Article

The Pangenome of Gram-Negative Environmental Bacteria Hides a Promising Biotechnological Potential

Cláudia Covas [ID], Gonçalo Figueiredo, Margarida Gomes, Tiago Santos, Sónia Mendo [ID] and Tânia S. Caetano *[ID]

CESAM and Department of Biology, University of Aveiro, 3810-193 Aveiro, Portugal; claudia.covas@ua.pt (C.C.); henriques.figueiredo@ua.pt (G.F.); smendo@ua.pt (S.M.)
* Correspondence: tcaetano@ua.pt

Abstract: Secondary metabolites (SMs) from environmental bacteria offer viable solutions for various health and environmental challenges. Researchers are employing advanced bioinformatic tools to investigate less-explored microorganisms and unearth novel bioactive compounds. In this research area, our understanding of SMs from environmental Gram-negative bacteria lags behind that of its Gram-positive counterparts. In this regard, *Pedobacter* spp. have recently gained attention, not only for their role as plant growth promoters but also for their potential in producing antimicrobials. This study focuses on the genomic analysis of *Pedobacter* spp. to unveil the diversity of the SMs encoded in their genomes. Among the 41 genomes analyzed, a total of 233 biosynthetic gene clusters (BGCs) were identified, revealing the potential for the production of diverse SMs, including RiPPs (27%), terpenes (22%), hybrid SMs (17%), PKs (12%), NRPs (9%) and siderophores (6%). Overall, BGC distribution did not correlate with phylogenetic lineage and most of the BGCs showed no significant hits in the MIBiG database, emphasizing the uniqueness of the compounds that *Pedobacter* spp. can produce. Of all the species examined, *P. cryoconitis* and *P. lusitanus* stood out for having the highest number and diversity of BGCs. Focusing on their applicability and ecological functions, we investigated in greater detail the BGCs responsible for siderophore and terpenoid production in these species and their relatives. Our findings suggest that *P. cryoconitis* and *P. lusitanus* have the potential to produce novel mixtures of siderophores, involving bifunctional IucAC/AcD NIS synthetases, as well as carotenoids and squalene. This study highlights the biotechnological potential of *Pedobacter* spp. in medicine, agriculture and other industries, emphasizing the need for a continued exploration of its SMs and their applications.

Keywords: specialized metabolites; siderophore; carotenoids; squalene; agriculture; PGPB

1. Introduction

Environmental natural products produced by microorganisms are and have been the most promising source for the discovery of new drugs [1,2]. The interest in exploring less-studied microorganisms for the production of novel molecules with relevant bioactivities has gained special importance in recent years, benefiting from the new bioinformatic tools and "omics" technologies available. Genome mining approaches, for instance, have been a powerful tool for identifying genes involved in the production of secondary metabolites (SMs) in genomes of poorly studied bacterial groups [3,4]. Such strategies lead to the discovery of a plethora of microbial compounds with applications in medicine and are being increasingly applied to other fields such as agriculture, nutrition, cosmetics and biomaterials, among others [5–8].

A less explored group of bacteria is *Pedobacter* spp., comprising Gram-negative bacteria from the Bacteroidota phylum, often found in soils or as members of the rhizosphere [9,10]. Moreover, this genus are known as environmental superbugs because they are capable of coping with the vast majority of known antibiotics [11,12]. Some *Pedobacter* spp. display antimicrobial activity, but the characterization of their bioactive

compounds is currently limited to an antifungal chitinase and nonribosomal peptides known as (iso)pedopeptins [12–16]. Notably, pedopeptins inhibit the growth of some antibiotic-resistant bacteria, which are classified by the WHO as a great threat to human health [17]. Other studies suggest *Pedobacter* spp. as producers of novel lanthipeptides, which are RiPPs described as having a wide variety of biological activities, including antimicrobial, antiviral and antiallodynic [18–20]. This highlights the potential of antimicrobials derived from *Pedobacter* spp. for diverse applications in the fields of medicine and food, including biopharmaceuticals and biopreservatives [21,22]. However, we anticipate that the biotechnological scope of this genus is considerably wider, potentially encompassing agricultural biotechnology, as certain species have been identified as plant growth-promoting bacteria (PGPB) while exhibiting the ability to suppress plant root diseases [23–26].

