

## Article

# Exploring *Micromonospora* as Phocoenamycin Producers

Maria Kokkini \*, Cristina González Heredia, Daniel Oves-Costales, Mercedes de la Cruz , Pilar Sánchez, Jesús Martín, Francisca Vicente , Olga Genilloud  and Fernando Reyes \*

Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, Parque Tecnológico Ciencias de la Salud, Avda. del Conocimiento 34, 18016 Armilla, Granada, Spain

\* Correspondence: maria.kokkini@medinaandalucia.es (M.K.); fernando.reyes@medinaandalucia.es (F.R.); Tel.: +34-958-993-965 (F.R.)

**Abstract:** Over the past few years, new technological and scientific advances have reinforced the field of natural product discovery. The spiro-tetronate class of natural products has recently grown with the discovery of phocoenamycins, natural actinomycete derived compounds that possess different antibiotic activities. Exploring the MEDINA's strain collection, 27 actinomycete strains, including three marine-derived and 24 terrestrial strains, were identified as possible phocoenamycins producers and their taxonomic identification by 16S rDNA sequencing showed that they all belong to the *Micromonospora* genus. Using an OSMAC approach, all the strains were cultivated in 10 different media each, resulting in 270 fermentations, whose extracts were analyzed by LC-HRMS and subjected to High-throughput screening (HTS) against methicillin-resistant *Staphylococcus aureus* (MRSA), *Mycobacterium tuberculosis* H37Ra and *Mycobacterium bovis*. The combination of LC-UV-HRMS analyses, metabolomics analysis and molecular networking (GNPS) revealed that they produce several related spiro-tetronates not disclosed before. Variations in the culture media were identified as the most determining factor for phocoenamycin production and the best producer strains and media were established. Herein, we reported the chemically diverse production and metabolic profiling of *Micromonospora* sp. strains, including the known phocoenamycins and maklamicin, reported for the first time as being related to this family of compounds, as well as the bioactivity of their crude extracts. Although our findings do not confirm previous statements about phocoenamycins production only in unique marine environments, they have identified marine-derived *Micromonospora* species as the best producers of phocoenamycins in terms of both the abundance in their extracts of some major members of the structural class and the variety of molecular structures produced.

**Keywords:** antimicrobial resistance; drug discovery; natural products; actinomycetes; *Micromonospora*; polyketides; spiro-tetronates; phocoenamycins



**Citation:** Kokkini, M.; González Heredia, C.; Oves-Costales, D.; de la Cruz, M.; Sánchez, P.; Martín, J.; Vicente, F.; Genilloud, O.; Reyes, F. Exploring *Micromonospora* as Phocoenamycin Producers. *Mar. Drugs* **2022**, *20*, 769. <https://doi.org/10.3390/md20120769>

Academic Editor: Jim La Clair

Received: 21 October 2022

Accepted: 3 December 2022

Published: 7 December 2022

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



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

## 1. Introduction

Antimicrobial resistance (AMR), the ability of microorganisms to progressively become resistant to known antimicrobials, is an urgent health challenge worldwide and can be considered a hidden global pandemic [1]. Historically, natural products (NPs) have played a major role in the discovery of antimicrobials. They have provided the pharmacophores for most of the major classes of antibiotics, including the  $\beta$ -lactams, aminoglycosides, macrolides, tetracyclines, rifamycins, glycopeptides, streptogramins, and lipopeptides [2]. In fact, according to Newman and Cragg's review analyzing the new drugs approved from January 1981 to September 2019, out of all the new antibiotics (vaccines excluded), 68.5% were natural products or natural product derivatives, while only 27.7% were totally synthetic antibiotics [3].

Natural products (NP) are structurally perfected by evolution to deliver particular biological functions [4] and are therefore an invaluable source of inspiration in antibiotic design and development [5], compared to the lack of success of synthetic molecules, which often miss the physicochemical properties needed to penetrate bacterial membranes [6].

These special characteristics in comparison with the synthetic compounds carry both advantages and challenges in the discovery of new antibiotic leads [4].

The challenges include various shortfalls and technical barriers, such as the inability to culture the majority of environmental microorganisms in the laboratory, the time-consuming dereplication process, and obtaining the optimal conditions to scale up their production and acquire a sufficient amount to isolate the active principles and elucidate their structures [4,7]. Since the 1990s and onwards, these challenges, along with other economic issues have led to the withdrawal of large pharmaceutical companies from antibiotic research and directed them towards other indications. As a result, there is an absence of new classes of antibiotics reaching the market, leaving antibiotic discovery and development to academic and governmental institutions, and small biotech companies [6,8,9].

Over the past few years, the development of technological and scientific advances such as improved analytical tools, genome mining and engineering, and microbial culturing advances, such as co-cultivation and in situ cultivation [10], are opening up new opportunities to explore new microbial sources and make NP drug discovery attractive again [4,7]. Metabolomics was developed as an approach for metabolic profiling and can provide information on the metabolite composition in NP extracts, thus helping to prioritize NP extracts for isolation and accelerate dereplication. Along with the one strain, many compound (OSMAC) approach, metabolomics can be used to explore and statistically validate relationships between culture parameters and chemical diversity in a microbial isolate. In this respect, the Global Natural Products Social (GNPS) molecular networking platform [11] can organize thousands of sets of MS/MS data recorded from a given set of extracts and visualize the relationship of the metabolites as clusters of structurally related molecules, annotating unknown analogues and new NP scaffolds [2,4].

Spirotetronates are a class of polyketides produced by actinomycetes, most of them belonging to the genera of *Streptomyces*, *Micromonospora* and *Actinomadura* [12,13]. The producing microorganisms have been isolated from many different sources, such as soil, plants, sea and marine sediments and organisms, and are geographically distributed worldwide. They have demonstrated a wide variety of biological activities, mostly antibiotic, against gram-positive bacteria and antitumor [12–14]. Despite their potent bioactivities, their structural complexity and the difficulties towards large-scale culturing, have been limiting their availability for further preclinical and clinical studies [3].

Recently, the spirotetronate class has grown with the discovery of the phocoenamycin family of compounds, some of them isolated in our laboratory [15,16]. In this work, a total of 27 actinomycete strains identified as possible phocoenamycin producers using our LC/MS analytical platform [17], were cultured using an OSMAC approach in order to select the most promising ones, as well as fermentation media for the production of members of this structural class. The analysis of the crude microbial extracts revealed that they produce several related compounds not disclosed before, and that their production is dependent on the culture medium. Herein, we reported the diverse metabolic production of these strains, as well as the bioactivity of their crude extracts against methicillin-resistant *Staphylococcus aureus* (MRSA), *Mycobacterium tuberculosis* H37Ra and *Mycobacterium bovis*.

## 2. Results

### 2.1. Taxonomy, Geography and Ecology of the Producing Microorganisms

The microbial strains used in this study belong to the culture collection of MED-INA foundation and were identified by searching an internal repository of LC/MS analyses of microbial extracts against an internal database of LC/UV/MS data of microbial natural products, which includes several members of the phocoenamycin family [17]. This search identified 27 microbial strains that potentially biosynthesize at least one of the phocoenamycins.

Taxonomic identification of the 27 strains was performed via sequencing of the 16S ribosomal gene with the universal primers fD1 and rP2. The sequences obtained were compared with the sequences of all prokaryotic species recognized as “type strains”

with validly published names registered in the EzBioCloud database. The closest species, the (%) percentage of similarity, the source of isolation and the geographic origin of each strain are shown in Table 1. In the case of the strain CA-259211, it was not possible to obtain amplification product of the 16S ribosomal gene in any of the PCR conditions tested.

**Table 1.** Closest species assignment using the EzBioCloud database, the (%) percentage of similarity, the source of isolation and the geographic origin of each strain.

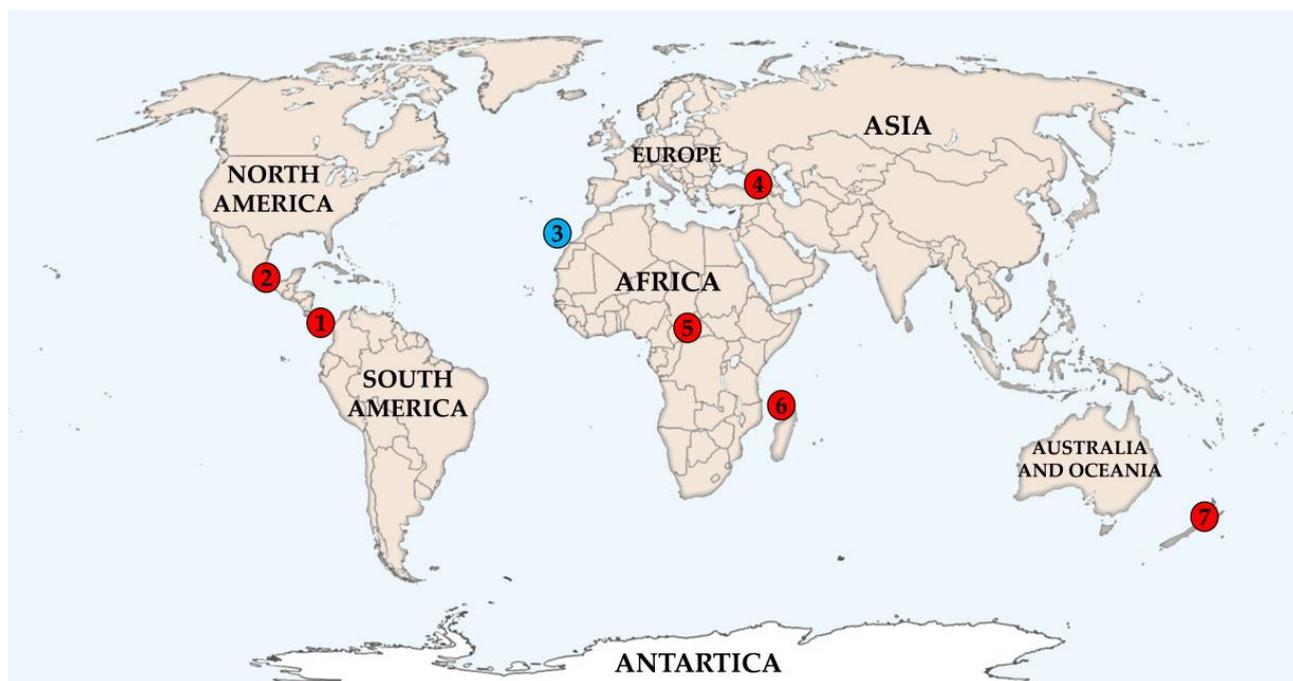
| Strain    | Geographic Origin *          | Ecology                                       | <i>Micromonospora</i> Species | Similarity (%) |
|-----------|------------------------------|---|-------------------------------|----------------|
| CA-107814 | Costa Rica (1)               | soil-rice cultivation                         | <i>M. terminaliae</i>         | 99.48          |
| CA-108000 | Costa Rica (1)               | soil-rice cultivation                         | <i>M. endophytica</i>         | 100            |
| CA-184181 | Mexico (2)                   | soil  | <i>M. siamensis</i>           | 99.32          |
| CA-214658 | Spain (3)                    | marine cave sediment                          | <i>M. endophytica</i>         | 99.63          |
| CA-214671 | Spain (3)                    | marine cave sediment                          | <i>M. chaiyaphumensis</i>     | 99.84          |
| CA-218877 | Spain (3)                    | marine invertebrate <i>Porifera</i> sp.       | <i>M. endophytica</i>         | 100            |
| CA-238377 | Georgia (4.a)                | rhizosphere soil of <i>Pteridium tauricum</i> | <i>M. terminaliae</i>         | 99.54          |
| CA-238397 | Georgia (4.a)                | rhizosphere soil of <i>Pteridium tauricum</i> | <i>M. chalcea</i>             | 100            |
| CA-238398 | Georgia (4.a)                | rhizosphere soil of <i>Pteridium tauricum</i> | <i>M. siamensis</i>           | 99.32          |
| CA-243027 | Georgia (4.a)                | rhizosphere soil of <i>Pteridium tauricum</i> | <i>M. terminaliae</i>         | 100            |
| CA-243168 | Central African Republic (5) | forest organic humid soil                     | <i>M. endophytica</i>         | 100            |
| CA-244160 | Union of the Comoros (6.a)   | forest organic dry soil                       | <i>M. terminaliae</i>         | 99.53          |
| CA-244161 | Union of the Comoros (6.a)   | forest organic dry soil                       | <i>M. terminaliae</i>         | 99.47          |
| CA-244669 | Union of the Comoros (6.b)   | forest organic humid soil                     | <i>M. terminaliae</i>         | 99.41          |
| CA-244673 | Union of the Comoros (6.b)   | forest organic humid soil                     | <i>M. chalcea</i>             | 100            |
| CA-244674 | Union of the Comoros (6.b)   | forest organic humid soil                     | <i>M. endophytica</i>         | 99.31          |
| CA-244675 | Union of the Comoros (6.b)   | forest organic humid soil                     | <i>M. terminaliae</i>         | 99.49          |
| CA-246501 | Georgia (4.a)                | rhizosphere soil of <i>Pteridium tauricum</i> | <i>M. terminaliae</i>         | 99.53          |
| CA-246506 | Georgia (4.a)                | rhizosphere soil of <i>Pteridium tauricum</i> | <i>M. aurantiaca</i>          | 99.64          |
| CA-248285 | New Zealand (7)              | swampy soil                                   | <i>M. soli</i>                | 99.07          |
| CA-248314 | New Zealand (7)              | swampy soil                                   | <i>M. endophytica</i>         | 99.66          |
| CA-248649 | Union of the Comoros (6.c)   | forest organic soil                           | <i>M. terminaliae</i>         | 100            |
| CA-249271 | Union of the Comoros (6.d)   | forest organic soil                           | <i>M. endophytica</i>         | 99.38          |
| CA-249379 | Union of the Comoros (6.a)   | forest organic dry soil                       | <i>M. terminaliae</i>         | 99.54          |
| CA-251294 | Union of the Comoros (6.c)   | forest organic soil                           | <i>M. terminaliae</i>         | 99.51          |
| CA-253038 | Georgia (4.b)                | rhizosphere soil of <i>Populus canescens</i>  | <i>M. siamensis</i>           | 99.29          |
| CA-259211 | Georgia (4.c)                | rhizosphere soil of <i>Ranunculus buhsei</i>  | No data                       | -              |

\* Places and dates of collection indicated in brackets are specified in Figure 1.

