992**, *174*, 144–154. [[CrossRef](#)]
87. Ochi, K.; Okamoto, S.; Tozawa, Y.; Inaoka, T.; Hosaka, T.; Xu, J.; Kurosawa, K. Ribosome Engineering and Secondary Metabolite Production. *Adv. Appl. Microbiol.* **2004**, *56*, 155–179.
88. Ashby, M.; Valley, M.; Shoemaker, D.D. Targeted Methods of Drug Screening Using Co-Culture Methods. U.S. Patent 6518035, 11 February 2003.
89. Malpartida, F.; Hopwood, D.A. Molecular Cloning of the Whole Biosynthetic Pathway of a Streptomyces Antibiotic and Its Expression in a Heterologous Host. *Nature* **1984**, *309*, 462–464. [[CrossRef](#)]
90. Wang, Y.; Tao, Z.; Zheng, H.; Zhang, F.; Long, Q.; Deng, Z.; Tao, M. Iteratively Improving Natamycin Production in *Streptomyces gilvosporeus* by a Large Operon-Reporter Based Strategy. *Metab. Eng.* **2016**, *38*, 418–426. [[CrossRef](#)] [[PubMed](#)]
91. Potharla, V.Y.; Wang, C.; Cheng, Y.Q. Identification and Characterization of the Spiruchostatin Biosynthetic Gene Cluster Enable Yield Improvement by Overexpressing a Transcriptional Activator. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 1457–1465. [[CrossRef](#)] [[PubMed](#)]
92. Baltz, R.H. Gifted Microbes for Genome Mining and Natural Product Discovery. *J. Ind. Microbiol. Biotechnol.* **2017**, *44*, 573–588. [[CrossRef](#)] [[PubMed](#)]
93. Okada, B.K.; Seyedsayamdost, M.R. Antibiotic Dialogues: Induction of Silent Biosynthetic Gene Clusters by Exogenous Small Molecules. *FEMS Microbiol. Rev.* **2017**, *41*, 19–33. [[CrossRef](#)] [[PubMed](#)]
94. Nah, H.J.; Pyeon, H.R.; Kang, S.H.; Choi, S.S.; Kim, E.S. Cloning and Heterologous Expression of a Large-Sized Natural Product Biosynthetic Gene Cluster in *Streptomyces* Species. *Front. Microbiol.* **2017**, *8*, 394. [[CrossRef](#)] [[PubMed](#)]
95. Lopatniuk, M.; Myronovskiy, M.; Nottebrock, A.; Busche, T.; Kalinowski, J.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A. Effect of “Ribosome Engineering” on the Transcription Level and Production of *S. albus* Indigenous Secondary Metabolites. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 7097–7110. [[CrossRef](#)] [[PubMed](#)]
96. Hoshino, S.; Onaka, H.; Abe, I. Activation of Silent Biosynthetic Pathways and Discovery of Novel Secondary Metabolites in Actinomycetes by Co-Culture with Mycolic Acid-Containing Bacteria. *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 363–374. [[CrossRef](#)] [[PubMed](#)]
97. Xu, M.; Wright, G.D. Heterologous Expression-Facilitated Natural Products’ Discovery in Actinomycetes. *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 415–431. [[CrossRef](#)]
98. Tomm, H.A.; Ucciferri, L.; Ross, A.C. Advances in Microbial Culturing Conditions to Activate Silent Biosynthetic Gene Clusters for Novel Metabolite Production. *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 1381–1400. [[CrossRef](#)]
99. Zammit, G.; Zammit, M.G.; Buttigieg, K.G. Emerging Technologies for the Discovery of Novel Diversity in Cyanobacteria and Algae and the Elucidation of Their Valuable Metabolites. *Diversity* **2023**, *15*, 1142. [[CrossRef](#)]
100. 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](#)] [[PubMed](#)]
101. Nowak, V. New Methods for the Discovery of Natural Products from Understudied and Uncultivated Bacterial Phyla. Ph.D. Dissertation, Open Access Te Herenga Waka-Victoria University of Wellington, Wellington, New Zealand, 2023.