Recognizing the biotechnological importance of the *Pedobacter* genus, we leveraged genomic analysis tools to reveal the array of secondary metabolites (SMs) encoded in its pangenome. These include the biosynthetic gene clusters (BGCs) responsible for the synthesis of terpenes (including carotenoids), nonribosomal peptides (NRPs), polyketides (PKs), siderophores, and ribosomally synthesized and post-translationally modified peptides (RiPPs), among other compounds.

2. Materials and Methods

2.1. Identification of BGCs

For all *Pedobacter* species genomes available in the National Center for Biotechnology Information (NCBI) database, a screening for candidate biosynthetic gene clusters (BGCs) was performed using the specialized metabolite identification pipeline antiSMASH version 5.0 [27]. *Pedobacter* species genome sequences were obtained from http://www.ncbi.nlm.nih.gov/genone/?term=pedobacter, accessed on 6 June 2019. The distribution of BGCs identified via the antiSMASH analysis among the *Pedobacter* species was visualized in a circular chord diagram, generated using Circos [28]. BiG-SCAPE version 20181005 (available at https://git.wageningenur.nl/medema-group/BiG-SCAPE/-/wikis/installation; accessed on 10 February 2019) was used locally to analyze the 233 BGCs as individual .gbk files that were downloaded from the antiSMASH database (10 March 2019) [29,30]. Phylogenetic trees provided by CORASON were generated during the BiG-SCAPE analysis. A BGC distribution analysis for the *Pedobacter* genus was performed using the ClustVis (https://biit.cs.ut.ee/clustvis/) tool, which allowed for the generation of the principal components analysis (PCA) plot and heatmaps [31].

2.2. Phylogenetic Tree Construction

A 16S rRNA gene tree was constructed in MEGA 7 [32] using the neighbor-joining (NJ) method [33]. The model applied was the Kimura 2-parameter with gamma distribution and invariant sites (G + I) since it was the model with the lowest BIC (Bayesian Information Criterion), determined via the model prediction option of MEGA 7. The *Pedobacter* species 16S rRNA gene sequences were retrieved from the GenBank database and aligned using the CLUSTALX algorithm [34]. The confidence level of the branches was determined using the bootstrap method with 1000 replications. To obtain the NIS synthetases phylogenetic tree, the protein sequences of the characterized proteins were obtained from the Uniprot database and aligned using MUSCLE [35] with the *P. cryoconitis*, *P. lusitanus*, *P. hartonius* and *P. himalayensis* NIS synthetases. Their evolutionary history was also inferred with MEGA 7 using the maximum likelihood method. A discrete gamma distribution was used to model the evolutionary rate differences among the sites (5 categories (+G, parameter = 1.5404)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 1.00% sites).

2.3. Analysis of Siderophore and Terpene BGCs

The sequence of all siderophores and terpenes BGCs analyzed via comparative genomics were retrieved directly from the antiSMASH database or taken directly from the Nucleotide database from the National Center for Biotechnology Information (NCBI) database [36]. The latter were identified following a microbial Blast analysis. The InterPro [37] and NCBI Conserved Domain [38] webtools were used to predict the protein domains and infer their functions. The putative genes involved in the synthesis of siderophores and terpenes were colored and viewed with the Artemis software tool [39]. A comparative genomics analysis was performed and the similarities between clusters were visualized with Easyfig. 2.1 with a tblastx analysis [40]. The final images were constructed with OmniGraffle Professional software 5.4.4.

3. Results and Discussion

3.1. Pedobacter spp. Encode the Production of a High Diversity of SMs

The potential of *Pedobacter* to produce SMs was evaluated by examining the genomes of *Pedobacter* spp. ($n = 41$) with antiSMASH [41]. This tool has gained widespread use in genome mining, allowing for the discovery of potential cryptic gene clusters that encode natural products that are not normally produced under laboratory conditions. For instance, it has been instrumental in uncovering various specialized metabolites in *Streptomyces* spp. and RiPPs in anaerobic bacteria, as well as shedding light on the unexplored metabolism of *Myxobacteria* [42–44]. Collectively, the genus *Pedobacter* harbored a total of 233 BGCs, with an average of 6 BGCs per genome (Figure 1a; Table S1). Most of the BGCs encode the potential to produce RiPPs (27%), followed by terpenes (22%), hybrid SMs (17%) and PKs (12%). Additionally, we identified BGCs responsible for the production of NRPs (9%), other SMs (7%) and siderophores (6%), but in lower percentages (Figure 1a; Table S1). The mean count of BGCs per genome in this study falls below the values reported in analogous research involving metabolically versatile Gram-positive bacteria, such as the genus *Rhodococcus* (mean count of 17) [45], or the Gram-negative genus as *Pseudomonas* (mean count of nine) [46]. It is worth noting that our study found a slightly higher average BCG count per genome compared to a broad analysis across 68 phyla, where the average was three [47].