All the strains identified as phocoenamycin producers belong to the *Micromonospora* genus, suggesting that the common pathway for the production of these molecules must be highly conserved within species of this genus. The closest type species were distributed into seven different species, with *M. terminaliae* and *M. endophytica* getting assigned more than half of all the strains (18 strains), followed by *M. siamensis*, *M. chalcea*, *M. solis*, *M. chaiyaphumensis* and *M. aurantiaca*. Given that all three known phocoenamycins were originally isolated as a product of *Micromonospora* sp., and that all strains identified in this study as possible producers also belong to this genus, it can be proposed that the biosynthesis of this family of compounds might be confined to the *Micromonospora* species.

Despite the close taxonomic relation of the strains, they displayed great ecological and geographical diversity. Even though the phocoenamycins were originally isolated from marine-derived actinomycetes, the producers identified in this study were isolated from many different sources, as shown in Table 1. These sources include forest organic soil, either dry or humid, rhizosphere soil from various plants, rice cultivations, soil in swamp ecosystems, marine sediments and finally a strain isolated from a marine invertebrate. This demonstrates that the production of phocoenamycins is not attributed to the unique environmental conditions, as previously believed [18,19]. Regarding their geographical origin, they were widely distributed worldwide, covering almost all continents. In particular, the 27 strains were isolated from Costa Rica, Mexico, Spain, Georgia, Central African

Republic, Comoros Islands and New Zealand, confirming the hypothesis that this family of compounds is highly conserved within the *Micromonospora* genus.



**Figure 1.** Geographic distribution of the 27 strains within the seven different countries worldwide. The blue colored circle corresponds to the three marine-derived strains and the red to the rest of the terrestrial strains. Numbers correspond to the different countries, regions, approximate Global Positioning System (GPS) coordinates and year of collection: 1. Costa Rica (Guanacaste, 11°2′12.80″ N 85°27′10.84″ W, 1998) 2. Mexico (Las Tuxtlas, Veracruz, 18°28′28.78″ N, 95°12′54.43″ W, 2001) 3. Spain (Gran Canaria, 27°44′42.04″ N, 15°37′42.28″ W, 2004) 4. Georgia, a. (Batumi, Adjara, 41°36′32.94″ N, 41°39′1.51″ E, 2004) b. (Poti, Samegrelo, 42°10′16.14″ N, 41°40′47.06″ E, 2005) c. (Khulo, Adjara, 41°38′51.94″ N, 42°18′51.74″ E, 2005) 5. Central African Republic (Damara/Sibut, 5°19′22.48″ N, 18°57′38.99″ E, 2005) 6. Union of the Comoros, a. (Mdji Diakagnoa, 11°53′2.65″ S, 43°24′44.32″ E, 2005) b. (Ouralé-Ouandaoé, 11°51′5.98″ S, 43°25′46.85″ E, 2005) c. (Tsinimouachongo, 11°50′38.11″ S, 43°22′35.29″ E, 2005) d. (Zikaledjou, 11°48′25.45″ S, 43°16′47.75″ E, 2005) 7. New Zealand (Pehitawa Forest Reserve, 38°21′26.21″ S, 175°11′35.88″ E, 2006).

## 2.2. OSMAC Approach, Extraction and Metabolomics Analysis of the Fermentation Broths

The importance of growth conditions in the amount and diversity of secondary metabolite production is well recorded. Therefore, each of the 27 strains was cultured in 10 different media (One Strain Many Compounds-OSMAC approach) [20] to identify the best conditions and maximize their biosynthetic potential.

After 14 days of fermentation, the culture broths were extracted with acetone for subsequent liquid chromatography coupled to high resolution mass spectrometry (LC-HRMS) analysis using positive ionization mode. A total of 280 samples were obtained, 270 extracts from the fermentation broths and 10 extracts from the corresponding control media. Initially, a study of the production of the three known phocoenamycin was carried out in each of the conditions analyzed.

All 27 strains produced the three compounds in at least one of the fermentation media. For phocoenamycin, the (+)-HRESIMS data showed a peak at a retention time (RT) of 6.45 min displaying an ammonium adduct ion at  $m/z$  1088.4982  $[M + NH_4]^+$ , corresponding to the molecular formula  $C_{56}H_{75}ClO_{18}$ . For phocoenamycin B, a peak with an adduct ion at  $m/z$  1104.4927  $[M + NH_4]^+$ , in agreement with the molecular formula  $C_{56}H_{75}ClO_{19}$  was detected at a RT of 5.62 min and for phocoenamycin C, a peak at a RT of 6.01 min with

an adduct ion at  $m/z$  1104.4933  $[M + NH_4]^+$ , corresponding to the molecular formula  $C_{56}H_{75}ClO_{19}$ . In all three cases, the characteristic UV spectrum of the phocoenamycin was observed and all the data were in good agreement with those previously published for the three compounds [15].

The manual exploration of the data suggested that phocoenamycin was the most abundant compound, detected in all strains and in almost every culture medium, with the exception of DEF-15, DEF-15S and DNPM for some of the strains. Phocoenamycin B was detected in lower amounts, but still significantly in all strains, and phocoenamycin C was the least produced metabolite and almost absent in extracts of some of the strains (CA-253038, CA-249379, CA-244669 and CA-108000).

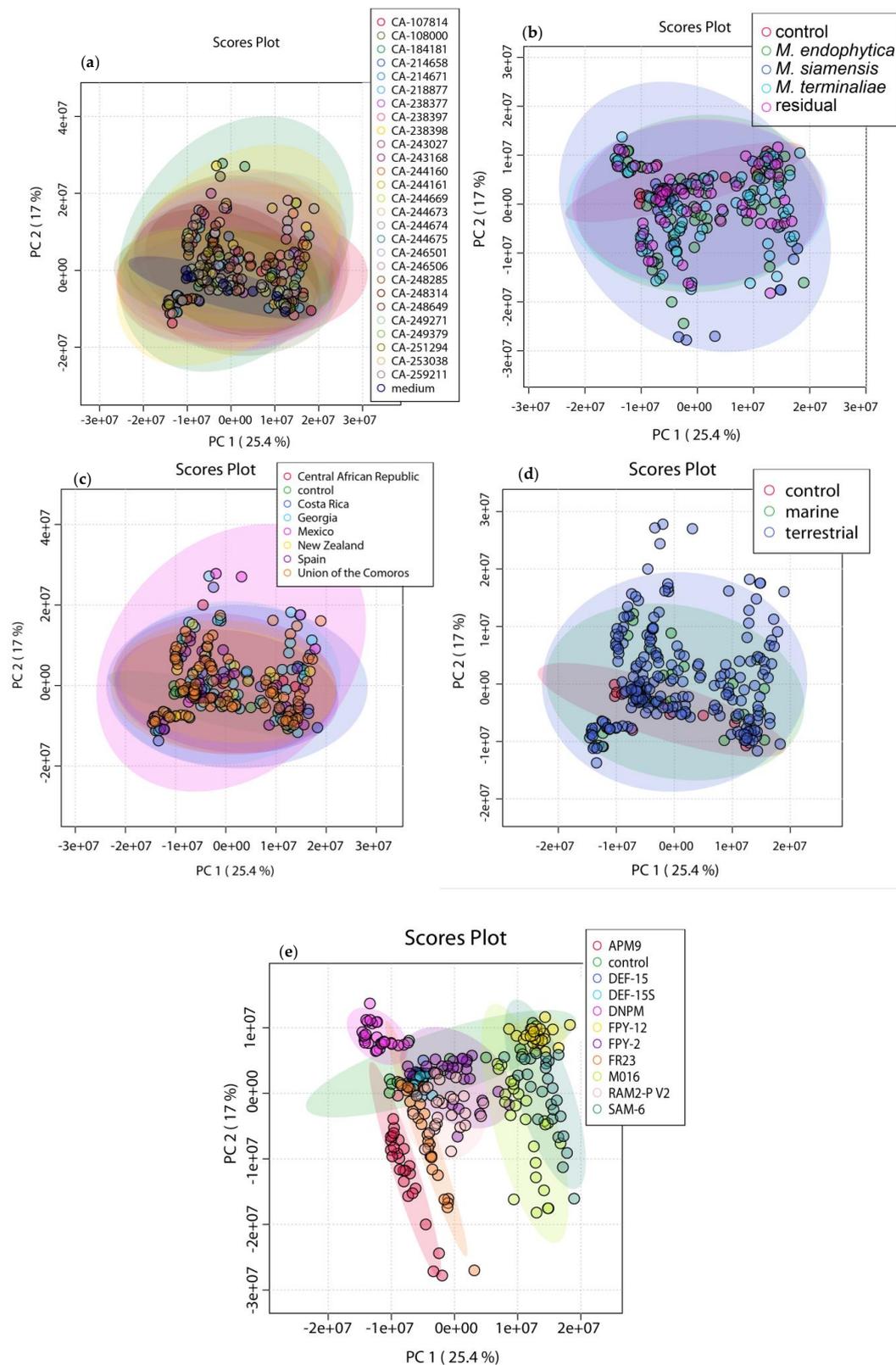
Furthermore, the qualitative study of the production of the three phocoenamycins suggested that, in general for all 27 strains, the most suitable fermentation media to obtain these compounds were RAM2-P V2, FR23, M016, and SAM-6. On the other hand, APM9, FPY-2 and FPY-12 exhibited significantly lower production levels, while in DEF-15, DEF-15S and DNPM the production was very low or non-existent.

The extracts were further subjected to multivariate data analysis (MVDA) in order to explore the production potential of the strains in depth. The raw data obtained from the analysis were preprocessed by MZmine 2 [21] (version 2.53). The treatment included mass detection, chromatogram building and deconvolution, noise reduction and data filtering, peak detection and integration, chromatographic alignment and gap filling [21], resulting in a data matrix of 6050 metabolite features. The samples were organized in groups according to their taxonomy, geography, ecology, culture medium, and the different strains. The processed data were then converted into a suitable format (comma-separated values, CSV file) for each group and was uploaded to the MetaboAnalyst 5.0 platform [22] to carry out the statistical analysis. Missing values were replaced, and further feature filtering based on interquartile range (IQR) reduced the data to 2500 features. The purpose of the data filtering was to identify and remove variables that were unlikely to be of use when modeling the data. Normalization was performed using Pareto scaling and log transformation to make individual features more comparable and transform the data matrix into a more Gaussian-type distribution.

Principal Component Analysis (PCA) was used as an exploratory tool to reveal possible patterns, trends or outliers. PCA reduces data by projecting them onto lower dimensions that are called principal components (PCs). As it is an unsupervised chemometric analysis, it ignores any information about the different groups of the samples and can only distinguish classes when the within-group variation is sufficiently lower than the between-group variation. The aim is to create the best summary of the data while using the least PCs possible [23,24].

The samples were the extracts of the *Micromonospora* fermentations, and the features were peaks defined by an  $m/z$  value and RT. The extracts that only contained the different media without bacterial cultivation were used as control samples. Five different parameters were tested as class labels, in particular the different strains, taxonomic species, culture media, source, and geography and PCA was generated for each one of them to explore if these parameters can influence the metabolite production.

As shown in Figure 2, there was a close clustering of the control samples in all cases, demonstrating that the discrimination observed was mainly due to biological factors. PC1 and PC2 were capable of discriminating samples depending on the culture medium Figure 2e, with a total variance explained of 42.4%. Furthermore, the extracts in media that were previously identified as poor producing, were clustered closer to the control media. There was no detectable grouping based on neither of the other parameters Figure 2a–d. This suggests that such parameters do not contribute to a different chemical profile, since they do not influence the distribution of objects.

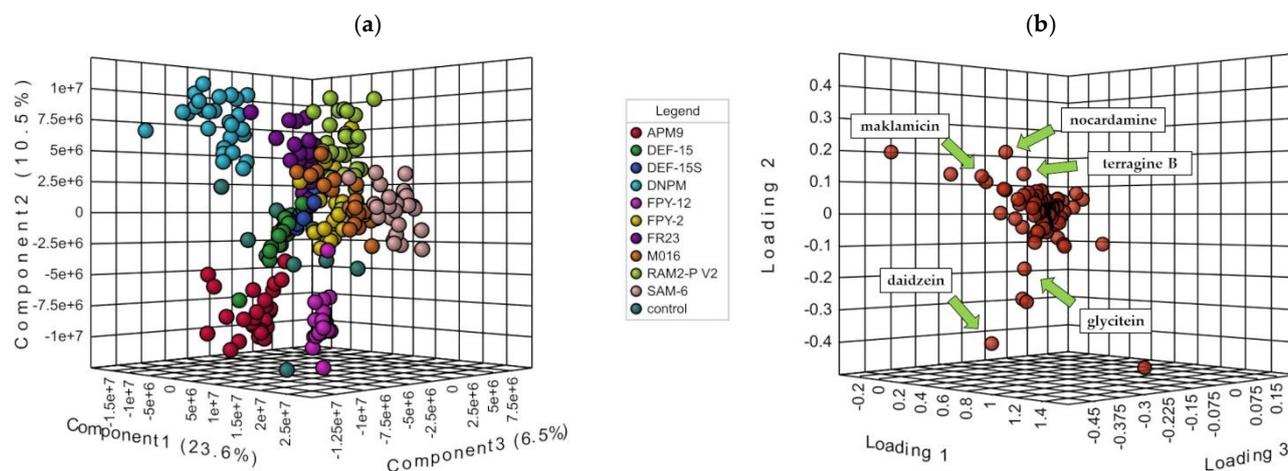


**Figure 2.** PCA 2D-Score plot of principal component 1 (PC1) (variance of 25.4%) and PC2 (variance of 17%), with a total variance of 42.4%. Five different models were tested, using the parameters: (a) different strains; (b) taxonomy species; (c) geographic origin; (d) ecology; and (e) culture medium. Only model (e) managed to separate the samples and distinguish them in groups according to the culture medium.