102. Craney, A.; Ozimok, C.; Pimentel-Elardo, S.M.; Capretta, A.; Nodwell, J.R. Chemical Perturbation of Secondary Metabolism Demonstrates Important Links to Primary Metabolism. *Chem. Biol.* **2012**, *19*, 1020–1027. [[CrossRef](#)] [[PubMed](#)]
103. Xu, F.; Wu, Y.; Zhang, C.; Davis, K.M.; Moon, K.; Bushin, L.B.; Seyedsayamdost, M.R. A Genetics-Free Method for High-Throughput Discovery of Cryptic Microbial Metabolites. *Nat. Chem. Biol.* **2019**, *15*, 161–168. [[CrossRef](#)]
104. Rosen, P.C.; Seyedsayamdost, M.R. Though Much Is Taken, Much Abides: Finding New Antibiotics Using Old Ones. *Biochemistry* **2017**, *56*, 4925–4926. [[CrossRef](#)] [[PubMed](#)]

105. Spraker, J.E.; Luu, G.T.; Sanchez, L.M. Imaging Mass Spectrometry for Natural Products Discovery: A Review of Ionization Methods. *Nat. Prod. Rep.* **2020**, *37*, 150–162. [[CrossRef](#)] [[PubMed](#)]
106. Nemes, P.; Vertes, A. Laser Ablation Electrospray Ionization for Atmospheric Pressure, *In Vivo*, and Imaging Mass Spectrometry. *Anal. Chem.* **2007**, *79*, 8098–8106. [[CrossRef](#)]
107. Li, H.; Balan, P.; Vertes, A. Molecular Imaging of Growth, Metabolism, and Antibiotic Inhibition in Bacterial Colonies by Laser Ablation Electrospray Ionization Mass Spectrometry. *Angew. Chem.* **2016**, *128*, 15259–15263. [[CrossRef](#)]
108. Fincher, J.A.; Korte, A.R.; Reschke, B.; Morris, N.J.; Powell, M.J.; Vertes, A. Enhanced Sensitivity and Metabolite Coverage with Remote Laser Ablation Electrospray Ionization-Mass Spectrometry Aided by Coaxial Plume and Gas Dynamics. *Analyst* **2017**, *142*, 3157–3164. [[CrossRef](#)] [[PubMed](#)]
109. Li, H.; Vertes, A. Solvent Gradient Electrospray for Laser Ablation Electrospray Ionization Mass Spectrometry. *Analyst* **2017**, *142*, 2921–2927. [[CrossRef](#)]
110. Etalo, D.W.; De Vos, R.C.; Joosten, M.H.; Hall, R.D. Spatially Resolved Plant Metabolomics: Some Potentials and Limitations of Laser-Ablation Electrospray Ionization Mass Spectrometry Metabolite Imaging. *Plant Physiol.* **2015**, *169*, 1424–1435. [[CrossRef](#)]
111. Stopka, S.A.; Agtuca, B.J.; Koppenaal, D.W.; Paša-Tolić, L.; Stacey, G.; Vertes, A.; Anderton, C.R. Laser-Ablation Electrospray Ionization Mass Spectrometry with Ion Mobility Separation Reveals Metabolites in the Symbiotic Interactions of Soybean Roots and Rhizobia. *Plant J.* **2017**, *91*, 340–354. [[CrossRef](#)] [[PubMed](#)]
112. Poksay, K.S.; Sheffler, D.J.; Spilman, P.; Campagna, J.; Jagodzinska, B.; Descamps, O.; Gorostiza, O.; Matalis, A.; Mullenix, M.; Bredesen, D.E.; et al. Screening for Small Molecule Inhibitors of Statin-Induced APP C-Terminal Toxic Fragment Production. *Front. Pharmacol.* **2017**, *8*, 46. [[CrossRef](#)] [[PubMed](#)]
113. Schramm, T.; Hester, Z.; Klinkert, I.; Both, J.P.; Heeren, R.M.; Brunelle, A.; Laprévote, O.; Desbenoit, N.; Robbe, M.F.; Stoeckli, M.; et al. imzML—A Common Data Format for the Flexible Exchange and Processing of Mass Spectrometry Imaging Data. *J. Proteomics* **2012**, *75*, 5106–5110. [[CrossRef](#)]
