

Review

Applications and Restrictions of Integrated Genomic and Metabolomic Screening: An Accelerator for Drug Discovery from Actinomycetes?

Janina Krause 

Abteilung Biomedizinische Grundlagen 1, Institut für Gesundheitsforschung und Bildung,
Universität Osnabrück, 49076 Osnabrück, Germany; janina.krause@uni-osnabrueck.de

Abstract: Since the golden age of antibiotics in the 1950s and 1960s actinomycetes have been the most prolific source for bioactive natural products. However, the number of discoveries of new bioactive compounds decreases since decades. New procedures (e.g., activating strategies or innovative fermentation techniques) were developed to enhance the productivity of actinomycetes. Nevertheless, compound identification remains challenging among others due to high rediscovery rates. Rapid and cheap genome sequencing as well as the advent of bioinformatical analysis tools for biosynthetic gene cluster identification in combination with mass spectrometry-based molecular networking facilitated the tedious process of dereplication. In recent years several studies have been dedicated to accessing the biosynthetic potential of *Actinomyces* species, especially streptomycetes, by using integrated genomic and metabolomic screening in order to boost the discovery rate of new antibiotics. This review aims to present the various possible applications of this approach as well as the newly discovered molecules, covering studies between 2014 and 2021. Finally, the effectiveness of this approach with regard to find new bioactive agents from actinomycetes will be evaluated.

Keywords: bioactive natural products; actinomycetes; genome mining; biosynthetic gene cluster; dereplication; molecular networking



Citation: Krause, J. Applications and Restrictions of Integrated Genomic and Metabolomic Screening: An Accelerator for Drug Discovery from Actinomycetes? *Molecules* **2021**, *26*, 5450. <https://doi.org/10.3390/molecules26185450>

Academic Editor: Anna Petruczynik

Received: 4 August 2021

Accepted: 6 September 2021

Published: 7 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. 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

Ever since Alexander Fleming discovered the ability of the fungus *Penicillium rubens* (*notatum*) to produce an antibiotic substance in 1929, microbes have been a rich source for bioactive molecules [1]. Among microbes actinomycetes are the most productive order known: e.g., between 1940 and 1970, 70% of all antibiotics or derivatives thereof have been isolated from actinomycetes [2]. Besides antibiotics, natural products with antiproliferative, immunosuppressant, herbicidal and antiviral activity have been isolated from actinomycetes [3,4].

While early searches for bioactive agents were based simply on testing the bioactive capacities of raw culture extracts on diffusion-assays [5], more elaborated methods, like target-based approaches and high-throughput screening, had to be applied when the rediscovery rate of already described natural products increased [6]. At the beginning of the century, the search for new bioactive substances was boosted by the possibility of relatively cheap whole-genome sequencing with next-generation sequencing techniques and the development of bioinformatic methods for genome mining [7]. To identify the genes, which are relevant for the synthesis of a secondary metabolite, it comes in handy that these genes generally accumulate in one contiguous region in the genome called biosynthetic gene cluster (BGC) [8].

In recent years, genome mining has been a main driver of natural product discovery [9]. Many bioinformatic tools have been developed to access the biosynthetic potential of microbes. Several of those tools will be referred to in the Section 2. These shall be introduced in the following (Figure 1).

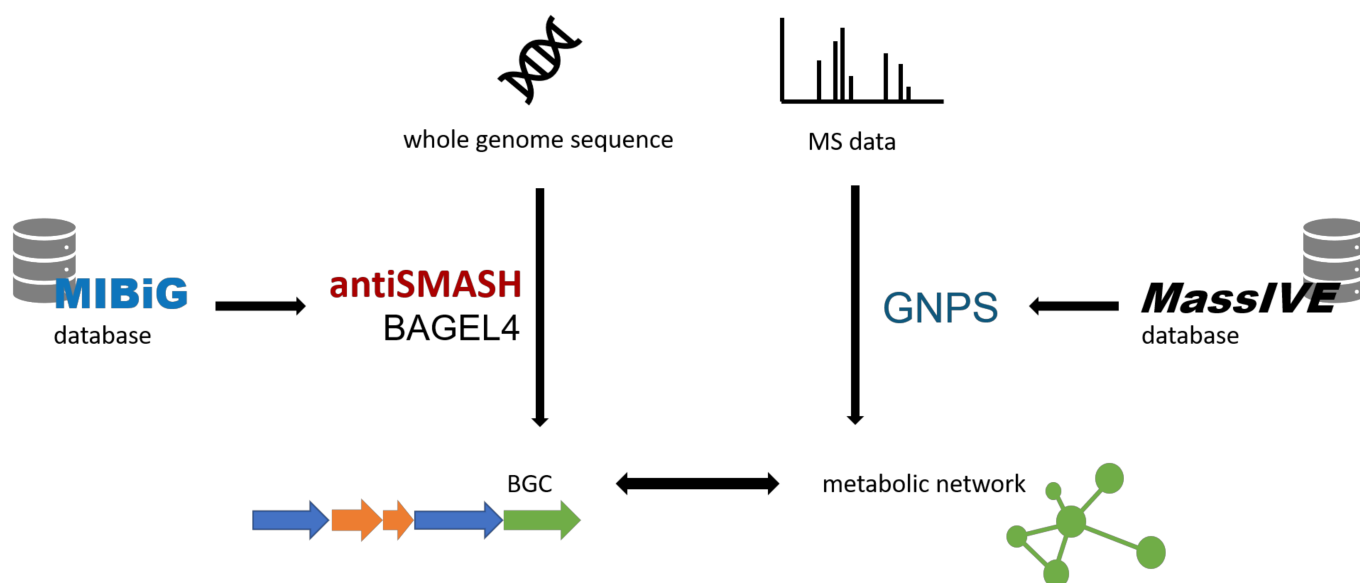


Figure 1. Overview over the bioinformatic tools involved in genomic and metabolomic screening, On the one hand, antiSMASH [10,11] and BAGEL4 [12,13] process the genome sequence to identify BGCs resorting to the MIBiG database [14]. On the other hand, MS data are converted into a molecular network by GNPS [15], which accesses deposited MS data from the MassIVE database [16].