PCA was combined with partial least squares-discriminant analysis (PLS-DA) to confirm the previous observation and review in detail the chemical variation within the extracts. PLS-DA can be considered as a “supervised” view of PCA as it takes into account the different groups within the samples and maximizes the covariance between the data matrix and the class label. It can also be used for the selection of the most important features that drive the separation between the groups [25].

Since it is prone to overfitting and may create clusters even in random data, cross-validation (CV) and permutation tests are necessary steps when using PLS-DA. A significantly important model shows  $R_2$  (explained variance) and  $Q_2$  (predicted variance) values  $> 0.7$  on cross-validation and  $p$ -value  $< 0.05$  on the permutation test [26].

Only the Y-model that clusters the samples according to the culture medium could be used to explain chemical variation with values of  $R_2 = 0.859$  and  $Q_2 = 0.841$  and  $p$ -value  $< 0.01$  on the permutation test (Figure 3a). Indeed, all the other PLS-DA attempts of the different Y-class models (strain, taxonomy, geography, and ecology) apart from the absence of discrimination (see Supplementary Material, Figure S7), resulted in overfitting and lack of validation. As a conclusion, the only parameter that could provoke a significant change in the metabolic profile of the different *Micromonospora* strains was the culture medium. As far as the control media were concerned, they were closely clustered for the media that were identified as potent phocoenamicin producers but scattered and close to all the extracts of the same medium for the poor producing ones, reaffirming the significant influence of the culture medium in the metabolite production of these strains.



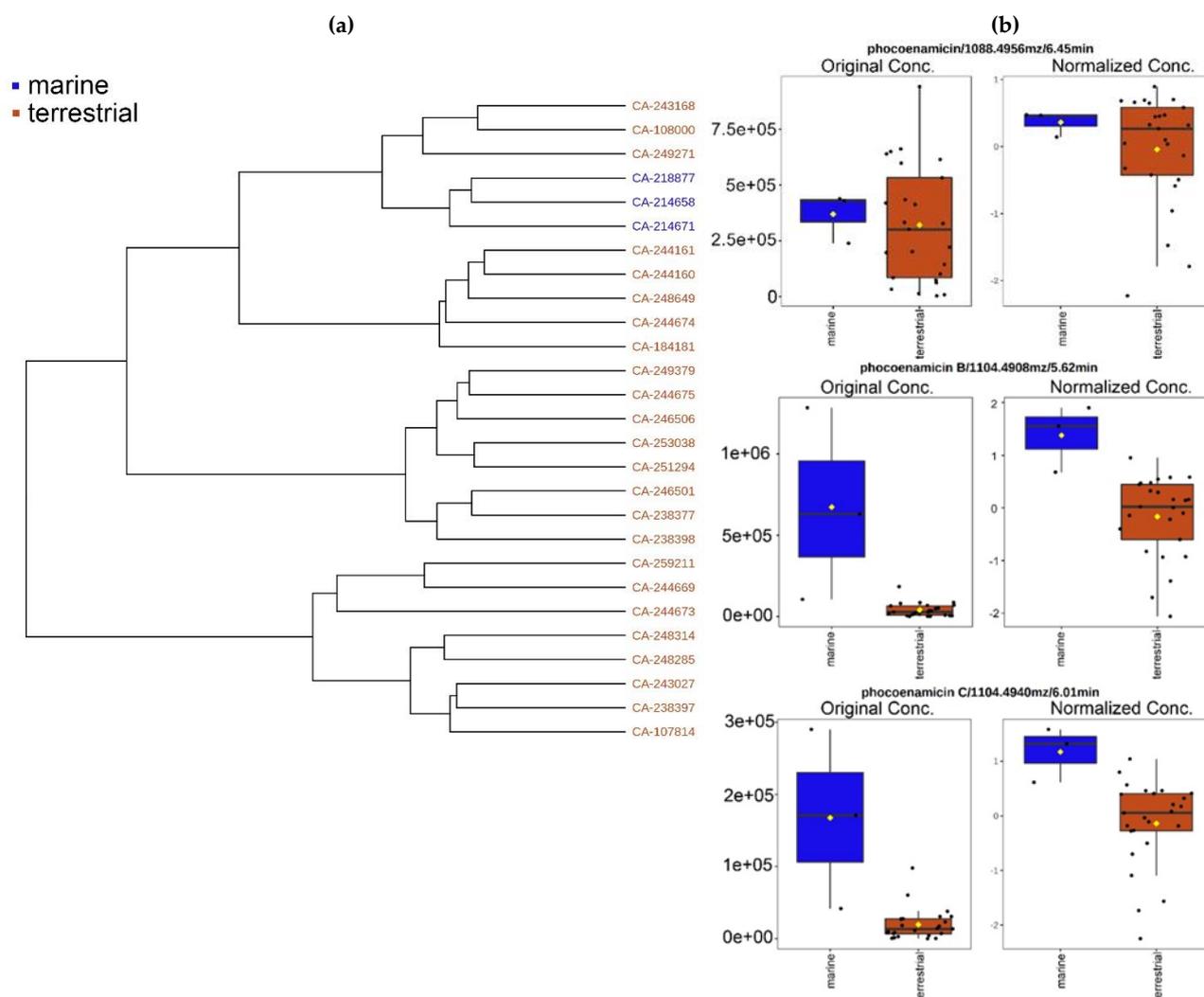
**Figure 3.** (a) PLS-DA 3D-Score plot of the 270 extracts grouped by culture medium and the 10 control extracts where the extracts are clustered according to the culture media; and (b) the main compounds that were responsible for the separation. The components that were further from the center of the plot contributed the most to the separation. Only the ones that accounted for Variable Importance in Projection (VIP) higher than 1.5 were considered.

Furthermore, the 3D loadings plot, as shown in Figure 3b, revealed the main compounds that were responsible for this separation. The components that were further from the center of the plot contributed the most to the separation. From the metabolites observed, only the ones that accounted for Variable Importance in Projection (VIP) higher than 1.5 were considered. Among them, the spirotetronate maklamicin was putatively identified via manual dereplication, as well as daidzein, genistein and the siderophores nocardamine and terragine B (see Section 2.4), confirming the chemical differences between the extracts cultivated in each medium.

As one of the most potent culture media, the RAM2-P V2 medium was selected for further analysis by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) to improve its resolving ability and explore more profoundly the metabolic potential of the strains. Following the same procedure, the data obtained were processed

with MZmine 2 and then analyzed using the MetaboAnalyst platform. The data matrix contained 1287 features and the samples were distributed into two groups according to their ecology: marine-derived, and terrestrial strains.

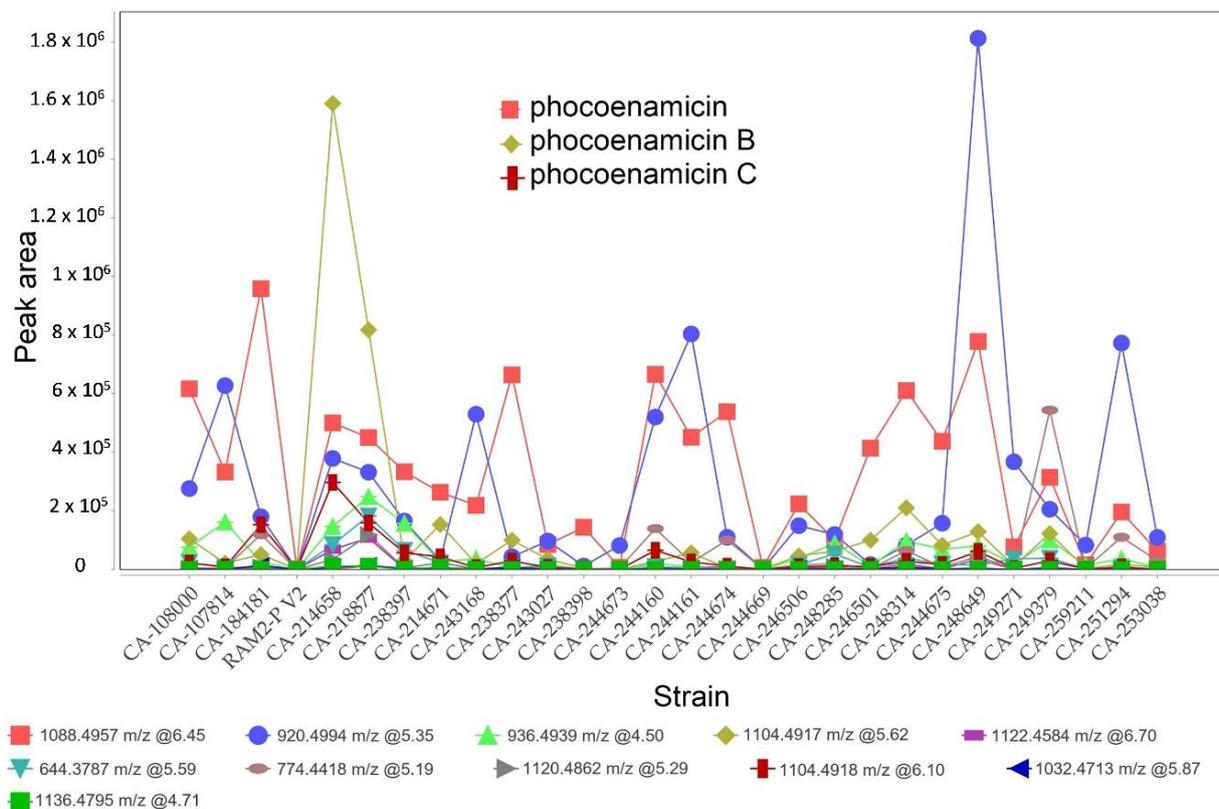
Hierarchical clustering was performed by MetaboAnalyst 5.0 to build a hierarchy of clusters according to the metabolic profile of the 27 strains. Hierarchical cluster analysis (HCA) is an unsupervised quantitative method used to assess the chemical similarity of different samples [27]. Ward's clustering algorithm and Euclidean distance measure were used to create the dendrogram that is shown in Figure 4a. The three marine-derived strains were closely clustered indicating that they have a similar chemical response to the RAM2-P V2 medium. Moreover, the relative peak intensities of the three phocoenamycin were explored comparing the marine and terrestrial strains. As shown in Figure 4b, phocoenamycin was produced more within the terrestrial strains, while phocoenamycin B and C within the marine.



**Figure 4.** (a) Dendrogram generated from the Hierarchical Cluster Analysis showing the close clustering of the three marine-derived strains (blue) compared to the terrestrial (brown); and (b) the relative concentration of the three phocoenamycin, before and after normalization techniques and based on their peak intensities and grouped by their source, marine (blue) and terrestrial strains (brown).

Lastly, keeping in mind the limitation of using non-quantitative mass spectrometry data, the analysis of the results allowed us to identify tendencies based on the peak area. The peak intensity plot of the processed data generated by MZmine 2 (Figure 5) suggested that the maximum levels of phocoenamycin production in RAM2-P V2 medium were obtained

in extracts of the terrestrial strains CA-184181 and CA-238649, and for phocoenamicins B and C in the marine CA-214658 strain, followed by the also marine CA-218877, reaffirming the previous results that phocoenamicin is produced more by the terrestrial strains while phocoenamicins B and C by the marine-derived ones.



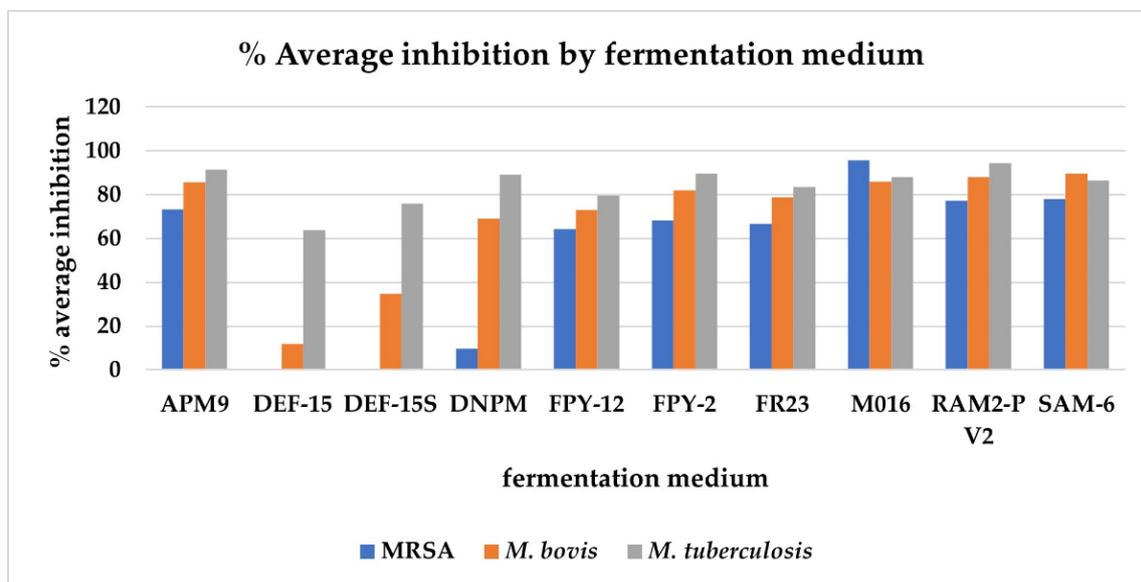
**Figure 5.** Peak intensity plot based on peak area generated by MZmine 2 for the three phocoenamicins along with the putative phocoenamicin analogues identified in RAM2-P V2 medium. The  $m/z$  values, RT and their corresponding colors-shapes are shown for each compound.