At least one BGC was found in all genomes (except *P. arcticus*), and *P. cryoconitis* and *P. lusitanus* have the highest number of clusters (18 and 17, respectively) (Figure 2; Table S1). There was no correlation between the genome size and the number of the BGCs per genome (Figure S1). The distribution of the BGCs across the genus, categorized by predicted product, displayed considerable heterogeneity (Figure 1). Nevertheless, terpene-related BGCs were prevalent in species that had only a single BGC per genome (Figure 1b).

3.2. The Pattern of Encoded SMs Is Phylogenetically Independent

The relationship between the distribution of biosynthetic gene clusters (BGCs) and phylogenetic lineage was examined, yielding an indiscernible pattern (Figure 1b). For instance, the species *P. lusitanus*, *P. cryoconitis*, *P. hartonius* and *P. himalayensis*, despite their close relationship, exhibited a distinct number and type of BGC arrays (Figures 1b and 2). Focusing only on the total BGCs, *P. lusitanus*, *P. cryoconitis* and *P. caeni* are the species with the highest number, despite not being phylogenetically close (Figure 2). On the other hand, *P. ruber*, *P. namyangjuensis* and *P. africanus* occupy distant branches in the phylogenetic tree and encode the same type of SMs (Figures 1b and 2). The same was observed for *P. nanyangensis* and *P. tournemirensis* that, despite being distantly related, both possess only one terpene BGC. Also, independently of the evolutionary similarities, almost all species have the potential to produce terpenes.
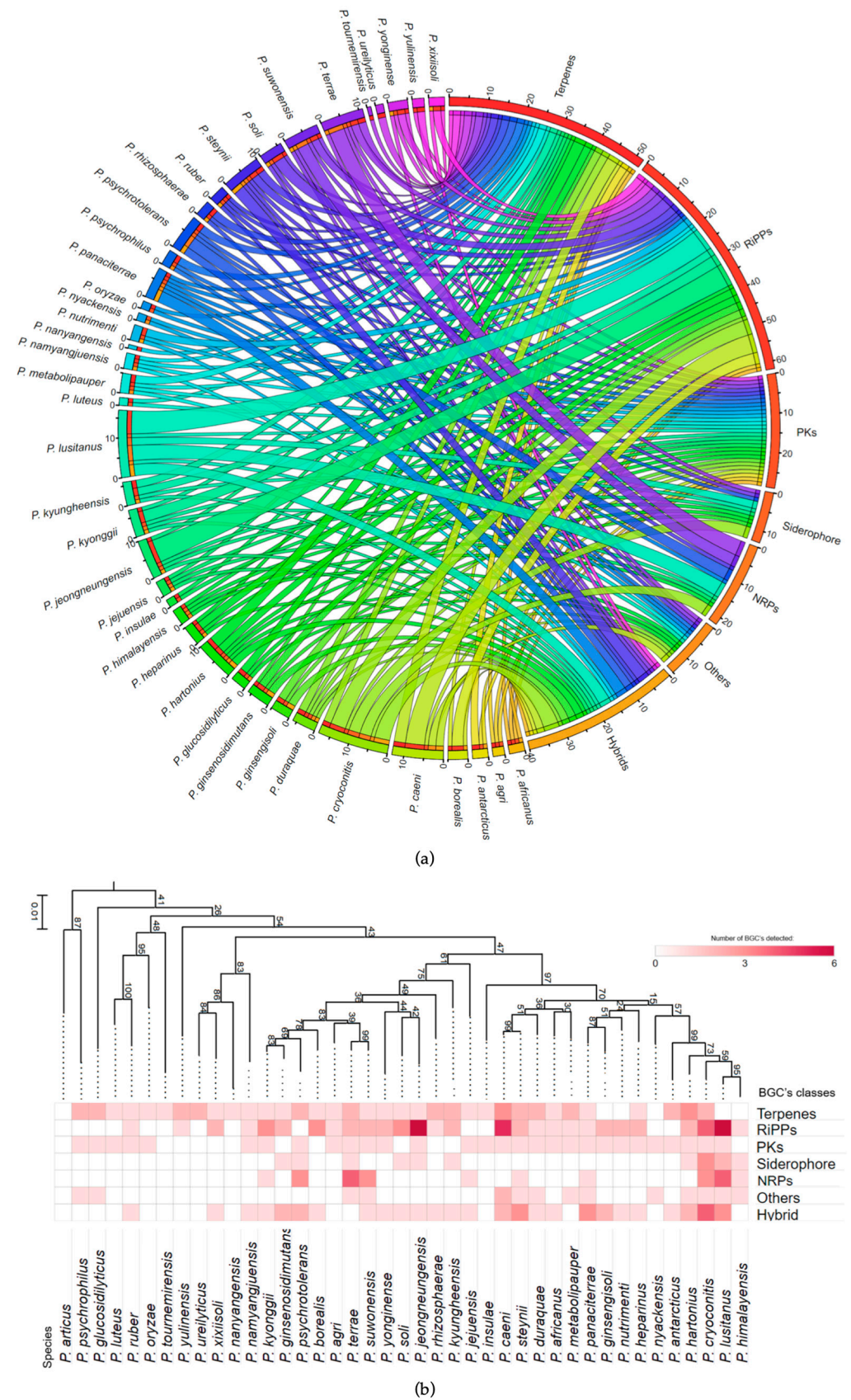
(a)