1.1. Bioinformatic Tools for Genome Mining

Several algorithms exist to screen whole genomes for BGCs. The most frequently used Webtool is the Antibiotics and Secondary Metabolites Analysis Shell (antiSMASH) [10,11] which was developed to integrate all existing methods for the identification of BGCs in one user friendly platform. The distinct regions identified in the genome are aligned to their nearest relatives deposited in the antiSMASH database. A forecast about the number of BGCs in the whole genome, the BGC's borders and the presumed class of secondary metabolite is given. Also, the similarity to identified clusters is displayed [17].

Similar to antiSMASH, the Bacteriocin Genome mining tool (BAGEL) serves to identify BGCs but with focus on Ribosomally synthesized and Post translationally modified Peptides (RiPPs) and bacteriocins. The whole genome in question is screened via comparison against a core peptide database [12,13].

The applicability of cluster finding tools like antiSMASH and BAGEL depends on the quality and quantity of the available genome data. These tools' databases revert to the Minimum Information about a Biosynthetic Gene Cluster (MIBiG) database, which is a manually curated repository for genome data. For standardization an own data standard has been developed [14].

1.2. Bioinformatic Tools for Metabolomic Analysis

The analysis of genome sequencing data demonstrates that on average about 20 secondary metabolites are encoded in the genome of actinomycetes [8], while only about 10% of these are actually expressed [18]. But the genome by itself does not give information about the metabolites that are actually produced by a strain. To investigate the metabolome, tandem mass spectroscopy has been established as the method of choice [19]. By comparison of the metabolites' spectra to spectral libraries, known compounds can be identified early in the process of dereplication and thus the tedious process of isolation and structure elucidation is avoided. The detected chemical entities can be visualized in a so-called molecular network. A molecular network displays the relations or similarities, respectively, between the spectra of related/similar molecules. The spectra are depicted as so called 'nodes' or 'parent ions', which are connected by 'edges', to visualize alignments between nodes. Prerequisite for this method is a common platform on which MS/MS-data

of preferably all known substances are stored and can be accessed for comparison with one's own data set [15].

The platform Global Natural Products Social Molecular Networking (GNPS) enables storage, analysis and sharing of MS/MS-data for molecular networking. Via GNPS-clustering MS/MS-data can be matched with the GNPS-database, which is built on user-data [15]. The Mass spectrometry Interactive Virtual Environment (MassIVE) database is the public repository of GNPS for MS/MS-data [16].

Other bioinformatic tools rely on the data processed by GNPS. DEREPLICATOR started as an identifier for peptidic natural products but was soon extended for other classes of natural products. Similar to GNPS, the aim of this tool is to identify known natural products before traversing an immense dereplication effort via compound isolation and structure elucidation. For this purpose, DEREPLICATOR accesses the large spectral library of GNPS [20]. DEREPLICATOR however requires at least one unmodified peptide within a node of a family of molecules as only spectral networks can be identified. To overcome this drawback, VarQuest was developed. VarQuest calculates the masses of mutated peptide variants from a peptidic database and matches these variants against recorded spectra to identify the best fit. The VarQuest algorithm's major advantage is to efficiently narrowing down the number of candidate peptidic variants, which accelerates the calculation immensely [21].

The approaches to study the genome as well as the metabolome have been successfully joined in an integrated genomic and metabolomic screening, also called metabologenomics [22] or comparative genomics and metabolomics [23]. With this method, the detected parent ions in a molecular network are linked to the corresponding biosynthetic pathway, derived from the BGC [24]. That offers the possibility to derive details about structural features, which cannot be derived from the mass-spectrum exclusively, like stereochemistry [25], or to identify targets for genetic engineering [26].

This review aims to highlight studies, in which integrated genomic and metabolomic screening was performed on actinomycetes, and to elaborate their various possible applications. Also, an overview over newly discovered molecules from actinomycetes detected by this method is provided. Finally, the review will demonstrate the usefulness of this method for the above-mentioned applications but rather not for the discovery of new chemical bioactive scaffolds.

2. Applications of Genomic and Metabolomic Screening

2.1. Discovery of New Analogs

Integrated genomic and metabolomic screening is a powerful tool for the detection of new derivatives of known compounds.

Liu et al. [27] investigated the genome and metabolome of the known daptomycin-producer *Streptomyces roseosporus* in search for nonribosomal peptide (NRP) antibiotics. They searched for specific peptidic signatures, that is masses of amino acid fragments, in the molecular network of the strain's metabolome. This way, they discovered subnodes of the daptomycin cluster, which correspond to daptomycin variants missing the N-terminal lipid chain and tryptophan. Additionally, sodium- and potassium- adducts of arylomycin [27], a lipohexapeptide with antibiotic activity [28], were discovered. The authors especially emphasized the existence of several arylomycin intermediates, which lack the typical biaryl linkage and a tryptophan residue at the C-terminus. In an antiSMASH cluster search for Non-ribosomal perptide synthase (NRPS) BGCs, those coding for the productive pathway of daptomycin and arylomycin could be assigned to the produced antibiotics. The production of another known natural product [27], the *Pseudomonas*-active peptidynucleoside napsamycin (Table 1) [29], was suggested by genome analysis and confirmed by the molecular network. Higher molecular weight variants seemed to be produced by *S. roseosporus* as well. Before it was unknown that *S. roseosporus* was able to produce napsamycins at all. The focus of this study lay on stenothricin, a barely characterized antibiotic that was discovered in 1974. So far, four analogs of stenothricin were known [30] and could be detected in

the metabolome of *S. roseosporus*. A corresponding BGC, which matches the metabolomic information, could be identified in an antiSMASH analysis. Additionally, the existence of analogs of stenothricin were indicated by the molecular network (Table 1). These analogs differ in the length of the lipid side chains and amino acid substitution. Also, hydrolysed and glycosylated products could be identified. Liu et al. verified activity of the extracted stenothricin variant mixture against Gram-negative and Gram-positive bacteria [27].

Table 1. List of newly discovered compounds from actinomycetes via combined genomic and metabolomic screening.