Moreover, as shown in Figure 5, the putative phocoenamicin derivatives identified (see Section 2.6) were also analyzed and their production levels were compared within the 27 strains in RAM2-P V2 medium. Among them, the compound with an adduct ion at  $m/z$  920.4935 was produced more by the strains CA-248649, CA-244161 and CA-251294 and the compound with an adduct ion at  $m/z$  774.4410 was produced more by the strain CA-249379. Lastly, the marine strains CA-214658 and CA-218877 demonstrated the highest diversity in producing these minor compounds, suggesting that they could be good candidates for the scale-up of the production and isolation of these putative minor analogues.

### 2.3. Antimicrobial Activity of the Crude Extracts

Antimicrobial activity of the crude extracts indicates the ability of the strains to produce metabolites of potential therapeutic interest [28]. The crude extracts of the 270 fermentations, as well as the 10 extracts from the control media, were subjected to high throughput screening against MRSA MB5393, *M. bovis* BCG and *M. tuberculosis* H37Ra to evaluate the antimicrobial potential of the strains and to study if a correlation could be established between antimicrobial activity and phocoenamicins production. The 10 extracts from the control media did not show any activity against the microorganisms tested, confirming that the bioactivity observed in the crude extracts can be exclusively attributed to the compounds produced by the *Micromonospora* strains and not to the fermentation media components. Figure 6 shows how bioactivity was influenced by the different fermentation

media. The graph represents the average of the percentages of inhibition (%) obtained from each of the extracts from the 27 strains versus the different media in which they had been grown.



**Figure 6.** The average of the percentages of inhibition (%) obtained against methicillin-resistant *Staphylococcus aureus* (MRSA) MB5393, *Mycobacterium bovis* BCG and *Mycobacterium tuberculosis* H37Ra from each of the extracts from the 27 strains versus the different media in which they were grown. Extracts were tested at a concentration of  $0.2 \times$  WBE.

All 27 strains exhibited some type of antimicrobial activity under some of the conditions and several trends were observed. It was observed that in the activity (% inhibition) of the extracts, the medium in which the *Micromonospora* strains were grown was clearly a more determining factor than the different strains themselves. So, the highest inhibition against the three pathogens was observed in the extracts from the fermentation medium M016, followed by the fermentation media RAM2-P V2, SAM-6 and APM9. Comparing the three pathogens, the highest activity was detected against *M. tuberculosis* and then *M. bovis* and MRSA. To the contrary, the lowest inhibition was found in the extracts from fermentation media DEF-15, DEF-15S and DNPM. A correlation between these low levels of inhibition and the low levels of phocoenamycin produced in these media can be established based on the fact that the phocoenamycin has been described as compounds with antimicrobial properties against gram-positive pathogens [15]. However, the same correlation was not observed when comparing other fermentation media, where significant differences in phocoenamycin production, as in the case of FPY-12 and FR23 media, did not translate into significant differences in inhibition. Nevertheless, this could be attributed to synergism between the phocoenamycin and/or other bioactive compounds present in the *Micromonospora* extracts, since they contain complex mixtures. Lastly, these results revealed a rich antimicrobial potential within the 27 strains that should be further investigated. More assays should be performed against more pathogens in order to discover new antimicrobial activities and determine the responsible bioactive compound(s) in each case.

#### 2.4. Dereplication of The Main Compounds Produced

The combination of the LC-UV-MS analyses, metabolomic profiling and molecular networking allowed the putative dereplication of the main compounds produced within the 27 strains in the different growing conditions, as shown in Table 2. The proposed dereplication was accomplished comparing the data against the in-house MEDINA database, the Dictionary of Natural Products (DNP), the COLleCtion of Open Natural ProdUCts

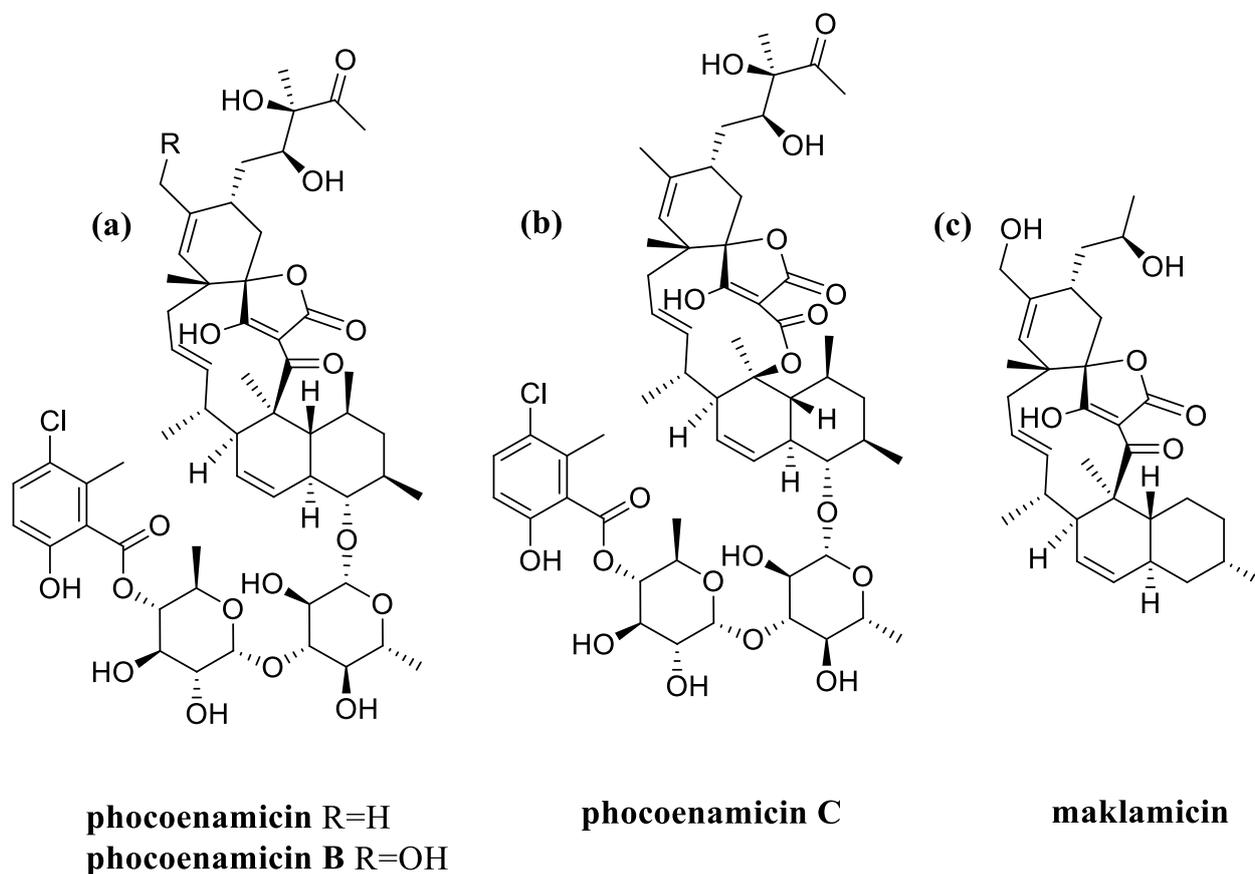
(COCONUT), the Natural Product Activity & Species Source database (NPASS) and the ChemSpider database. The comparative analyses led to the tentative identification of multiple, chemically diverse, compounds. The  $m/z$  detected, the corresponding adduct ion, the theoretical mass of each adduct and its deviation in ppm, the molecular formulae, and the RT are shown for each compound. Due to the detection of the molecules included in the table in multiple samples, representative values of  $m/z$  and retention time are displayed for each compound.

**Table 2.** Putatively dereplicated metabolites detected in the crude extracts of the fermentation broths.

| Compound                         | $m/z$ Detected, Adduction                      | Theoretical Mass ( $\Delta$ ppm) | Molecular Formula  | Retention Time |
|----------------------------------|--|----------------------------------|--|----------------|
| <b>spirotetronates</b>           |  |                                  |  |                |
| phocoenamicin                    | 1088.4982, [M + NH <sub>4</sub> ] <sup>+</sup> | 1088.4980 (+0.2)                 | C <sub>56</sub> H <sub>75</sub> ClO <sub>18</sub>              | 6.44           |
| phocoenamicin B                  | 1104.4927, [M + NH <sub>4</sub> ] <sup>+</sup> | 1104.4929 (−0.2)                 | C <sub>56</sub> H <sub>75</sub> ClO <sub>19</sub>              | 5.53           |
| phocoenamicin C                  | 1104.4933, [M + NH <sub>4</sub> ] <sup>+</sup> | 1104.4929 (+0.4)                 | C <sub>56</sub> H <sub>75</sub> ClO <sub>19</sub>              | 6.02           |
| maklamicin                       | 542.3478, [M + NH <sub>4</sub> ] <sup>+</sup>  | 542.3476 (+0.4)                  | C <sub>32</sub> H <sub>44</sub> O <sub>6</sub>                 | 6.38           |
| <b>siderophores</b>              |  |                                  |  |                |
| nocardamine                      | 601.3551, [M + H] <sup>+</sup>                 | 601.3556 (−0.8)                  | C <sub>27</sub> H <sub>48</sub> N <sub>6</sub> O <sub>9</sub>  | 1.26           |
| deoxynocardamine                 | 585.3589, [M + H] <sup>+</sup>                 | 585.3606 (−2.9)                  | C <sub>27</sub> H <sub>48</sub> N <sub>6</sub> O <sub>8</sub>  | 1.02           |
| demethylenocardamine             | 587.3396, [M + H] <sup>+</sup>                 | 587.3399 (−0.5)                  | C <sub>26</sub> H <sub>46</sub> N <sub>6</sub> O <sub>9</sub>  | 0.95           |
| terragine D                      | 479.2848, [M + H] <sup>+</sup>                 | 479.2864 (−3.3)                  | C <sub>24</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>  | 2.54           |
| terragine B                      | 319.1652, [M + H] <sup>+</sup>                 | 319.1652 (0.0)                   | C <sub>17</sub> H <sub>22</sub> N <sub>2</sub> O <sub>4</sub>  | 2.80           |
| IC 202B                          | 533.3275, [M + H] <sup>+</sup>                 | 533.3293 (−3.4)                  | C <sub>23</sub> H <sub>44</sub> N <sub>6</sub> O <sub>8</sub>  | 1.51           |
| deferoxamine                     | 561.3604, [M + H] <sup>+</sup>                 | 561.3606 (−0.4)                  | C <sub>25</sub> H <sub>48</sub> N <sub>6</sub> O <sub>8</sub>  | 0.70           |
| proferrioxamine A1               | 547.3450, [M + H] <sup>+</sup>                 | 547.3450 (0.0)                   | C <sub>24</sub> H <sub>46</sub> N <sub>6</sub> O <sub>8</sub>  | 0.65           |
| desferrioxamine D1               | 603.3695, [M + H] <sup>+</sup>                 | 603.3712 (−1.1)                  | C <sub>27</sub> H <sub>50</sub> N <sub>6</sub> O <sub>9</sub>  | 1.06           |
| legonoxamine A                   | 637.3914, [M + H] <sup>+</sup>                 | 637.3919 (−0.8)                  | C <sub>31</sub> H <sub>52</sub> N <sub>6</sub> O <sub>8</sub>  | 2.34           |
| legonoxamine G                   | 623.3763, [M + H] <sup>+</sup>                 | 623.3763 (0.0)                   | C <sub>30</sub> H <sub>50</sub> N <sub>6</sub> O <sub>8</sub>  | 2.21           |
| legonoxamine H                   | 665.3861, [M + H] <sup>+</sup>                 | 665.3869 (−1.2)                  | C <sub>32</sub> H <sub>52</sub> N <sub>6</sub> O <sub>9</sub>  | 2.47           |
| microferrioxamine B              | 729.5473, [M + H] <sup>+</sup>                 | 729.5484 (−1.5)                  | C <sub>37</sub> H <sub>72</sub> N <sub>6</sub> O <sub>8</sub>  | 4.78           |
| microferrioxamine C              | 743.5632, [M + H] <sup>+</sup>                 | 743.5641 (−1.2)                  | C <sub>38</sub> H <sub>74</sub> N <sub>6</sub> O <sub>8</sub>  | 4.97           |
| microferrioxamine D              | 757.5792, [M + H] <sup>+</sup>                 | 757.5797 (−0.7)                  | C <sub>39</sub> H <sub>76</sub> N <sub>6</sub> O <sub>8</sub>  | 5.21           |
| acyl ferrioxamine 2              | 679.4015, [M + H] <sup>+</sup>                 | 679.4025 (−1.4)                  | C <sub>33</sub> H <sub>54</sub> N <sub>6</sub> O <sub>9</sub>  | 2.78           |
| <b>various</b>                   |  |                                  |  |                |
| daidzein                         | 255.0648, [M + H] <sup>+</sup>                 | 255.0652 (−1.6)                  | C <sub>15</sub> H <sub>10</sub> O <sub>4</sub>                 | 2.55           |
| genistein                        | 271.0594, [M + H] <sup>+</sup>                 | 271.0601 (−2.6)                  | C <sub>15</sub> H <sub>10</sub> O <sub>5</sub>                 | 3.20           |
| glycitein                        | 285.0754, [M + H] <sup>+</sup>                 | 285.0757 (−1.1)                  | C <sub>16</sub> H <sub>12</sub> O <sub>5</sub>                 | 2.67           |
| antascomicin D                   | 646.3946, [M + H] <sup>+</sup>                 | 646.3950 (−0.6)                  | C <sub>36</sub> H <sub>55</sub> NO <sub>9</sub>                | 3.80           |
| 21-demethyl-leptomycin A         | 513.3208, [M + H] <sup>+</sup>                 | 513.3211 (−0.6)                  | C <sub>31</sub> H <sub>44</sub> O <sub>6</sub>                 | 5.07           |
| indothiazinone-4-carboxylic acid | 273.0327, [M + H] <sup>+</sup>                 | 273.0328 (−0.4)                  | C <sub>13</sub> H <sub>8</sub> N <sub>2</sub> O <sub>3</sub> S | 2.98           |
| anandin A                        | 358.2743, [M + H] <sup>+</sup>                 | 358.2741 (+0.6)                  | C <sub>23</sub> H <sub>35</sub> NO <sub>2</sub>                | 4.52           |
| ganefromycin epsilon             | 644.3794, [M + H] <sup>+</sup>                 | 644.3793 (+0.2)                  | C <sub>36</sub> H <sub>53</sub> NO <sub>9</sub>                | 4.70           |
| actiphenol                       | 276.1238, [M + H] <sup>+</sup>                 | 276.1230 (+2.9)                  | C <sub>15</sub> H <sub>17</sub> NO <sub>4</sub>                | 2.48           |
| N <sup>b</sup> -Acetyltryptamine | 203.1177, [M + H] <sup>+</sup>                 | 203.1179 (−1.0)                  | C <sub>12</sub> H <sub>14</sub> N <sub>2</sub> O               | 2.41           |
| N-Acetyltyramine                 | 180.1023, [M + H] <sup>+</sup>                 | 180.1019 (+2.2)                  | C <sub>10</sub> H <sub>13</sub> NO <sub>2</sub>                | 0.92           |
| actiphenamide                    | 294.1331, [M + H] <sup>+</sup>                 | 294.1336 (−1.7)                  | C <sub>15</sub> H <sub>19</sub> NO <sub>5</sub>                | 2.48           |
| antibiotic BE 54476              | 392.2429, [M + H] <sup>+</sup>                 | 392.2431 (−0.5)                  | C <sub>22</sub> H <sub>33</sub> NO <sub>5</sub>                | 4.14           |