114. Konishi, Y.; Kiyota, T.; Draghici, C.; Gao, J.M.; Yeboah, F.; Acoca, S.; Jarussophon, S.; Purisima, E. Molecular Formula Analysis by an MS/MS/MS Technique to Expedite Dereplication of Natural Products. *Anal. Chem.* **2007**, *79*, 1187–1197. [[CrossRef](#)] [[PubMed](#)]
115. Allard, P.M.; Genta-Jouve, G.; Wolfender, J.L. Deep Metabolome Annotation in Natural Products Research: Towards a Virtuous Cycle in Metabolite Identification. *Curr. Opin. Chem. Biol.* **2017**, *36*, 40–49. [[CrossRef](#)] [[PubMed](#)]
116. Kind, T.; Tsugawa, H.; Cajka, T.; Ma, Y.; Lai, Z.; Mehta, S.S.; Wohlgemuth, G.; Barupal, D.K.; Showalter, M.R.; Arita, M.; et al. Identification of Small Molecules Using Accurate Mass MS/MS Search. *Mass Spectrom. Rev.* **2018**, *37*, 513–532. [[CrossRef](#)] [[PubMed](#)]
117. Mohimani, H.; Gurevich, A.; Shlemov, A.; Mikheenko, A.; Korobeynikov, A.; Cao, L.; Shcherbin, E.; Nothias, L.F.; Dorrestein, P.C.; Pevzner, P.A. Dereplication of Microbial Metabolites through Database Search of Mass Spectra. *Nat. Commun.* **2018**, *9*, 4035. [[CrossRef](#)] [[PubMed](#)]
118. Kildgaard, S.; Subko, K.; Phillips, E.; Goidts, V.; De la Cruz, M.; Díaz, C.; Gotfredsen, C.H.; Andersen, B.; Frisvad, J.C.; Nielsen, K.F.; et al. A Dereplication and Bioguided Discovery Approach to Reveal New Compounds from a Marine-Derived Fungus *Stilbella fimetaria*. *Mar. Drugs* **2017**, *15*, 253. [[CrossRef](#)]
119. Zani, C.L.; Carroll, A.R. Database for Rapid Dereplication of Known Natural Products Using Data from MS and Fast NMR Experiments. *J. Nat. Prod.* **2017**, *80*, 1758–1766. [[CrossRef](#)]
120. Gaudêncio, S.P.; Pereira, F. Dereplication: Racing to Speed up the Natural Products Discovery Process. *Nat. Prod. Rep.* **2015**, *32*, 779–810. [[CrossRef](#)]
121. Hubert, J.; Nuzillard, J.M.; Renault, J.H. Dereplication Strategies in Natural Product Research: How Many Tools and Methodologies Behind the Same Concept? *Phytochem. Rev.* **2017**, *16*, 55–95. [[CrossRef](#)]
122. Ramos, A.E.; Evanno, L.; Poupon, E.; Champy, P.; Benididir, M.A. Natural Products Targeting Strategies Involving Molecular Networking: Different Manners, One Goal. *Nat. Prod. Rep.* **2019**, *36*, 960–980. [[CrossRef](#)] [[PubMed](#)]
123. Olivon, F.; Allard, P.M.; Koval, A.; Righi, D.; Genta-Jouve, G.; Neyts, J.; Apel, C.; Pannecouque, C.; Nothias, L.F.; Cachet, X.; et al. Bioactive Natural Products Prioritization Using Massive Multi-Informational Molecular Networks. *ACS Chem. Biol.* **2017**, *12*, 2644–2651. [[CrossRef](#)]
124. Hartmann, A.C.; Petras, D.; Quinn, R.A.; Protsyuk, I.; Archer, F.I.; Ransome, E.; Williams, G.J.; Bailey, B.A.; Vermeij, M.J.; Alexandrov, T.; et al. Meta-Mass Shift Chemical Profiling of Metabolomes from Coral Reefs. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 11685–11690. [[CrossRef](#)]
125. 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](#)] [[PubMed](#)]
126. Yang, J.Y.; Sanchez, L.M.; Rath, C.M.; Liu, X.; Boudreau, P.D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; et al. Molecular Networking as a Dereplication Strategy. *J. Nat. Prod.* **2013**, *76*, 1686–1699. [[CrossRef](#)]
127. 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.

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.