Compound	Strain	Activity	Reference
napsamycin analogs stenothricin analogs	<i>Streptomyces roseosporus</i>	antibiotic	[27]
spiroindimicins E and F lagunapyrones D and E	<i>Streptomyces</i> sp. MP131-18	none	[31]
strecacansamycin A, B, C	<i>Streptomyces cacaoi</i> subsp. <i>asoensis</i>	antiproliferative	[25]
valinomycin derivatives	<i>Streptomyces</i> sp. CBMAI 2042	antibiotic	[26]
cyclomarin analogs arenicolide analogs retimycin A	<i>Salinispora</i> sp.	suggested antiproliferative	[24]

This study demonstrates that it is worth investigating old strains with new methods, such as integrated genomic and metabolomic screening. As shown here, such an approach can result in an extended production profile of the known strain as well as point out bioactivities which have not been observed before.

In a study of Duncan et al. [24], 30 isolates of *Salinispora* strains were examined via combined genomic and metabolomic screening, also in order to detect new analogs. The species of the strains were identified as *Salinispora arenicola*, *Salinispora tropica* and *Salinispora pacifica*. Like *Streptomyces*, *Salinispora* belongs to the order of actinomycetes [32]. For several identified BGCs in the genome of the strains, no corresponding compound could be associated indicating that these clusters are not expressed under the cultured conditions or that the product is not detectable with the applied method. The raw extracts of the cultures were submitted to HR-MS/MS and a molecular network was calculated from the resulting ions. To identify the microbial products the MS-spectra were compared to a library of standard *Salinispora* biosynthetic products [24]. Besides the known compounds cyclomarin A [33] and D [34] (Table 1), the existence of a methylated, a demethylated and a hydrated analog is implied by MS-data. Also, a hydroxylated, dehydrogenated and a methylated version [24] of arenicolide A [35] could be detected (Table 1). Eight of the detected compounds could be linked to *Salinispora*-BGCs, among them the so far unidentified BGC of arenicolide A. The BGC could be associated with the production of arenicolide A as, on the one hand, this BGC was present exclusively in arenicolide producing strains and, on the other hand, the annotated ketosynthase-domains correspond to the sequence of the enzyme required for arenicolide production. Another BGC could be linked bioinformatically to a compound newly identified in this study, named retimycin A, an analog of the quinomycins [24], which display antiproliferative activity [36] (Table 1). Here, also the singleness of the parent ion as well as of the BGC in one single strain indicated the match. Additionally, similarity to a BGC of a known quinomycin-like compound was observed with variations in an adenylation-domain and a domain coding for an oxygenase. The MS/MS analysis indicated the presence of an oxidized, methylated thioacetal-version of retimycin A. This functional group has not been observed before for quinomycins [24]. So far, no bioactivity testing has been performed on retimycin A.

One advantage of genomic and metabolomic screening over classical bioactivity-based screening is the possibility to detect inactive analogs. Paulus et al. [31] investigated the biosynthetic potential of strain MP131-18, which was sampled at a Norwegian fjord. 16S phylogeny showed highest similarity to *Streptomyces specialis* and *Streptomyces avicenniae*.

Culture extracts displayed activity against *Bacillus subtilis* and *Pseudomonas putida*. Analysis with antiSMASH revealed 36 gene clusters. To identify the corresponding metabolites HR-LC-QTOF MS was performed followed by dereplication with the Dictionary of Natural Products (DNP) database. This way, the production of the bisindole pyrrole antibiotics lynamycins A to G [37] and spiroindimicins B and C [38] could be confirmed. Known biosynthetic genes from the production of bisindole pyrroles were identified in BGC 36, so the cluster could be associated with the production of these compounds. Additionally, analogs of the known secondary metabolites (named lynamycin H, spiroindimicin E and spiroindimicin F; Table 1) could be identified in the culture extract [31]. Besides these, the polyketides lagunapyrones A-C [39] accumulated. Genes coding for type I and type III Polyketidesynthase (PKS) were present in BGC 3. This indicated that this cluster is responsible for lagunapyrone-production, as lagunapyrone is a polyketide. As for this type III PKS high flexibility in the choice of the acyl-CoA unit was predicted. In the following, the production of two more lagunapyrones D and E (Table 1) with C2 and C5 acyl-chains, of which masses could be found in the culture extract, was indeed confirmed [31]. The new compounds did not display antimicrobial activity with the exception of spiroindimicin B, which showed moderate activity against *Bacillus subtilis* [31]. Due to their lack of bioactivity, the analogs would not have been detected by classical activity-based screening of HPLC-fractions.

All three studies presented in this section demonstrate that integrated genomic and metabolomic screening represents a useful tool for the discovery of unknown analogs of known compound classes and how the linkage between BGC and parent ion can be established.

2.2. Exploring the Productive Spectrum

Integrated genomic and metabolomic screening can be used to explore the productive capability of strains and prioritize those strains, which display the highest amount of uncharacterized genomic and metabolomic entities.

Ishaque et al. [40] investigated the crude extract of a novel *Streptomyces* isolate named *Streptomyces tendae* VITAKN. The culture extract showed quorum sensing inhibition (QSI), thus the group aimed to identify the compound responsible for this effect. Hence, the group performed a whole genome analysis with antiSMASH, in which 33 BGCs could be detected. Only nine of these clusters showed more than 75% similarity to those deposited in the antiSMASH database. The remaining clusters were suspected to code for the production of so far unknown chemical entities. The crude extract was examined via LC-HRMS and LC-HRMS/MS. This resulted in a molecular network consisting of 327 nodes of which four correlated to the spectra of cyclic dipeptides (2,5-diektopiperazines) [40]. Cyclic dipeptides act as LuxR-type activators or inhibitors and exhibit antiproliferative, antibiotic and anti-inflammatory activity [41]. The genes coding for the key enzymes for the formation of 2,5-diektopiperazines, CDPS [42], were identified in the genome. A comparison of the spectra with data from the GNPS-MassIVE database did not result in exact matches. This indicates that the wanted compound has not been characterized before [40].

Nevertheless, the compound with QSI activity could not be identified via this combined genome and metabolomic screening approach. But as no corresponding MS-data for the detected parent ions could be found, an uncharacterized compound with QSI-activity is likely produced by *S. tendae* VITAKN. The existence of unexplored chemical entities also predestines the metabolome of *S. tendae* VITAKN as an object of further investigation [40].