As mentioned above, the three phocoenamicins were putatively detected within all the extracts. Along with the spirotetronates of interest, maklamicin was also tentatively identified to be produced by all 27 strains. Maklamicin is also a spirotetronate first found to be produced by an endophytic *Micromonospora* sp. collected in Thailand and has shown strong to modest antimicrobial activities against various gram-positive bacteria [29]. As shown in Figure 7, structurally, maklamicin is closely related to phocoenamicins, having 11 carbons on its macrocycle and a decalin unit, as phocoenamicins. The closest proximity is with phocoenamicin B, their major differences being the absence of the two 6-deoxyglucose

moieties, the 3-chloro-6-hydroxy-2-methylbenzoate moiety, and the diol side chain, unique in the phocoenamincins [15,29]. Within the 270 extracts of this study, the production of one of the two families conditioned the production of the other in all cases. Furthermore, phocoenamincin and maklamicin co-eluted in the crude extracts of the fermentation broths and further purification would be needed for their separation.



**Figure 7.** Structures of: (a) phocoenamincin and phocoenamincin B; (b) phocoenamincin C; and (c) maklamicin.

Apart from the spirotetronates, the majority of the compounds putatively dereplicated belonged to the hydroxamate siderophores, and in particular to the desferrioxamine family, with the most common being deferoxamine [30], nocardamine [31], terragine B [32] and IC 202B [33]. The hydroxamate siderophores have therapeutic and diagnostic importance and have demonstrated antimicrobial and antitumor activities. Furthermore, antibiotics linked to siderophores, natural conjugates named sideromycins, have shown a higher antibacterial efficacy compared to normal antibiotics due to an enhanced uptake using the siderophore-mediated iron active transport and their use has been explored in a “Trojan horse” approach to overcome bacterial resistance to antibiotics [34,35]. As metal chelators, Desferal, the brand name of desferrioxamine B, was the first antibiotic used to remove excess iron in patients suffering from iron toxicity or overload, such as in the case of  $\beta$ -thalassemia patients that depend on blood transfusions [36].

Finally, various other compounds were tentatively identified within the extracts. The most abundant were the isoflavones daidzein [37], genistein [38] and glycitein [39] that have antioxidant, anticancer and antimicrobial properties [37,40] and several more metabolites with antibiotic properties, such as 21-demethyl-leptomycin A [41], that could explain the differences observed between the antimicrobial activity and the production levels of phocoenamincins in some of the extracts. The wide and diverse metabolic profile of these strains highlights their therapeutic potential and the need for further investigation.

### 2.5. Molecular Networking Analysis

Over the last few years, molecular networking (MN) has gained a lot of attention in the field of natural products. MN is a bioinformatic tool that explores chemical diversity visualizing the entire metabolome detected in a dataset. It is based on the assumption that chemistry influences how the molecules will be fragmented by tandem mass spectrometry. Therefore, structurally related compounds will have a similar fragmentation pattern and will be connected whereas the unrelated will be separated [42].

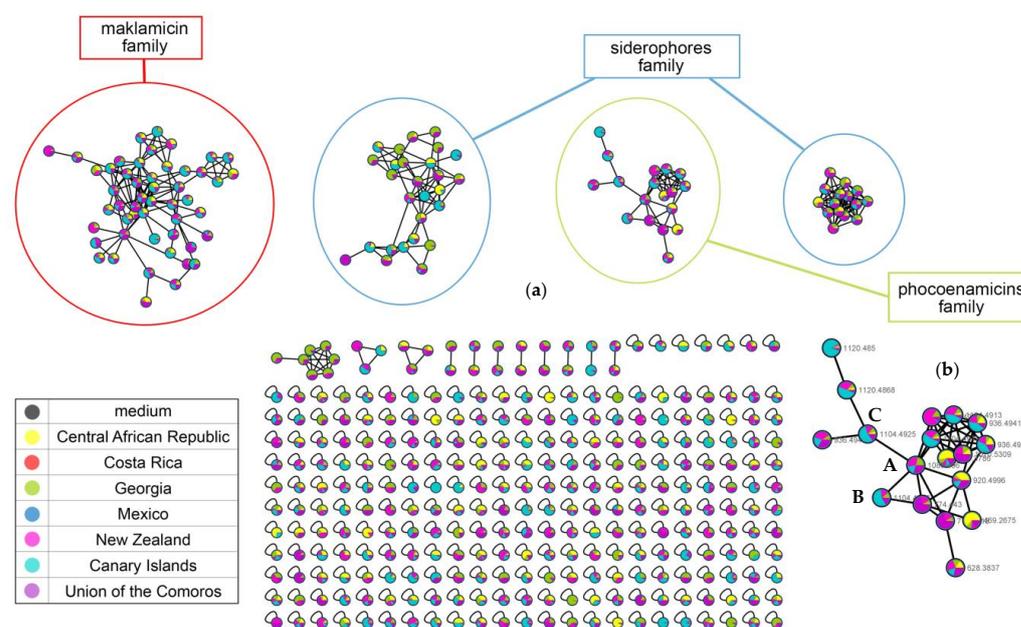
In a molecular network, each node represents a compound and the nodes with similar spectra are linked to form clusters or “molecular families”. Moreover, these links provide valuable structural information. Compounds that belong to the same cluster usually share the same core structure but differ due to simple transformations such as alkylation or oxidation. For example, a difference of 14 Da could suggest putative CH<sub>2</sub> homologues and a difference of 16 Da an oxygenated analogue [43].

Molecular networking has led to the creation of Global Natural Products Social Molecular Networking (GNPS), a web-based platform (<http://gnps.ucsd.edu>, accessed on 2 December 2022) developed by researchers at the University of California at San Diego. GNPS uses an algorithm to compare the MS/MS spectra and is based on various parameters like the cosine score. After the analysis, searching of the GNPS libraries, in-house databases, and other public spectral libraries can lead to the rapid annotation of known compounds, making GNPS an important dereplication tool [11].

GNPS offers two main workflows, the classical MN and feature-based molecular networking (FBMN). In classical MN, raw MS/MS spectra files are processed directly to generate a MN. This often leads to multiple nodes for the same compound when detected over a large retention time span and creates huge MNs that are not fully representative of the actual number of compounds in the dataset. In order to overcome this problem, FBMN uses a feature detection and alignment tool, such as MZmine 2, to preprocess the raw data and is used for advanced molecular networking analysis, enabling the relative quantification of the compounds and resolution of isomers [44].

MN was used to organize the MS/MS data obtained from the crude extracts of the 27 strains grown in the medium RAM2-P V2 that was used as well as a control. Classical MN was dismissed because of the massive and difficult task to interpret the molecular network it created, with 2514 nodes organized in 65 clusters (at least two nodes connected). The results of the analysis are publicly available at: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=37e734c22f5f43cab7f8983d5121cea6>, accessed on 2 December 2022. Even though an overview of the metabolites and their relationships could be observed, FBMN workflow was chosen for further analysis, leaving behind many nodes that were a replicate of the same compound and compounds that were only produced in traces within the extracts, therefore providing a more realistic view of the *Micromonospora* production. The raw data were preprocessed by MZmine 2 [21] and then uploaded to the GNPS platform to perform the analysis. A metadata file was also uploaded to describe the properties of each sample-strain (taxonomy, geography, ecology). The results were visualized using the Cytoscape software (version 3.8.2) and are publicly available at: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=2dfa24f8cd0c404eb980cef3580d3eac>, accessed on 2 December 2022.

The FBMN analysis resulted in 447 nodes (parent ions) that were organized in 15 clusters (molecular families, at least two nodes connected). The four main clusters were putatively identified. Two of them were created by the siderophores, one by the phocoenamycin and one by maklamicin and related compounds (Figure 8a). Many of the nodes remained unannotated, highlighting the need for further study of the potential of these strains. Some of them that were related to known compounds were tentatively characterized as new according to the manual dereplication process (see Section 2.6).



**Figure 8.** (a) The molecular network generated from the FBMN analysis of the 27 strains grown in the RAM2-P V2 medium. The medium was used as control to remove the compounds that were not produced by the strains. The nodes are colored according to the geographic distribution of the strains that produced them. Four main clusters were detected, two created by the siderophores, one by maklamicin and one more from the phocoenamycins; and (b) the phocoenamycins cluster revealed 18 nodes and the phocoenamycins were putatively identified as phocoenamycin (A), phocoenamycin B (B) and phocoenamycin C (C). *m/z* values of new derivatives in the network are detailed in Table 3.

**Table 3.** Potential derivatives detected in the crude extracts of the fermentation broths.

| Compound                 | <i>m/z</i> Detected, Adduction                 | Theoretical Mass ( $\Delta$ ppm) | Molecular Formula Predicted                                     | Retention Time |
|--------------------------|--|----------------------------------|---|----------------|
| <b>spirotetronates</b>   |  |                                  |   |                |
| phocoenamycin derivative | 1120.4862, [M + NH <sub>4</sub> ] <sup>+</sup> | 1120.4878 (−1.4)                 | C <sub>56</sub> H <sub>75</sub> ClO <sub>20</sub>               | 5.29           |
| phocoenamycin derivative | 1032.4713, [M + NH <sub>4</sub> ] <sup>+</sup> | 1032.4718 (−0.5)                 | C <sub>53</sub> H <sub>71</sub> ClO <sub>17</sub>               | 5.70           |
| phocoenamycin derivative | 1122.4584, [M + NH <sub>4</sub> ] <sup>+</sup> | 1122.4590 (−0.5)                 | C <sub>56</sub> H <sub>74</sub> Cl <sub>2</sub> O <sub>18</sub> | 6.71           |
| phocoenamycin derivative | 1136.4795, [M + NH <sub>4</sub> ] <sup>+</sup> | 1136.4828 (−2.9)                 | C <sub>56</sub> H <sub>75</sub> ClO <sub>21</sub>               | 4.54           |
| phocoenamycin derivative | 920.4994, [M + NH <sub>4</sub> ] <sup>+</sup>  | 920.5002 (−0.9)                  | C <sub>48</sub> H <sub>70</sub> O <sub>16</sub>                 | 5.24           |
| phocoenamycin derivative | 936.4939, [M + NH <sub>4</sub> ] <sup>+</sup>  | 936.4951 (−1.3)                  | C <sub>48</sub> H <sub>70</sub> O <sub>17</sub>                 | 4.46           |
| phocoenamycin derivative | 644.3787, [M + NH <sub>4</sub> ] <sup>+</sup>  | 644.3793 (−0.9)                  | C <sub>36</sub> H <sub>50</sub> O <sub>9</sub>                  | 5.55           |
| phocoenamycin derivative | 774.4418, [M + NH <sub>4</sub> ] <sup>+</sup>  | 744.4423 (−0.7)                  | C <sub>42</sub> H <sub>60</sub> O <sub>12</sub>                 | 5.19           |
| maklamicin derivative    | 509.3250, [M + H] <sup>+</sup>                 | 509.3262 (−2.4)                  | C <sub>32</sub> H <sub>44</sub> O <sub>5</sub>                  | 7.30           |
| maklamicin derivative    | 495.3091, [M + H] <sup>+</sup>                 | 495.3105 (−2.8)                  | C <sub>31</sub> H <sub>42</sub> O <sub>5</sub>                  | 7.09           |
| maklamicin derivative    | 511.3060, [M + H] <sup>+</sup>                 | 511.3054 (+1.2)                  | C <sub>31</sub> H <sub>42</sub> O <sub>6</sub>                  | 6.12           |
| maklamicin derivative    | 539.3350, [M + H] <sup>+</sup>                 | 539.3367 (−3.2)                  | C <sub>33</sub> H <sub>46</sub> O <sub>6</sub>                  | 6.60           |
| maklamicin derivative    | 523.3049, [M + H] <sup>+</sup>                 | 523.3054 (−1.0)                  | C <sub>32</sub> H <sub>42</sub> O <sub>6</sub>                  | 6.91           |
| <b>siderophores</b>      |  |                                  |   |                |
| nocardamine derivative   | 599.3379, [M + H] <sup>+</sup>                 | 599.3399 (−3.3)                  | C <sub>27</sub> H <sub>46</sub> N <sub>6</sub> O <sub>9</sub>   | 1.85           |
| deferroxamine derivative | 631.4384, [M + H] <sup>+</sup>                 | 631.4389 (−0.8)                  | C <sub>30</sub> H <sub>58</sub> N <sub>6</sub> O <sub>8</sub>   | 2.88           |
| deferroxamine derivative | 645.4531, [M + H] <sup>+</sup>                 | 645.4545 (−2.2)                  | C <sub>31</sub> H <sub>60</sub> N <sub>6</sub> O <sub>8</sub>   | 3.23           |
| deferroxamine derivative | 687.4638, [M + H] <sup>+</sup>                 | 687.4651 (−1.9)                  | C <sub>33</sub> H <sub>62</sub> N <sub>6</sub> O <sub>9</sub>   | 3.44           |
| deferroxamine derivative | 659.4342, [M + H] <sup>+</sup>                 | 659.4338 (+0.6)                  | C <sub>31</sub> H <sub>58</sub> N <sub>6</sub> O <sub>9</sub>   | 2.94           |
| deferroxamine derivative | 671.4694, [M + H] <sup>+</sup>                 | 671.4702 (−1.0)                  | C <sub>33</sub> H <sub>62</sub> N <sub>6</sub> O <sub>8</sub>   | 3.66           |