(b)

Figure 1. Representation of the various classes of BGCs identified in *Pedobacter* spp.; (**a**) their numbers according to the species and (**b**) according to their phylogeny and species. The neighbor-joining (NJ) tree (**b**) was based on a 16S rRNA gene sequence analysis and its scale corresponds to one nucleotide substitution per 100 nucleotides.
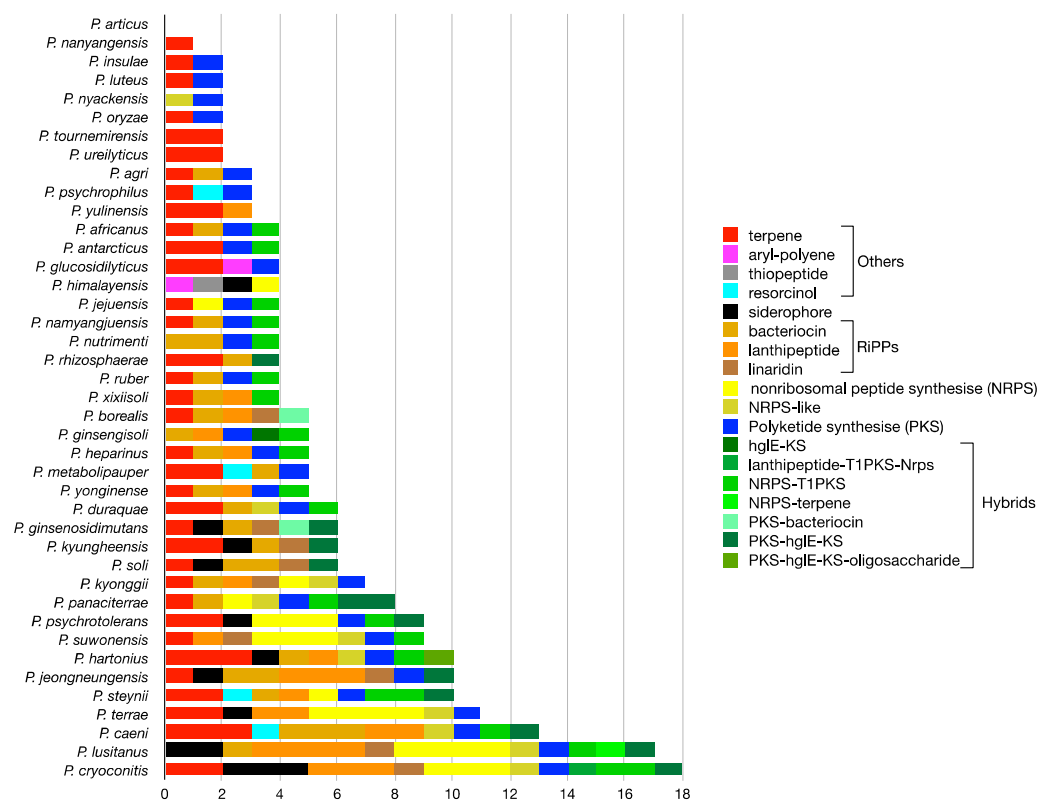
Figure 2. Occurrence of BGCs predicted using antiSMASH in the genomes of *Pedobacter* spp. according to their classes and sub-classes.