A detailed screening of both, genome and metabolome, can not only help to estimate the amount of unknown natural products but also to elucidate the full biosynthetic potential of putative producer strains. *Streptomyces clavuligerus*, *Streptomyces jumonjinensis*, and *Streptomyces katsurahamanus* [23] are all known to produce the β -lactamase inhibitor clavulanic acid [43]. The following study by AbuSara et al. [23] aimed to examine if the three species produce other secondary metabolites in common. Therefore, a comparative analysis of the metabolome as well as of the genome was performed. Via LC-MS and LC-

MS/MS analysis, it was observed that all three species produce desferrioxamines [44] and ectoine [45], which are very common in streptomycetes as these metabolites are required for general cell functions. The antibiotics holomycin [46] and thiolutin [47] are exclusively produced by *S. clavuligerus* [23]. In this study, the production of the antiproliferative nucleoside pentostatin [48] could be reported for the first time in *S. clavuligerus*, though the corresponding BGC had already been discovered in this strain. However, no production or BGC of pentostatin could be detected in *S. jumonjinensis* or *S. katsurahamanus* [23]. Naringenin is a flavonoid previously only known from plants [49] but was found here to be produced in all three *Streptomyces* species. The same is true for the plant-associated monoterpenes carveol and cuminyl alcohol. The terpene hydroxyvaleric acid could be found in metabolome of *S. clavuligerus* exclusively. To elucidate the corresponding BGCs, an analysis with antiSMASH was performed. This way, 49 BGCs could be detected in the genomes of *S. jumonjinensis* and *S. katsurahamanus*, of which 44 could be associated with known clusters. Terpene like BGCs were observed in all three species, which could be the corresponding BGCs to the above-mentioned plant-derived metabolites. Besides clavulanic acid, *S. clavuligerus* is also able to produce its analog 5S clavam [23], which displays no inhibition of β -lactamases due to the inversed stereochemistry, and another unknown paralog of clavulanic acid [50]. While *S. jumonjinensis* and *S. katsurahamanus* are producers of clavulanic acid as well, no production of 5S clavam or the paralog could be detected. This is reflected in the genomes of the producer strains, which lack the according BGCs [23]. All three species are capable of producing cephalosporin C [51], which is linked to the production of clavulanic acid in *S. clavuligerus* [52]. The corresponding BGCs of *S. jumonjinensis* and *S. katsurahamanus* lack one gene, *blp*, of the cephalosporin-BGC in contrast to *S. clavuligerus*, which indicates that this gene is not essential for the production of cephalosporin C [23].

Though the group performed both, metabolomic and genomic analysis, still little connections could be drawn between the detected metabolites and the BGCs. Many nodes of the metabolomic clusters are unrelated to known BGCs, which reveals potential for further investigation [23]. However, the two methods complemented each other to create a more complete picture of the productive capability of the underexplored strains *S. jumonjinensis* and *S. katsurahamanus*.

Handayani et al. [53] aimed to explore the biosynthetic capabilities of *Streptomyces* isolates from rarely explored habitats in Indonesian terrestrial and marine environments. Firstly, 422 isolates were prioritized according to their bioactivity against a board of test microbes. Nine isolates were selected for integrated genomic and metabolomic screening, using antiSMASH and GNPS, respectively. Alongside metabolites required for basic cell functions (like desferrioxamines or spore pigments) several known compounds (ferrioxamine D1, naphthyridinomycin A, amicetin, antimycin A2, echinoserine, echinomycin, tirandamycin A, staurosporine) could be detected and associated with corresponding BGCs. The ion cluster of ECO-0501 could be detected but no corresponding BGC could be identified. In addition to these characterized compounds, four ion clusters were observed, to which no known compound could be associated. One of these clusters was probably attributed to peptides. The authors compared the BGCs of the three isolates, in which's metabolome the peptidic ion clusters were observed, and identified two BGCs uniquely present in these three isolates. These BGCs presumably code for peptidic compounds, as they were identified as a bacteriocin and a NRPS/ectoine/butyrolactone/other/T1PKS gene cluster. Two more unidentifiable ion clusters were detected, for which a number of BGCs could code so that matching was not yet possible. The detected substances' structures have not been elucidated yet [53].

By assigning the identifiable ion clusters and BGCs, the authors could prioritize strains and within the strains the ion clusters and BGCs, which probably code for new substances, for further research. The investigated strains displayed some unknown ion clusters and even more unknown BGCs, which could not be associated with the production of any known compounds. With these findings, the authors proved their point that underexplored

habitats harbor untapped biosynthetic potential and demonstrated that an integrated genomic and metabolomic screening approach eases dereplication and prioritization.

Integrated genomic and metabolomic screening can also be applied the other way around, not to detect a new compound in a known strain but to find novel producers of known metabolites. In a study of Gosse et al. [54], the microbial biodiversity of the Iron Curtain Cave, Canada should be explored. Among the isolates, a strain displaying antimicrobial activity against multi-resistant *Escherichia coli* and non-resistant *Staphylococcus aureus* was selected for whole genome sequencing, followed by comparative genome analysis for taxonomic classification, which revealed high sequence identity to *Streptomyces lavendulae*. The investigated strain was accordingly named '*Streptomyces*' sp. ICC1. For general cluster investigation, an antiSMASH analysis was performed. Additionally, the BAGEL4 algorithm was deployed. All in all, 37 gene clusters could be identified, the majority of which represent terpene clusters. The metabolome of ICC1 was investigated via UHPLC-ESI-HRMS followed by GNPS clustering. Nodes for several dipeptides were detected but no corresponding BGCs (NRPS or a cyclodipeptide synthase containing cluster) could be identified. On the other hand, a BGC for the siderophore coelichelin could be identified with 100% similarity via antiSMASH analysis, though the metabolite was not present in the culture extract. But then, siderophores are mainly produced under iron limiting conditions, while culture conditions for strains from an iron ore cave are maintained rich in iron. Additionally, some antibacterial compounds could be identified in the culture extract: 2', 5'-dimethoxyflavone and nordentatin were detected by GNPS and a corresponding type III PKS cluster coding for both compounds could be identified via antiSMASH analysis [54]. Also, the antitumor [55] terpene diazepinomicin, which was only known to be produced by *Micromonospora* sp. DPJ12 so far, could be identified in the culture extract. The core genes for the key enzymes for the production of this compound, DkpF, DkpD, and DkpE, which form one of the two main building blocks of diazepinomicin [56], are present in one of the terpene clusters detected with antiSMASH [54].