Focusing on the phocoenamycins cluster, the three phocoenamycins were putatively detected clustering together (A–C, Figure 8b). As shown in Figure 8b, the cluster was

composed of 18 nodes in total. Some of the nodes represented adducts of the same molecule and this was taken into account upon their tentative identification, based on the RT span and the MS/MS fragments, resulting in 10 different compounds. The nodes are marked by experimental  $m/z$  values and colored according to the geographic distribution of the strains that produced them. As it is shown, they are widely distributed in almost all continents. Moreover, taking into account that the FBMN analysis provides relative quantification, in the culture medium RAM2-P V2, phocoenamicins B and C were produced more by the marine-derived strains isolated from the Canary Islands, while phocoenamicin is produced more by the rest of the strains, in agreement with the previously mentioned observations. Furthermore, some nodes could not be identified, indicating that they could putatively represent new analogues of this family of compounds. For example, the parent ion with  $m/z$  1120.4868 that was directly connected with phocoenamicin C showed a difference of 16 Da, indicating a possible oxygenation and the proposed molecular formula  $C_{56}H_{75}ClO_{20}$  generated for the compound was in agreement (see Section 2.6).

### 2.6. Putative Identification of Possible New Compounds

The LC-UV-HRMS analysis and the manual dereplication tentatively detected several peaks corresponding to compounds with molecular formulae that were not found within various databases, therefore are putatively new metabolites related to the phocoenamicins, the maklamicin or the siderophores.

In particular, the annotations were based on GNPS, searching the compounds that were in each molecular cluster and combining the manual putative dereplication of  $m/z$ , UV profile, RT and proposed molecular formula against the databases previously mentioned (DNP, COCONUT, NPASS and ChemSpider). Table 3 summarizes the putatively unknown compounds grouped by family. Their  $m/z$  values and adduct ions detected, theoretical masses and deviation in ppm, the predicted molecular formulae and the chromatographic retention time are shown. As in Table 2, due to the detection of these molecules in multiple analytical runs, representative values of  $m/z$  and retention times for each compound are displayed.

Several phocoenamicin derivatives were tentatively detected. Some of the molecular formulae predicted ( $C_{48}H_{70}O_{16}$ ,  $C_{48}H_{70}O_{17}$ ,  $C_{36}H_{50}O_9$ ,  $C_{42}H_{60}O_{12}$ ) proposed a possible change in the oligosaccharide motif of phocoenamicins that was not observed before within the family. Spirotetronates have exhibited a huge structural variety and it has been demonstrated that even a small change in the structure of a spirotetronate can result in changing its bioactivity [45]. Therefore, large-scale fermentation should be performed in order to isolate these compounds, elucidate their structure and evaluate their potential bioactivity.

Maklamicin [29] is the only member of the family naturally produced so far. 29-deoxymaklamicin, that tentatively matched the accurate mass (508.3169) and the proposed molecular formula ( $C_{32}H_{44}O_5$ ) of one of these putatively new compounds, was also reported before, but as a product of a genetically engineered strain [46]. The GNPS analysis revealed five assumably new analogues of maklamicin in total and further research should be carried out to explore the full potential of this family. Likewise, various siderophore-related and putatively new compounds were detected in agreement with the literature reporting that the fountain of actinomycetes producing new siderophores is infinite [47].

## 3. Discussion

The microbial strains explored in this study as possible phocoenamicins producers belonged to the *Micromonospora* genus, with *M. terminaliae* and *M. endophytica* getting assigned more than half of all the strains and it can be proposed that the biosynthesis of the phocoenamicins is restrained within *Micromonospora* sp. Despite the close taxonomic relation of the strains, they displayed great ecological and geographical diversity. Even though the phocoenamicins were originally isolated from marine actinomycetes, the producers identified in this study were isolated from very diverse ecosystems, resulting in 24 terrestrial strains and only three marine-derived strains. Regarding their geographical

origin, they were widely distributed throughout the planet, covering almost all continents, confirming the hypothesis that this family of compounds is highly conserved within the *Micromonospora* community.

Using an OSMAC approach, the 27 strains were cultivated in 10 different media each, resulting in 270 fermentations, whose crude extracts were analyzed by LC-HRMS and explored by metabolomics analysis and molecular networking. All 27 strains produced the three known phocoenamycin in at least one of the fermentation media. The manual exploration of the data suggested that phocoenamycin was the most abundant compound, followed by phocoenamycin B and C. Furthermore, the analysis of the production of the three phocoenamycins suggested that, in general for all the 27 strains, the most suitable culture media were RAM2-P V2, FR23, M016 and SAM-6. The influence of the different strains, taxonomic species, culture media, ecological source and geography was studied and the only parameter that could provoke a significant change in the metabolic profile was the culture medium.

The potent RAM2-P V2 medium was selected for further analysis by LC-MS/MS. The three marine-derived strains had a similar chemical response to the medium in comparison to the terrestrial ones. As a result, the relative peak intensities of the three phocoenamycins were analyzed comparing the marine and terrestrial strains. Phocoenamycin was produced more within the terrestrial strains (maximum levels from the strains CA-184181 and CA-238377), while phocoenamycins B and C within the marine strains (CA-214658 and CA-218877).

The crude extracts of the 270 fermentations were tested against methicillin-resistant *S. aureus* (MRSA) MB5393, *M. bovis* BCG and *M. tuberculosis* H37Ra to evaluate the antimicrobial potential of the strains. All 27 strains exhibited some type of antimicrobial activity under some of the conditions and it was observed that the medium in which the *Micromonospora* strains were grown was clearly a more determining factor than the different strains themselves. The results revealed a rich antimicrobial potential within the 27 strains that should be further investigated.

The combination of the LC-UV-HRMS analysis, metabolomic profiling and molecular networking allowed the putative dereplication of the main compounds produced within the 27 strains in the different growing conditions. The comparative analyses led to the tentative identification of multiple, chemically diverse compounds that were divided into three main groups, the spirotetronates, the siderophores and various other compounds.

The spirotetronate maklamicin, structurally related to phocoenamycins, was also tentatively identified to be produced by all 27 strains. Within the 270 extracts of this study, the production of maklamicin conditioned the production of the phocoenamycins in all cases. The coexistence of maklamicin and the phocoenamycins in the extracts analyzed confirmed the obvious common biosynthetic origin of both families of compounds.

Apart from the spirotetronates, many of the compounds putatively dereplicated belong to the hydroxamate siderophores, as well as various other compounds that were tentatively identified within the extracts with the most abundant being the isoflavones daidzein, genistein and glycitein. This wide and diverse metabolic profile of these strains highlights the need for further research.

Lastly, the analyses tentatively revealed the presence of various peaks of compounds with predicted molecular formulae that were not found in the literature, putatively new metabolites related to the phocoenamycins (8 analogues), the maklamicin (5 analogues) or siderophores (6 analogues) and thus candidates for isolation and discovery of new bioactive compounds. Large-scale fermentation should be performed in order to isolate these compounds, elucidate their structure and evaluate their potential bioactivity.

In this study, we have clearly demonstrated that the combination of manual exploration, metabolomics analysis and molecular networking can be a rapid and efficient way to prioritize strains and further research should be carried out to unveil the chemical diversity of these promising *Micromonospora* strains. Additionally, our research reveals a wide geographical distribution of phocoenamycin producing microbial strains, including

both marine and terrestrial strains, opposed to initial statements about phocoenamycin production only in unique marine environments. Our findings have also identified marine-derived *Micromonospora* species as the best producers of phocoenamycins in terms of both the abundance in their extracts of some major members of the structural class and in the variety of molecular structures produced.

#### 4. Materials and Methods

##### 4.1. Taxonomical Identification of the Producing Microorganisms

The high-performance liquid chromatography coupled to mass spectrometry (HPLC-UV-MS) profile of the phocoenamycins, including RT, UV spectrum and positive and negative mass spectra, was used to search an internal repository of LC/UV/MS analytical data corresponding to 74,647 extracts originated from 16,366 prokaryotic strains (95.6% actinomycetes). This search identified 27 microbial strains belonging to the MEDINA culture collection that potentially biosynthesized at least one of the phocoenamycins.

Genomic DNA from the 27 strains was isolated as previously described [48]. The 16S rRNA gene was PCR-amplified employing the universal eubacterial primers fd1 (50-AGAGTTTGATCCTGGCTCAG-30) and rP2 (50-ACGGCTACCTTGTTACGACTT-30). PCR reactions were carried out in a final volume of 50  $\mu$ L containing 2  $\mu$ L of dNTPs (10 mM each), 2  $\mu$ L of each of the primers (10  $\mu$ M), 2  $\mu$ L of the DNA dilution, 5  $\mu$ L of PCR buffer (10 $\times$ ), and 0.4  $\mu$ L of Taq Polymerase (5 U/ $\mu$ L). Amplifications were performed on an iCycler iQ<sup>TM</sup> Real-Time PCR Detection System (Bio-Rad Laboratories, Inc, Hercules, CA, USA).

The PCR products were purified and sequenced at Secugen S. L. (Madrid, Spain). For each product the two strands were sequenced employing the primers mentioned above. The resulting DNA sequence lectures were aligned and visually inspected with Bionumerics 6.6 (Applied Maths, Sint-Martens-Latem, Belgium).

##### 4.2. Fermentation of the Producing Microorganisms and OSMAC Approach

The fermentation for each of the 27 strains was generated as follows: a seed culture of the strain was obtained by inoculating a 25  $\times$  150 mm tube containing 10 mL of ATCC-2 medium (potato starch 2 g/L, dextrose 1 g/L, NZ Amine Type E 5 g/L, meat extract 3 g/L, peptone 5 g/L, yeast extract 5 g/L, calcium carbonate 1 g/L, pH 7) with 0.5 mL of a freshly thawed inoculum stock of the producing strain. The tubes were incubated in a rotary shaker at 28  $^{\circ}$ C, 70% relative humidity and 220 rpm for about 7 days (Kühner incubator model ISF-4-V, Adolf Kühner AG, Brisfelden, Switzerland).

The fresh inoculum thus generated was mixed and employed to inoculate (5% *v/v*) EPA vials (28  $\times$  90 mm glass tubes) each containing 10 mL of fermentation medium. The fermentation media used were as follows: APM9, DEF-15, DEF-15S, DNPM, FPY-12, FPY-2, FR23, M016, RAM2-P V2 and SAM-6. Their composition is summarized in Supplementary Material, Table S1. In parallel, a negative control of 10 uninoculated culture media was performed to ensure that any compounds of interest are produced by the microorganisms. All EPA vials were incubated in a rotary shaker at 28  $^{\circ}$ C, 70% relative humidity, and 220 rpm for 14 days before harvesting (Kühner ISF-4-V incubator, Adolf Kühner AG, Brisfelden, Switzerland).

##### 4.3. Extraction of the Fermentation Broths and Analysis by LC-HRMS

The fermentation broths were extracted by adding 10 mL of acetone to each 10 mL fermentation, including the negative control media. The mixtures were then vortexed and shaken in an orbital shaker incubator (Kühner ISF-1-W incubator, Adolf Kühner AG, Brisfelden, Switzerland) at 200 rpm for 1 h, centrifuged for 10 min for the separation of the mycelium, and 12 mL of the supernatant were transferred to 16 mm TurboVap<sup>®</sup> tubes (Biotage, Uppsala, Sweden) to which 600  $\mu$ L of dimethyl sulfoxide (DMSO) were added. Finally, evaporation of the samples was carried out in a hot nitrogen flow cabinet, until a final volume of 3.0 mL to reach a final concentration of 2  $\times$  WBE (Whole Broth Equivalent,

that is, the natural concentration of each metabolite in the culture broth) in 20% DMSO in water. The contents of the tubes were transferred to 0.8 mL 96-Well Deep Well plates (Abgene™, Waltham, MA, USA) using the MultiPROBE® II HT robotic liquid handling system (Packard Bioscience Co, Meriden, CO, USA), transferring 540 µL to each well in the plate.

The extracts obtained (as well as the control media) were subjected to liquid chromatography-high resolution mass spectrometry (LC-HRMS) and furthermore tandem mass spectrometry for the extracts obtained from the RAM2-P V2 medium. ESI-TOF and MS/MS spectra were acquired using a Bruker maXis QTOF (Bruker Daltonik GmbH, Bremen, Germany) mass spectrometer coupled to an Agilent 1200 LC (Agilent Technologies, Waldbronn, Germany) with a standard 10-min reversed-phase gradient chromatographic run, as described before [49].