3.3. Most BGCs from Pedobacter spp. Do Not Share Homology with Each Other

To identify the orthologous relationship of the BGCs, sequence similarity networks were generated based on the BiG-SCAPE analysis and visualized with Cytoscape (Figure 3). Out of the 233 BGCs identified, 169 were considered singletons and 64 were integrated in 18 undirected networks, making a total of 187 gene cluster families (GCFs; Figure 3; Tables S1 and S2). These results show that most BGCs from *Pedobacter* spp. show little or no homology with each other, and a significant proportion of the GCFs (73%) are unique to individual species (singletons). An exception is the largest identified GCF, which includes 10 BGCs that encode terpenes (Figure 3; Table S2). However, GCFs for other SMs consisting of at least two BGCs were also found, except for NRPs, all of which were singletons (Figure 3). The GCFs including only two clusters are derived primarily from closely related species (Table S2). Therefore, despite a general lack of correlation between the quantities and types of BGCs, certain species (phylogenetically related or not) might have the potential to produce identical or very similar SMs. We also found that only two species (*P. agri* and *P. ginsenosidimutans*) have all their BGCs integrated in GCF networks (Tables S1 and S2). On the contrary, *P. cryoconitis*, *P. lusitanus*, and *P. caeni* emerged as the species with highest count of GCFs as singletons, followed by *P. hartonius*, *P. psychrotolerans* and *P. steynii* (Table S1).

3.4. Most of the BGCs from Pedobacter spp. Should Encode the Production of New Compounds

The perspective of *Pedobacter* spp. producing novel SMs becomes apparent, as only 23% of their BGCs have any level of similarity to the BGCs found in the MIBiG repository, which is a reference database for BGCs with known functions [29]. The highest similarity identified (84%) was between the aryl polyene (APE) BGC of *P. himalayensis* and *E. coli* (Figure S2). APEs share structural and functional similarities with carotenoid pigments, despite their distinct biosynthetic pathways. Notably, both compounds contribute to protecting bacteria from oxidative stress [48,49].

Also, in *P. himalayensis*, a cluster encoding a nonribosomal peptide synthetase (NRPS), which are large enzymatic complexes engaged in the production of NRPs, showed a 57% similarity to the amonabactin NRPS of *A. hydrophila* ATCC 7966 (Table S1; Figure S2). Additionally, this cluster also includes enterobactin-related genes. Both amonabactin and enterobactin function as siderophores. Enterobactin is the strongest siderophore known, binding to the ferric ion (Fe^{3+}) with a high affinity (Raines et al., 2016). We conducted an in silico analysis of this *P. himalayensis* NRPS and determined that its domains incorporate Ser and Dhb amino acids into the final peptide structure, which corresponds to enterobactin monomers [50]. Consequently, we hypothesize that *P. himalayensis* produces a enterobactin-like siderophore, probably in response to iron deficiency, as seen in *E. coli* (Hantash et al., 1997).