Though no new antibacterial compounds could be detected, a new producer of the known antitumor agent diazepinomicin was identified. Additionally, the productive capabilities of the investigated strain indicate a high metabolomic potential, which should be explored further by established methods (fractioning culture extracts and prioritization according to their bioactivity). The study also demonstrates, that it is worth investigating both, genome and metabolome, on the one hand, to detect the potential for the production of secondary metabolites, even if they are not actually produced, and on the other hand, to detect metabolites even if the cluster has not been characterized so far.

2.3. Elucidation of Biosynthesis

Not only can new compounds be detected via combined genomic and metabolomic screening, it is also possible to identify their biosynthetic origin. For a study by Paulo et al. [57], *Streptomyces* sp. CBMAI 2042 was isolated from *Citrus sinensis* branches. This strain inhibited growth of *Bacillus megaterium*, *Staphylococcus aureus* and *Candida albicans* and suppressed the proliferation of *Citrus* pathogens. Whole genome sequencing followed by an analysis with antiSMASH revealed 35 BGCs, among them the NRPS-cluster for the depsipeptide valinomycin. Valinomycin has proven antibacterial [58], antiviral [59] and antiproliferative [60] properties. Via UHPLC-QTOF-MS/MS the metabolite itself and its ammonium adduct were detected and the identities verified via matching with the GNPS database. Additionally, the bioinformatic tools DEREPLICATOR and VarQuest, which are specialized in the detection of peptide natural products, were used. Both, valinomycin and its ammonium adduct, appeared as distinct nodes in the molecular network. Besides valinomycin, *Streptomyces* sp. CBMAI 2042 was able to produce the cyclic depsipeptide montanastatin [57], which also displays antiproliferative activity [58]. Though montanastatin has been known before, the biosynthetic origin has never been elucidated. But according to its structure, montanastatin could stem from the same cluster as valinomycin. Thus, the authors expressed the cluster for valinomycin in *S. coelicolor* as

heterologous host and performed a metabolomic analysis. Montanastatin was detected as well as valinomycin and five so far unknown valinomycin-analogs (Table 1). This demonstrates, that all mentioned metabolites originate from the same NRPS-cluster.

Here, metabolomic and genomic approaches were combined in order to elucidate the common biosynthetic origin of valinomycin and montanastatin. Nevertheless, heterologous expression for confirmation remained indispensable.

Liu et al. [25] also aimed to elucidate biosynthetic details by using combinatorial genetic and metabolomic screening to determine the absolute configuration of the stereochemical centers of the newly isolated streccacansamycins A, B and C (Table 1), which are produced by *Streptomyces cacaoi* subsp. *asoensis* H2S5. Streccacansamycins belong to the class of aliphatically bridged aromatic ansamycins [61]. In activity tests in vivo against PC-3, HepG2, and U87-MG cells, respectively, the isolated analogs displayed antiproliferative properties [25]. LC-HR-ESI-MS data of the culture extract were evaluated with GNPS. This way, nodes for ansamycin-analogs were detected. MS data, however, give no information about the absolute stereochemistry of a molecule. So, Liu et al. used the genetic information to reconstruct the production pipeline and derive which stereochemistry would be provided by the modules. For this purpose, the whole genome was sequenced and analyzed with antiSMASH. The analysis revealed 31 BGCs. A type I PKS-NRPS hybrid cluster is probably responsible for the production of streccacansamycins. PKSs and NRPSs are composed of several biosynthetic units called modules, which contain a set of catalytically active domains. The type I PKS-NRPS hybrid cluster contains acyltransferase-domains in module 4, which are stereoselective for *S*-methylmalonyl-CoA, but stereoinversion occurs in the subsequent condensation reaction catalyzed by a ketosynthase-domain so the final configuration at C-12 is *R*. The configuration of the methoxy or hydroxyl-groups at C-3, C-11 and C-13 could be determined as *R*, *S* and *R*, based on the direction of the hydride-addition at the ketoreductase-domains. One ketoreductase-domain type is also responsible for the formation of *cis* or *trans* double bonds depending on the direction of the reduced hydroxy-group. This way, it could be deduced that the double bonds at C-5, C-7, and C-9 exhibit *trans*- and those at C-15 *cis*-configuration [25].

The study demonstrates the usefulness of integrated genomic and metabolomic screening for structure elucidation beyond the scope of mass spectrometry, especially concerning stereochemistry. This is of special relevance considering the stereospecificity of most mechanisms of action.

3. Conclusions

The listed examples demonstrate that integrated genomic and metabolomic screening has already been successfully applied for various purposes: The strategy has proven useful for comparative genomic and metabolomic studies as to explore the full biosynthetic potential of known or newly isolated *Actinomyces* strains and to detect similarities in the production profile of distinct strains [23,40,53]. This way, new producers of known compounds could be found [24,54]. Based on this overview of their productive capabilities, strains could be prioritized according to the presence of unidentifiable clusters [40]. Also, bioactive as well as inactive analogs of known compounds have been found [27,31]. Therefore, integrated genomic and metabolomic screening also facilitates the elucidation of biosynthesis [25,57].

Nevertheless, integrated genomic and metabolomic screening should not be overestimated, for it will not enable the discovery of new scaffolds, but just support this objective by easing the process of dereplication. No completely new lead structures could be discovered with the method so far, due to the fact that the compound and BGC identification is based on a pattern-based search, which relies on acquired data about known compounds and BGCs. For the identification of new compound-classes laborious bioactivity-guided screening, genetic engineering, and fractionation of culture extracts, as well as heterologous expression will still be required.