#### 4.4. Multivariate Data Analysis (MVDA) with MZmine 2 and MetaboAnalyst 5.0

The LC-MS raw data were initially converted into mzXML files using MSConvert (ProteoWizard, <http://proteowizard.sourceforge.net>, accessed on 2 December 2022). The converted datasets were imported into MZmine v2.53, a bioinformatics tool for differential analysis of mass spectrometry data [21]. Peak detection was achieved by noise removal, chromatogram construction, and peak deconvolution. Firstly, the mass values were detected using the centroid mode in each spectrum and the noise level of the peaks was set to 100. Then, chromatograms were constructed for each of the mass values using the ADAP chromatogram builder [50], where group intensity threshold and minimum highest intensity were set to 300 and  $m/z$  tolerance to 20 ppm. Next, the “local minimum search” deconvolution algorithm was applied to each constructed chromatogram of each mass ion to detect the individual peaks. For this algorithm, the chromatographic threshold was set to 10%, the search minimum in RT range to 0.2 min, the minimum relative height to 10%, the minimum absolute height to 100 and the minimum ratio of peak top/edge to 2. The separated peaks were then deisotoped using the function of isotopic peaks grouper in which  $m/z$  tolerance was set to 20 ppm, RT tolerance to 0.5 absolute (min), maximum charge of 3 and representative isotope as the most intense. The remaining peaks in different samples were aligned based on the mass and RT of each peak, creating a peak list. The ion  $m/z$  tolerance for alignment was set to 20 ppm, while RT to 0.5 min, and weight for  $m/z$  and RT were 80 and 20, respectively. Finally, the resulting peak list was gap-filled with missing peaks using intensity tolerance of 10%,  $m/z$  tolerance of 20 ppm and RT tolerance of 0.5 min.

At the same time, using the MZmine’s setting of sample parameters, the samples were organized in groups, according to their taxonomy, geography, ecology, culture medium and the different strains. In the end of the above processing, the resulting data were converted into a MetaboAnalyst-CSV (comma-separated values) file, a text file that allows data to be saved in a table-structured format, taking into account the grouping parameters. As a result, five .csv files were generated, one for each of the grouping parameters to study their influence in the metabolic profile obtained. Each of these files was uploaded to the MetaboAnalyst 5.0 platform [22] to carry out the statistical analysis.

The MetaboAnalyst module of the One Factor-Statistical Analysis was used, and each file was uploaded as peak intensities data type and samples in columns (unpaired) format. Missing values were replaced by 1/5 of min positive values of their corresponding values and further feature filtering based on interquartile range (IQR) reduced the dataset to 2500 features. Normalization was performed using total area sums, Pareto scaling, and log transformation to transform the data matrix into a more Gaussian-type distribution and make individual features more comparable. Using the above parameters, PCA and PLS-DA analyses were generated to explore the metabolomics production. Within the PLS-DA analysis, the variables importance in projection (VIP) were studied, and cross validation and permutation test (separation distance (B/W), 100 permutation numbers) were performed.

Furthermore, the LC-MS/MS raw data of the 27 strains cultured in the RAM2-P V2 medium were processed by MZmine 2, employing the same parameters and noise level of 50 for the MS2 level. The processed data were used to visualize the production of the phocoenamycin by the peak intensities plot. Finally, after uploading the MetaboAnalyst file created and using the ecology of the strains as the grouping parameter, the HCA dendrogram was generated.

#### 4.5. High-Throughput Screening Assay of the Fermentation Extracts for Antimicrobial Activities

The crude extracts of the 270 fermentations were tested against the growth of methicillin-resistant *S. aureus* (MRSA) MB5393, *M. tuberculosis* H37Ra and *M. bovis* BCG. For the preparation of the inocula and the HTS assays, previously described methodologies were followed [51,52].

Briefly, 90  $\mu\text{L}$ /well of the diluted inocula were mixed with 10  $\mu\text{L}$ /well of extracts (final concentration of  $0.2 \times \text{WBE}$  of each extract in the assay). A dose–response curve of a reference standard compound (vancomycin for MRSA assay and streptomycin for both *Mycobacterium* assays) was included as internal assay plate control. The assays were performed in duplicate for each microorganism.

For MRSA assay, absorbance at OD 612 nm was measured at  $T_0$  (zero time) and immediately after that, plates were statically incubated at 37 °C for 20 h. After this period, the assay plates were shaken using the DPC Micromix-5 and once more the absorbance at OD 612 nm was measured at  $T_f$  (final time). For *Mycobacterium* assays, plates were incubated for 7 days at 5%  $\text{CO}_2$  and 95% humidity and 37 °C. After this incubation, 30  $\mu\text{L}$  of 0.02% resazurin and 15  $\mu\text{L}$  of Tween 20 were added to each well, incubated 24 h and assessed for color development. A change from blue to pink indicates reduction of resazurin and therefore bacterial growth. The wells were read for color change and the data were quantified by measuring fluorescence (excitation 570 nm, emission 615 nm). Both readouts absorbance and fluorescence were measured using an EnVision<sup>®</sup> multimode plate reader (Perkin Elmer, Waltham, MA, USA).

In order to process and analyze the data and calculate the RZ' factor (which predicts the robustness of an assay), the Genedata Screener software (Genedata, Inc., Basel, Switzerland) was employed. For an assay to be accepted, the RZ' factor must be close to 1 [53].

An extract was considered active when the percentage of inhibition was greater than 50%. In all experiments performed in this work, the RZ' factor obtained was between 0.85 and 0.95.

#### 4.6. Molecular Networking

##### 4.6.1. Classical Molecular Networking

The LC-MS/MS raw data were converted into mzXML files using MSConvert (ProteoWizard, proteowizard.sourceforge.net), uploaded to the Global Natural Products Social Molecular Networking (<http://gnps.ucsd.edu>, accessed on 2 December 2022) platform using WinSCP (<https://winscp.net/eng/index.php>, accessed on 2 December 2022) and then processed by GNPS to generate a MS/MS molecular network. The molecular network was created using the online workflow (<https://ccms-ucsd.github.io/GNPSDocumentation/>, accessed on 2 December 2022) on the GNPS website (<http://gnps.ucsd.edu>, accessed on 2 December 2022) [54]. The parameters are publicly available at: [https://gnps.ucsd.edu/ProteoSAFe/result.jsp?task=37e734c22f5f43cab7f8983d5121cea6&view=written\\_description](https://gnps.ucsd.edu/ProteoSAFe/result.jsp?task=37e734c22f5f43cab7f8983d5121cea6&view=written_description), accessed on 2 December 2022.

##### 4.6.2. Preprocessing by MZmine 2 and Feature-Based Molecular Networking

The data files previously converted to mzXML format were imported to MZmine v2.53 for preprocessing [21]. For each sample, the mass detection was set to a noise level of 100 for the MS1 level and 50 for MS2 levels. The chromatograms were built using the ADAP chromatogram builder [50] with a minimum group size set to 5 scans, group intensity threshold and minimum highest intensity set to 300 and  $m/z$  tolerance to 20 ppm. The

chromatograms were deconvoluted with the local minimum search algorithm (chromatographic threshold of 10%, search minimum in RT range of 0.1 min, minimum relative height of 10% and minimum absolute height of 100). Deisotoping of the chromatograms was achieved by the isotope peak grouper algorithm with  $m/z$  tolerance set to 20 ppm and RT tolerance to 0.5 min. All samples were combined in a peak list using the join aligner algorithm. The data were filtered in order to keep only peaks with minimum 2 peaks in a row and minimum 2 peaks in an isotope pattern. Lastly, the peak list was gap-filled with an intensity tolerance of 10%,  $m/z$  tolerance of 20 ppm and RT tolerance of 0.5 min. The final peak list was exported in two different format files (.csv and .mgf).

The resulting files were uploaded to GNPS for feature-based molecular networking (FBMN) analysis. A metadata file was also created and uploaded as a text file, describing the properties of each sample-strain (taxonomy, geography, ecology). A molecular network was created with the feature-based molecular networking (FBMN) workflow [44] on GNPS (<https://gnps.ucsd.edu>, accessed on 2 December 2022, [53] with the following parameters: The data were filtered by removing all MS/MS fragment ions within  $\pm 17$  Da of the precursor  $m/z$  and the MS/MS spectra were window filtered by choosing only the top 6 fragment ions in the  $\pm 50$  Da window throughout the spectrum. The precursor ion mass tolerance was set to 1 Da and the MS/MS fragment ion tolerance to 0.5 Da. In the molecular network, the edges had a cosine score above 0.6 and more than 6 matched peaks. Edges between two nodes were kept only if each of the nodes appeared in each other's respective top 10 most similar nodes and the maximum size of a molecular family was set to 100. Then, the spectra in the network were searched against GNPS spectral libraries [54,55] and the library spectra were filtered in the same manner as the data results. The molecular network generated was analyzed and visualized using Cytoscape software (version 3.8.2) [56]. The number of nodes and edges, as well as the average number of neighbors were generated using the Network Analyzer in Cytoscape software.

#### 4.7. Dereplication Process and Identification of Putative New Analogues

The LC-UV-HRMS analyses were carried out as previously reported [49]. The major peaks in each chromatogram were initially searched manually against the in-house MEDINA's LC-HRMS library. Then, the proposed molecular formula, the accurate mass and the UV profile were compared against the Dictionary of Natural Products (DNP), the COleCtion of Open Natural ProdUcTs (COCONUT), the Natural Product Activity & Species Source database (NPASS) and the ChemSpider database. Additional information was taken into account, such as the biological source and the structure of the compound, where available within these databases. The combination of this manual comparative search along with the metabolomic profiling and molecular networking allowed the putative dereplication of the main compounds produced within the 27 strains in the different growing conditions.

Furthermore, this manual dereplication tentatively detected various peaks of compounds with molecular formulae that were not found within the above-mentioned databases, therefore are putatively new metabolites related to the phocoenamycin, maklamycin or siderophores. These annotations were also based on GNPS, searching the compounds that were in each molecular cluster and combining all the data obtained.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/md20120769/s1>, Table S1: Composition of the culture media ingredients; Figure S1: Total Ion Chromatogram (TIC) of the crude extract from the strain CA-214658 grown in RAM2-P V2 medium; Figures S2–S4: ESI-TOF spectra of the three phocoenamycin; Figures S5 and S6: Boxplots and density plots showing the normalization results in MetaboAnalyst 5.0; Figure S7: PLS-DA 2D Score plots of the 270 extracts for the parameters (a) different strains, (b) taxonomy species, (c) geographic origin and (d) ecology; Figure S8: (a) Cross-validation test (CV) and (b) permutation test of the PLS-DA analysis of the 270 extracts grouped by culture medium; Figure S9: Classical molecular network (GNPS) generated from the extracts of the 27 strains cultured in RAM2-P V2 medium; Table S2. Average % inhibition of the 270 extracts against MRSA MB5393, *M. bovis* BCG and *M. tuberculosis* H37Ra.

**Author Contributions:** Conceptualization, M.K., D.O.-C., O.G. and F.R.; Data curation, M.K.; Formal analysis, M.K., D.O.-C., M.d.I.C., J.M. and F.V.; Funding acquisition, O.G. and F.R.; Investigation, M.K., C.G.H., D.O.-C., M.d.I.C. and J.M.; Methodology, M.K., C.G.H., D.O.-C., P.S. and J.M.; Supervision, F.V., O.G. and F.R.; Writing—original draft, M.K. and F.R.; Writing—review & editing, D.O.-C. and M.d.I.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has received funding from the European Union’s H2020 program under grant agreement No. 721421 (MarPipe).

**Data Availability Statement:** Data availability is indicated through links within the article or included as Supplementary Material.

**Acknowledgments:** The published results are part of the doctoral thesis of Maria Kokkini (Doctoral Programme in Pharmacy (B15.56.1), Doctoral School in Health Sciences (EDCS), University of Granada, University Campus of Cartuja, Prof. Vicente Callao Street, 3, 18011 Granada, Spain).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Adebisi, Y.A.; Alaran, A.J.; Okereke, M.; Oke, G.I.; Amos, O.A.; Olaoye, O.C.; Oladunjoye, I.; Olanrewaju, A.Y.; Ukor, N.A.; Lucero-Prisno, D.E. COVID-19 and Antimicrobial Resistance: A Review. *Infect. Dis. Res. Treat.* **2021**, *14*, 117863372110338. [[CrossRef](#)] [[PubMed](#)]
2. Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **2015**, *14*, 111–129. [[CrossRef](#)]
3. Newman, D.J.; Cragg, G.M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803. [[CrossRef](#)] [[PubMed](#)]
4. Atanasov, A.G.; Zotchev, S.B.; Dirsch, V.M.; Supuran, C.T. Natural products in drug discovery: Advances and opportunities. *Nat. Rev. Drug Discov.* **2021**, *20*, 200–216. [[CrossRef](#)] [[PubMed](#)]
5. Davison, E.K.; Brimble, M.A. Natural product derived privileged scaffolds in drug discovery. *Curr. Opin. Chem. Biol.* **2019**, *52*, 1–8. [[CrossRef](#)] [[PubMed](#)]
6. Genilloud, O. Natural products discovery and potential for new antibiotics. *Curr. Opin. Microbiol.* **2019**, *51*, 81–87. [[CrossRef](#)]
7. Niu, G.; Li, W. Next-Generation Drug Discovery to Combat Antimicrobial Resistance. *Trends Biochem. Sci.* **2019**, *44*, 961–972. [[CrossRef](#)]
8. De Mol, M.L.; Snoeck, N.; De Maeseeneire, S.L.; Soetaert, W.K. Hidden antibiotics: Where to uncover? *Biotechnol. Adv.* **2018**, *36*, 2201–2218. [[CrossRef](#)]
9. Mohana, C.; Rao, H.Y.; Rakshith, D.; Mithun, P.; Nuthan, B.R.; Satish, S. Omics based approach for biodiscovery of microbial natural products in antibiotic resistance era. *J. Genet. Eng. Biotechnol.* **2018**, *16*, 1–8. [[CrossRef](#)]
10. Abdel-Razek, A.S.; El-Naggar, M.E.; Allam, A.; Morsy, O.M.; Othman, S.I. Microbial Natural Products in Drug Discovery. *Processes* **2020**, *8*, 470. [[CrossRef](#)]
11. Quinn, R.A.; Nothias, L.-F.; Vining, O.; Meehan, M.; Esquenazi, E.; Dorrestein, P.C. Molecular Networking As a Drug Discovery, Drug Metabolism, and Precision Medicine Strategy. *Trends Pharmacol. Sci.* **2017**, *38*, 143–154. [[CrossRef](#)] [[PubMed](#)]
12. Lacoske, M.H.; Theodorakis, E.A. Spirotrionate Polyketides as Leads in Drug Discovery. *J. Nat. Prod.* **2015**, *78*, 562–575. [[CrossRef](#)] [[PubMed](#)]
13. Iorio, M.; Maffioli, S.I.; Gaspari, E.; Rossi, R.; Mauri, P.; Sosio, M.; Donadio, S. Chrolactomycins from the Actinomycete *Actinospica*. *J. Nat. Prod.* **2012**, *75*, 1991–1993. [[CrossRef](#)]
14. Braddock, A.A.; Theodorakis, E.A. Marine Spirotrionates: Biosynthetic Edifices That Inspire Drug Discovery. *Mar. Drugs* **2019**, *17*, 232. [[CrossRef](#)] [[PubMed](#)]
15. Pérez-Bonilla, M.; Oves-Costales, D.; de la Cruz, M.; Kokkini, M.; Martín, J.; Vicente, F.; Genilloud, O.; Reyes, F. Phocoenamycin B and C, New Antibacterial Spirotrionates Isolated from a Marine *Micromonospora* sp. *Mar. Drugs* **2018**, *16*, 95. [[CrossRef](#)] [[PubMed](#)]
16. Ochoa, J.L.; Sanchez, L.M.; Koo, B.-M.; Doherty, J.S.; Rajendram, M.; Huang, K.C.; Gross, C.A.; Lington, R.G. Marine Mammal Microbiota Yields Novel Antibiotic with Potent Activity Against *Clostridium difficile*. *ACS Infect. Dis.* **2018**, *4*, 59–67. [[CrossRef](#)] [[PubMed](#)]
17. Pérez-Victoria, I.; Martín, J.; Reyes, F. Combined LC/UV/MS and NMR Strategies for the Dereplication of Marine Natural Products. *Planta Med.* **2016**, *82*, 857–871. [[CrossRef](#)]
18. Demain, A.L.; Gómez-Ortiz, B.; Ruiz-Villafán, B.; Rodríguez-Sanoja, R.; Sánchez, S. Recent findings of molecules with anti-infective activity: Screening of non-conventional sources. *Curr. Opin. Pharmacol.* **2019**, *48*, 40–47. [[CrossRef](#)]
19. Ossai, J.; Khatabi, B.; Nybo, S.E.; Kharel, M.K. Renewed interests in the discovery of bioactive actinomycete metabolites driven by emerging technologies. *J. Appl. Microbiol.* **2022**, *132*, 59–77. [[CrossRef](#)]
20. 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](#)]

21. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform.* **2010**, *11*, 395. [[CrossRef](#)] [[PubMed](#)]
22. Pang, Z.; Chong, J.; Zhou, G.; de Lima Morais, D.A.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P.-É.; Li, S.; Xia, J. MetaboAnalyst 5.0: Narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* **2021**, *49*, W388–W396. [[CrossRef](#)] [[PubMed](#)]
23. Lever, J.; Krzywinski, M.; Altman, N. Points of Significance: Principal component analysis. *Nat. Methods* **2017**, *14*, 641–642. [[CrossRef](#)]
24. Worley, B. Multivariate Analysis in Metabolomics. *Curr. Metabolomics* **2012**, *1*, 92–107. [[CrossRef](#)]
25. Ruiz-Perez, D.; Guan, H.; Madhivanan, P.; Mathee, K.; Narasimhan, G. So you think you can PLS-DA? *BMC Bioinform.* **2020**, *21*, 2. [[CrossRef](#)]
26. Chen, Y.; Li, E.-M.; Xu, L.-Y. Guide to Metabolomics Analysis: A Bioinformatics Workflow. *Metabolites* **2022**, *12*, 357. [[CrossRef](#)]
27. Caesar, L.K.; Kvalheim, O.M.; Cech, N.B. Hierarchical cluster analysis of technical replicates to identify interferents in untargeted mass spectrometry metabolomics. *Anal. Chim. Acta* **2018**, *1021*, 69–77. [[CrossRef](#)]
28. Shamikh, Y.I.; El Shamy, A.A.; Gaber, Y.; Abdelmohsen, U.R.; Madkour, H.A.; Horn, H.; Hassan, H.M.; Elmaidomy, A.H.; Alkhalifah, D.H.M.; Hozzein, W.N. Actinomycetes from the Red Sea Sponge *Coscinoderma matheusi*: Isolation, Diversity, and Potential for Bioactive Compounds Discovery. *Microorganisms* **2020**, *8*, 783. [[CrossRef](#)]
29. Thamchaipenet, A. Maklamicin, an Antibacterial Polyketide from an Endophytic Micromonospora sp. *J. Nat. Prod.* **2011**, *74*, 670–674.
30. Holden, P.; Nair, L.S. Deferoxamine: An Angiogenic and Antioxidant Molecule for Tissue Regeneration. *Tissue Eng. Part B Rev.* **2019**, *25*, 461–470. [[CrossRef](#)]
31. Normant, V.; Josts, I.; Kuhn, L.; Perraud, Q.; Fritsch, S.; Hammann, P.; Mislin, G.L.A.; Tidow, H.; Schalk, I.J. Nocardamine-Dependent Iron Uptake in *Pseudomonas aeruginosa*: Exclusive Involvement of the FoxA Outer Membrane Transporter. *ACS Chem. Biol.* **2020**, *15*, 2741–2751. [[CrossRef](#)] [[PubMed](#)]
32. Wang, G.-Y.; Graziani, E.; Waters, B.; Pan, W.; Li, X.; McDermott, J.; Meurer, G.; Saxena, G.; Andersen, R.J.; Davies, J. Novel Natural Products from Soil DNA Libraries in a Streptomyces Host. *Org. Lett.* **2000**, *2*, 2401–2404. [[CrossRef](#)] [[PubMed](#)]
33. Iijima, M.; Someno, T.; Ishizuka, M.; Sawa, R.; Naganawa, H.; Takeuchi, T. IC202B and C, new siderophores with immunosuppressive activity produced by *Streptoalloteichus* sp. 1454-19. *J. Antibiot.* **1999**, *52*, 775–780. [[CrossRef](#)] [[PubMed](#)]
34. Al Shaer, D.; Al Musaimi, O.; de la Torre, B.G.; Albericio, F. Hydroxamate siderophores: Natural occurrence, chemical synthesis, iron binding affinity and use as Trojan horses against pathogens. *Eur. J. Med. Chem.* **2020**, *208*, 112791. [[CrossRef](#)] [[PubMed](#)]
35. Acquah, K.S.; Beukes, D.R.; Warner, D.F.; Meyers, P.R.; Sunassee, S.N.; Maglangit, F.; Deng, H.; Jaspars, M.; Gammon, D.W. Novel South African Rare Actinomycete *Kribbella speibonae* Strain SK5: A Prolific Producer of Hydroxamate Siderophores Including New Dehydroxylated Congeners. *Molecules* **2020**, *25*, 2979. [[CrossRef](#)] [[PubMed](#)]
36. Bellotti, D.; Remelli, M. Deferoxamine B: A Natural, Excellent and Versatile Metal Chelator. *Molecules* **2021**, *26*, 3255. [[CrossRef](#)]
37. Sun, M.-Y.; Ye, Y.; Xiao, L.; Rahman, K.; Xia, W.; Zhang, H. Daidzein: A review of pharmacological effects. *Afr. J. Tradit. Complement. Altern. Med.* **2016**, *13*, 117–132. [[CrossRef](#)]
38. Sharifi-Rad, J.; Quispe, C.; Imran, M.; Rauf, A.; Nadeem, M.; Gondal, T.A.; Ahmad, B.; Atif, M.; Mubarak, M.S.; Sytar, O.; et al. Genistein: An Integrative Overview of Its Mode of Action, Pharmacological Properties, and Health Benefits. *Oxidative Med. Cell. Longev.* **2021**, *2021*, 3268136. [[CrossRef](#)]
39. Wang, J.-F.; Liu, S.-S.; Song, Z.-Q.; Xu, T.-C.; Liu, C.-S.; Hou, Y.-G.; Huang, R.; Wu, S.-H. Naturally Occurring Flavonoids and Isoflavonoids and Their Microbial Transformation: A Review. *Molecules* **2020**, *25*, 5112. [[CrossRef](#)]
40. Ondrejčíková, P.; Šturdíková, M.; Hushegyi, A.; Švajdenka, E.; Markošová, K.; Čertík, M. Endophytic *Streptomyces* sp. AC35, a producer of bioactive isoflavone aglycones and antimycins. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 1333–1344. [[CrossRef](#)]
41. Hamamoto, T.; Seto, H.; Beppu, T. Leptomycins A and B, new antifungal antibiotics. II. Structure elucidation. *J. Antibiot.* **1983**, *36*, 646–650. [[CrossRef](#)] [[PubMed](#)]
42. Afoullouss, S.; Balsam, A.; Allcock, A.L.; Thomas, O.P. Optimization of LC-MS<sup>2</sup> Data Acquisition Parameters for Molecular Networking Applied to Marine Natural Products. *Metabolites* **2022**, *12*, 245. [[CrossRef](#)] [[PubMed](#)]
43. Aron, A.T.; Gentry, E.C.; McPhail, K.L.; Nothias, L.-F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J.M.; Sikora, N.; Vargas, F.; et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat. Protoc.* **2020**, *15*, 1954–1991. [[CrossRef](#)] [[PubMed](#)]
44. Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; et al. Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **2020**, *17*, 905–908. [[CrossRef](#)] [[PubMed](#)]
45. Jia, X.-Y.; Tian, Z.-H.; Shao, L.; Qu, X.-D.; Zhao, Q.-F.; Tang, J.; Tang, G.-L.; Liu, W. Genetic Characterization of the Chlorothricin Gene Cluster as a Model for Spirotetronate Antibiotic Biosynthesis. *Chem. Biol.* **2006**, *13*, 575–585. [[CrossRef](#)]
46. Daduang, R.; Kitani, S.; Sudoh, Y.; Pait, I.G.U.; Thamchaipenet, A.; Ikeda, H.; Igarashi, Y.; Nihira, T. 29-Deoxymaklamicin, a new maklamicin analogue produced by a genetically engineered strain of Micromonospora sp. NBRC 110955. *J. Biosci. Bioeng.* **2015**, *120*, 608–613. [[CrossRef](#)]
47. Wang, W.; Qiu, Z.; Tan, H.; Cao, L. Siderophore production by actinobacteria. *BioMetals* **2014**, *27*, 623–631. [[CrossRef](#)]

48. Lacret, R.; Pérez-Victoria, I.; Oves-Costales, D.; de la Cruz, M.; Domingo, E.; Martín, J.; Díaz, C.; Vicente, F.; Genilloud, O.; Reyes, F. MDN-0170, a New Napyradiomycin from *Streptomyces* sp. Strain CA-271078. *Mar. Drugs* **2016**, *14*, 188. [[CrossRef](#)]
49. Martín, J.; Crespo, G.; González-Menéndez, V.; Pérez-Moreno, G.; Sánchez-Carrasco, P.; Pérez-Victoria, I.; Ruiz-Pérez, L.M.; González-Pacanowska, D.; Vicente, F.; Genilloud, O.; et al. MDN-0104, an Antiplasmodial Betaine Lipid from *Heterospora chenopodii*. *J. Nat. Prod.* **2014**, *77*, 2118–2123. [[CrossRef](#)]
50. Myers, O.D.; Sumner, S.J.; Li, S.; Barnes, S.; Du, X. One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks. *Anal. Chem.* **2017**, *89*, 8696–8703. [[CrossRef](#)]
51. Audoin, C.; Bonhomme, D.; Ivanisevic, J.; De la Cruz, M.; Cautain, B.; Monteiro, M.C.; Reyes, F.; Rios, L.; Perez, T.; Thomas, O.P. Balibalosides, an Original Family of Glucosylated Sesterterpenes Produced by the Mediterranean Sponge *Oscarella balibaloi*. *Mar. Drugs* **2013**, *11*, 1477–1489. [[CrossRef](#)] [[PubMed](#)]
52. Lauritano, C.; Martin, J.; De La Cruz, M.; Reyes, F.; Romano, G.; Ianora, A. First identification of marine diatoms with anti-tuberculosis activity. *Sci. Rep.* **2018**, *8*, 2284. [[CrossRef](#)] [[PubMed](#)]
53. Zhang, J.-H.; Chung, T.D.Y.; Oldenburg, K.R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **1999**, *4*, 67–73. [[CrossRef](#)] [[PubMed](#)]
54. Wang, M.; Carver, J.J.; Phelan, V.V.; Sanchez, L.M.; Garg, N.; Peng, Y.; Nguyen, D.D.; Watrous, J.; Kapon, 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](#)]
55. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; et al. MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **2010**, *45*, 703–714. [[CrossRef](#)]
56. Ono, K.; Demchak, B.; Ideker, T. Cytoscape tools for the web age: D3. js and Cytoscape. js exporters version 2. *F1000Research* **2014**, *3*, 1–13. [[CrossRef](#)]