In conclusion, integrated genomic and metabolomic screening is a useful tool for simplified dereplication and as such will help to accelerate drug discovery through identification of new strains and analogs of known bioactive compounds.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Acknowledgments: Thanks to Irene Reimche and Franziska Handel for proofreading.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Takahashi, Y.; Nakashima, T. Actinomycetes, an inexhaustible source of naturally occurring antibiotics. *Antibiotics* **2018**, *7*, 45. [[CrossRef](#)] [[PubMed](#)]
2. Bérdy, J. Thoughts and facts about antibiotics: Where we are now and where we are heading. *J. Antibiot.* **2012**, *65*, 385–395. [[CrossRef](#)] [[PubMed](#)]
3. Wibowo, J.T.; Kellermann, M.Y.; Köck, M.; Putra, M.Y.; Murniasih, T.; Mohr, K.I.; Wink, J.; Praditya, D.F.; Steinmann, E.; Schupp, P.J. Anti-infective and antiviral activity of valinomycin and its analogues from a sea cucumber-associated bacterium, *Streptomyces* Sp. SV 21. *Mar. Drugs* **2021**, *19*, 81. [[CrossRef](#)] [[PubMed](#)]
4. Bèhal, V. Bioactive products from *Streptomyces*. In *Advances in Applied Microbiology*; Academic Press Inc.: Cambridge, MA, USA, 2000; pp. 113–156. [[CrossRef](#)]
5. Hutchings, M.; Truman, A.; Wilkinson, B. Antibiotics: Past, present and future. In *Current Opinion in Microbiology*; Elsevier Ltd.: Amsterdam, The Netherlands, 2019; pp. 72–80. [[CrossRef](#)]
6. Katz, L.; Baltz, R.H. Natural product discovery: Past, present, and future. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 155–176. [[CrossRef](#)] [[PubMed](#)]
7. Baltz, R.H. Natural product drug discovery in the genomic era: Realities, conjectures, misconceptions, and opportunities. *J. Ind. Microbiol. Biotechnol.* **2018**, *46*, 1–19. [[CrossRef](#)] [[PubMed](#)]
8. Traxler, M.F.; Kolter, R. Natural products in soil microbe interactions and evolution. *Nat. Prod. Rep.* **2015**, *32*, 956–970. [[CrossRef](#)]
9. Kenshole, E.; Herisse, M.; Michael, M.; Pidot, S.J. Natural product discovery through microbial genome mining. *Curr. Opin. Chem. Biol.* **2021**, *60*, 47–54. [[CrossRef](#)]
10. Niu, G. Genomics-driven natural product discovery in actinomycetes. *Trends Biotechnol.* **2018**, *36*, 238–241. [[CrossRef](#)]
11. 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](#)]
12. Van Heel, A.J.; de Jong, A.; Song, C.; Viel, J.H.; Kok, J.; Kuipers, O.P. BAGEL4: A user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res.* **2018**, *46*, W278–W281. [[CrossRef](#)]
13. Van Heel, A.J.; De Jong, A.; Montalbà N-Ló Pez, M.; Kok, J.; Kuipers, O.P. BAGEL3: Automated identification of genes encoding bacteriocins and (non-)bacteriocidal posttranslationally modified peptides. *Nucleic Acids Res.* **2013**, *41*, W448–W453. [[CrossRef](#)] [[PubMed](#)]
14. Kautsar, S.A.; Blin, K.; Shaw, S.; Navarro-Muñoz, J.C.; Terlouw, B.R.; Van Der Hooft, J.J.J.; Van Santen, J.A.; Tracanna, V.; Suarez Duran, H.G.; Pascal Andreu, V.; et al. MIBiG 2.0: A repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* **2020**, *48*, D454–D458. [[CrossRef](#)]
15. 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](#)] [[PubMed](#)]
16. 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](#)]
17. Blin, K.; Wolf, T.; Chevrette, M.G.; Lu, X.; Schwalen, C.J.; Kautsar, S.A.; Suarez Duran, H.G.; de Los Santos, E.L.C.; Kim, H.U.; Nave, M.; et al. AntiSMASH 4.0—Improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, *45*, W36–W41. [[CrossRef](#)]
18. Walsh, C.; Tang, Y. Natural product biosynthesis: Chemical logic and enzymatic machinery. *Chem. Rev.* **2017**, *117*, 5226–5333.
19. Van Agthoven, M.A.; Lam, Y.P.Y.; O'Connor, P.B.; Rolando, C.; Delsuc, M.A. Two-dimensional mass spectrometry: New perspectives for tandem mass spectrometry. *Eur. Biophys. J.* **2019**, *48*, 213–229. [[CrossRef](#)]
20. 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*, 1–12. [[CrossRef](#)] [[PubMed](#)]

21. Gurevich, A.; Mikheenko, A.; Shlemov, A.; Korobeynikov, A.; Mohimani, H.; Pevzner, P.A. Increased diversity of peptidic natural products revealed by modification-tolerant database search of mass spectra. *Nat. Microbiol.* **2018**, *3*, 319–327. [[CrossRef](#)] [[PubMed](#)]
22. Van Der Hooft, J.J.J.; Mohimani, H.; Bauermeister, A.; Dorrestein, P.C.; Duncan, K.R.; Medema, M.H. Linking genomics and metabolomics to chart specialized metabolic diversity. *Chem. Soc. Rev.* **2020**, *49*, 3297–3314. [[CrossRef](#)]
23. AbuSara, N.F.; Piercey, B.M.; Moore, M.A.; Shaikh, A.A.; Nothias, L.F.; Srivastava, S.K.; Cruz-Morales, P.; Dorrestein, P.C.; Barona-Gómez, F.; Tahlan, K. Comparative genomics and metabolomics analyses of clavulanic acid-producing *Streptomyces* species provides insight into specialized metabolism. *Front. Microbiol.* **2019**, *10*, 2550. [[CrossRef](#)]
24. Duncan, K.R.; Crüsemann, M.; Lechner, A.; Sarkar, A.; Li, J.; Ziemert, N.; Wang, M.; Bandeira, N.; Moore, B.S.; Dorrestein, P.C.; et al. Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from *Salinispora* species. *Chem. Biol.* **2015**, *22*, 460–471. [[CrossRef](#)]
25. Liu, L.-L.; Chen, Z.-F.; Liu, Y.; Tang, D.; Gao, H.-H.; Zhang, Q.; Gao, J.-M. Molecular networking-based for the target discovery of potent antiproliferative polycyclic macrolactam ansamycins from *Streptomyces Cacaoi* subsp. *asoensis*. *Org. Chem. Front.* **2020**, *7*, 4008–4018. [[CrossRef](#)]
26. Sigrist, R.; Paulo, B.S.; Angolini, C.F.F.; De Oliveira, L.G. Mass spectrometry-guided genome mining as a tool to uncover novel natural products. *J. Vis. Exp.* **2020**, *2020*, e60825. [[CrossRef](#)]
27. Liu, W.T.; Lamsa, A.; Wong, W.R.; Boudreau, P.D.; Kersten, R.; Peng, Y.; Moree, W.J.; Duggan, B.M.; Moore, B.S.; Gerwick, W.H.; et al. MS/MS-based networking and peptidogenomics guided genome mining revealed the stenothricin gene cluster in *Streptomyces roseosporus*. *J. Antibiot.* **2014**, *67*, 99–104. [[CrossRef](#)]
28. Schimana, J.; Gebhardt, K.; Hölzel, A.; Schmid, D.G.; Süßmuth, R.; Müller, J.; Pukall, R.; Fiedler, H.-P. Arylomycins A and B, new biaryl-bridged lipopeptide antibiotics produced by *Streptomyces* Sp. tue 6075. I. Taxonomy, fermentation, isolation and biological activities. *J. Antibiot.* **2002**, *55*, 565–570. [[CrossRef](#)]
29. Chatterjee, S.; Nadkarni, S.R.; Vijayakumar, E.K.S.; Patel, M.V.; Gangul, B.N.; Fehlhauer, H.-W.; Vertesy, L. Napsamycins, new *Pseudomonas* active antibiotics of the mureidomycin family from *Streptomyces* sp. HIL Y-82, 11372. *J. Antibiot.* **1994**, *47*, 595–598. [[CrossRef](#)]
30. Rincken, M.; Lehmann, W.D.; König, W.A. Die Struktur von stenothricin—Korrektur eines früheren strukturvorschlags. *Liebigs Ann. Chem.* **1984**, *1984*, 1672–1684. [[CrossRef](#)]
31. Paulus, C.; Rebets, Y.; Tokovenko, B.; Nadmid, S.; Terekhova, L.P.; Myronovskiy, M.; Zotchev, S.B.; Rückert, C.; Braig, S.; Zahler, S.; et al. New natural products identified by combined genomics-metabolomics profiling of marine *Streptomyces* Sp. MP131-18. *Sci. Rep.* **2017**, *7*, 1–11. [[CrossRef](#)]
32. Jensen, P.R.; Moore, B.S.; Fenical, W. The marine actinomycete genus *Salinispora*: A model organism for secondary metabolite discovery. *Nat. Prod. Rep.* **2015**, *32*, 738–751. [[CrossRef](#)]
33. Renner, M.K.; Shen, Y.C.; Cheng, X.C.; Jensen, P.R.; Frankmoelle, W.; Kauffman, C.A.; Fenical, W.; Lobkovsky, E.; Clardy, J. Cyclomarins A-C, new antiinflammatory cyclic peptides produced by a marine bacterium (*Streptomyces* Sp.). *J. Am. Chem. Soc.* **1999**, *121*, 11273–11276. [[CrossRef](#)]
34. Gontang, E.A.; Fenical, W.; Jensen, P.R. Phylogenetic diversity of gram-negative positive bacteria cultured from marine sediments. *Appl. Environ. Microbiol.* **2007**, *73*, 3272–3282. [[CrossRef](#)]
35. Williams, P.G.; Miller, E.D.; Asolkar, R.N.; Jensen, P.R.; Fenical, W. Arenicolides A-C, 26-Membered ring macrolides from the marine actinomycete *Salinispora Arenicola*. *J. Org. Chem.* **2007**, *72*, 5025–5034. [[CrossRef](#)] [[PubMed](#)]
36. Wang, Y.; Liu, Y.; Tang, F.; Bernot, K.M.; Schore, R.; Marcucci, G.; Caligiuri, M.A.; Zheng, P.; Liu, Y. echinomycin protects mice against relapsed acute myeloid leukemia without adverse effect on hematopoietic stem cells. *Blood* **2014**, *124*, 1127–1135. [[CrossRef](#)]
37. McArthur, K.A.; Mitchell, S.S.; Tsueng, G.; Rheingold, A.; White, D.J.; Grodberg, J.; Lam, K.S.; Potts, B.C.M. Lynamycins A-E, chlorinated bisindole pyrrole antibiotics from a novel marine actinomycete. *J. Nat. Prod.* **2008**, *71*, 1732–1737. [[CrossRef](#)]
38. Zhang, W.; Liu, Z.; Li, S.; Yang, T.; Zhang, Q.; Ma, L.; Tian, X.; Zhang, H.; Huang, C.; Zhang, S.; et al. Spiroindimycins A-D: New bisindole alkaloids from a deep-sea-derived actinomycete. *Org. Lett.* **2012**, *14*, 3364–3367. [[CrossRef](#)]
39. Lindel, T.; Jensen, P.R.; Fenical, W. Lagunapyrones A-C: Cytotoxic acetogenins of a new skeletal class from a marine sediment bacterium. *Tetrahedron Lett.* **1996**, *37*, 1327–1330. [[CrossRef](#)]
40. Ishaque, N.M.; Burgsdorf, I.; Limlingan Malit, J.J.; Saha, S.; Teta, R.; Ewe, D.; Kannabiran, K.; Hrouzek, P.; Steindler, L.; Costantino, V.; et al. Isolation, genomic and metabolomic characterization of *Streptomyces Tendae* VITAKN with quorum sensing inhibitory activity from Southern India. *Microorganisms* **2020**, *8*, 121. [[CrossRef](#)]
41. Borthwick, A.D. 2,5-Diketopiperazines: Synthesis, reactions, medicinal chemistry, and bioactive natural products. *Chem. Rev.* **2012**, *112*, 3641–3716. [[CrossRef](#)]
42. Yao, T.; Liu, J.; Liu, Z.; Li, T.; Li, H.; Che, Q.; Zhu, T.; Li, D.; Gu, Q.; Li, W. Genome mining of cyclodipeptide synthases unravels unusual tRNA-dependent diketopiperazine-terpene biosynthetic machinery. *Nat. Commun.* **2018**, *9*, 1–12. [[CrossRef](#)]
43. Reading, C.; Cole, M. Clavulanic acid: A beta lactamase inhibiting beta lactam from *Streptomyces Clavuligerus*. *Antimicrob. Agents Chemother.* **1977**, *11*, 852–857. [[CrossRef](#)] [[PubMed](#)]
44. Keberle, H. The biochemistry of desferrioxamine and its relation to iron metabolism. *Ann. N. Y. Acad. Sci.* **1964**, *119*, 758–768. [[CrossRef](#)]

45. Galinski, E.A.; Pfeiffer, H.-P.; Truper, H.G. 1,4,5,6-tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. *Eur. J. Biochem.* **1985**, *149*, 135–139. [[CrossRef](#)] [[PubMed](#)]
46. Oliva, B.; O'Neill, A.; Wilson, J.M.; O'Hanlon, P.J.; Chopra, I. Antimicrobial properties and mode of action of the pyrrothine holomycin. *Antimicrob. Agents Chemother.* **2001**, *45*, 532–539. [[CrossRef](#)]
47. Lauinger, L.; Li, J.; Shostak, A.; Cemel, I.A.; Ha, N.; Zhang, Y.; Merkl, P.E.; Obermeyer, S.; Stankovic-Valentin, N.; Schafmeier, T.; et al. Thiolutin is a zinc chelator that inhibits the Rpn11 and other JAMM metalloproteases. *Nat. Chem. Biol.* **2017**, *13*, 709–714. [[CrossRef](#)]
48. Dillman, R.O. Pentostatin (Nipent®) in the treatment of chronic lymphocyte leukemia and hairy cell leukemia. *Expert Rev. Anticancer Ther.* **2004**, *4*, 27–36. [[CrossRef](#)]
49. Zeng, W.; Jin, L.; Zhang, F.; Zhang, C.; Liang, W. Naringenin as a potential immunomodulator in therapeutics. *Pharmacol. Res.* **2018**, *135*, 122–126. [[CrossRef](#)]
50. Tahlan, K.; Anders, C.; Wong, A.; Mosher, R.H.; Beatty, P.H.; Brumlik, M.J.; Griffin, A.; Hughes, C.; Griffin, J.; Barton, B.; et al. 5S clavam biosynthetic genes are located in both the clavam and paralog gene clusters in *Streptomyces Clavuligerus*. *Chem. Biol.* **2007**, *14*, 131–142. [[CrossRef](#)] [[PubMed](#)]
51. Newton, G.G.F.; Abraham, E.P. Cephalosporin C, a new antibiotic containing sulphur and D- α -amino adipic acid. *Nature* **1955**, *175*, 548. [[CrossRef](#)]
52. Romero, J.; Liras, P.; Martín, J.F. Dissociation of cephamycin and clavulanic acid biosynthesis in streptomyces clavuligerus. *Appl. Microbiol. Biotechnol.* **1984**, *20*, 318–325. [[CrossRef](#)]
53. Handayani, I.; Saad, H.; Ratnakomala, S.; Lisdiyanti, P.; Kusharyoto, W.; Krause, J.; Kulik, A.; Wohlleben, W.; Aziz, S.; Gross, H.; et al. Mining Indonesian microbial biodiversity for novel natural compounds by a combined genome mining and molecular networking approach. *Mar. Drugs* **2021**, *19*, 316. [[CrossRef](#)] [[PubMed](#)]
54. Gosse, J.T.; Ghosh, S.; Sproule, A.; Overy, D.; Cheeptham, N.; Boddy, C.N. Whole genome sequencing and metabolomic study of cave *Streptomyces* isolates ICC1 and ICC4. *Front. Microbiol.* **2019**, *10*, 1020. [[CrossRef](#)] [[PubMed](#)]
55. Gourdeau, H.; McAlpine, J.B.; Ranger, M.; Simard, B.; Berger, F.; Beaudry, F.; Falardeau, P. Identification, characterization and potent antitumor activity of ECO-4601, a novel peripheral benzodiazepine receptor ligand. *Cancer Chemother. Pharmacol.* **2008**, *61*, 911–921. [[CrossRef](#)] [[PubMed](#)]
56. McAlpine, J.B.; Banskota, A.H.; Charan, R.D.; Schlingmann, G.; Zazopoulos, E.; Pirae, M.; Janso, J.; Bernan, V.S.; Aouidate, M.; Farnet, C.M.; et al. Biosynthesis of diazepamycin/ECO-4601, a *Micromonospora* secondary metabolite with a novel ring system. *J. Nat. Prod.* **2008**, *71*, 1585–1590. [[CrossRef](#)] [[PubMed](#)]
57. Paulo, B.S.; Sigrist, R.; Angolini, C.F.F.; De Oliveira, L.G. New cyclodepsipeptide derivatives revealed by genome mining and molecular networking. *ChemistrySelect* **2019**, *4*, 7785–7790. [[CrossRef](#)]
58. Pettit, G.R.; Tan, R.; Melody, N.; Kielty, J.M.; Pettit, R.K.; Herald, D.L.; Tucker, B.E.; Mallavia, L.P.; Doubek, D.L.; Schmidt, J.M. Antineoplastic agents. Part 409: Isolation and structure of montanastatin from a terrestrial actinomycete. *Bioorganic Med. Chem.* **1999**, *7*, 895–899. [[CrossRef](#)]
59. Wu, C.Y.; Jan, J.T.; Ma, S.H.; Kuo, C.J.; Juan, H.F.; Cheng, Y.S.E.; Hsu, H.H.; Huang, H.C.; Wu, D.; Brik, A.; et al. Small molecules targeting severe acute respiratory syndrome human Coronavirus. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10012–10017. [[CrossRef](#)]
60. Iacobazzi, R.M.; Annese, C.; Azzariti, A.; D'accolti, L.; Franco, M.; Fusco, C.; la Piana, G.; Laquintana, V.; Denora, N. Antitumor potential of conjugable valinomycins bearing hydroxyl sites: In vitro studies. *ACS Med. Chem. Lett.* **2013**, *4*, 1189–1192. [[CrossRef](#)]
61. Rinehart, K.L.; Shield, L.S. Chemistry of the ansamycin antibiotics. *Fortschr. Chem. Org. Nat.* **1976**, *33*, 231–307. [[CrossRef](#)]