The third highest MIBiG similarity detected was between a terpene BGC of *P. oryzae* and a carotenoid BGC of *Algoriphagus* spp. (57%) [51]. Terpene BGCs of 22 other species, exhibited similarities to the same carotenoid BGC, although at percentages of less than 50% (Figure S2). This finding suggests that a considerable number of the terpene BGCs identified in this study may be linked to the biosynthesis of carotenoids, which are natural pigments that play diverse and essential roles in various biological systems. Next, eight siderophore BGCs showed a 50% similarity to a siderophore BGC responsible for the production of desferrioxamine E (Figure S2). These clusters encode enzymes belonging to the NRPS-independent siderophore (NIS) synthetases (conserved domain IucA/IucC) involved in the adenylation of a carboxylic acid substrate, typically citrate, or a derivative [35]. The BGCs encoding the production of these two types of SMs (siderophores and terpenes) were selected for an in-depth analysis, given their economic importance and promising biotechnological application, which will be discussed in more detail in the subsequent sections. The species selected for this analysis were *P. cryoconitis* and *P. lusitanus*, mainly due to their high number of BGCs, many of which did not yield matches with the MiBIG database. As these species are phylogenetically related to *P. hartonius* and *P. himalayensis*; these were also included in the analysis.
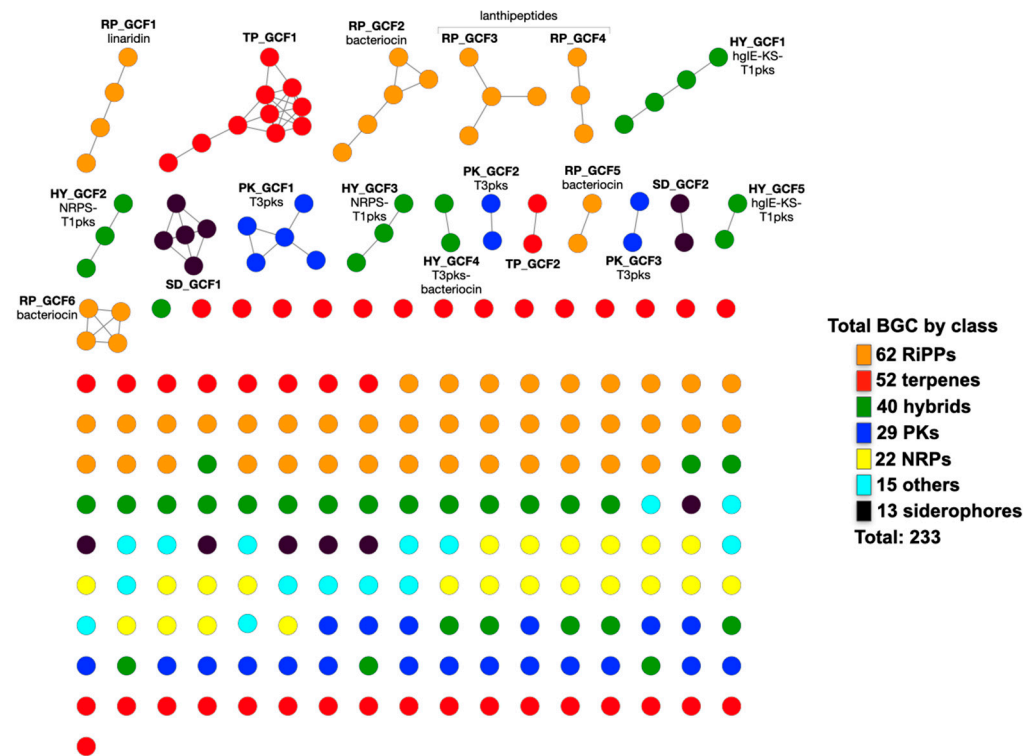

Figure 3. Overview of the sequence similarity network of 233 *Pedobacter* BGCs based on the BiG-SCAPE analysis output and visualized with Cytoscape [52]. Each node represents a BGC. On the right,

the number of BGCs per class is provided. GCFs with more than one node were provided with a designation, as described in Table S2. The prefixes HY, PK, RP, SD and TP represent hybrid, PK, RiPPs, siderophore and terpene, respectively.

3.5. The Siderophore BGCs of P. cryoconitis and P. lusitanus Encode Bifunctional IucAC/AcD NIS Synthetases

A total of seven siderophore BGCs encoding the NIS biosynthetic pathway were identified in *P. cryoconitis*, *P. lusitanus* and their closely related strains (Figure 4b). Of these, three BGCs were from *P. cryoconitis*, two from *P. lusitanus*, and *P. himalayensis* and *P. hartonius* have one each (Figure 2; Table S1). In an attempt to simplify the description, we tried to categorize the BGCs based on their similarity (Figure 2) and identified that only two BGCs from *P. cryoconitis* and *P. lusitanus* showed similarity, and, therefore, were grouped together (Figure 4b). Considering the core biosynthetic enzymes, it was observed that each BGC, except for one from *P. cryoconitis* (Group 5 in Figure 4b), contained two NIS synthetases (Figure 4b). All the synthetases were subjected to phylogenetic analysis (Figure 4a), revealing that the *P. hartonius* cluster exclusively contains type C' synthetases, whereas the two groups of *P. lusitanus* and *P. cryoconitis* (Group 2 and Group 3; Figure 4b), as well as the *P. himalayensis* BGC (Group 4; Figure 4b), harbor both type A and type C' synthetases.
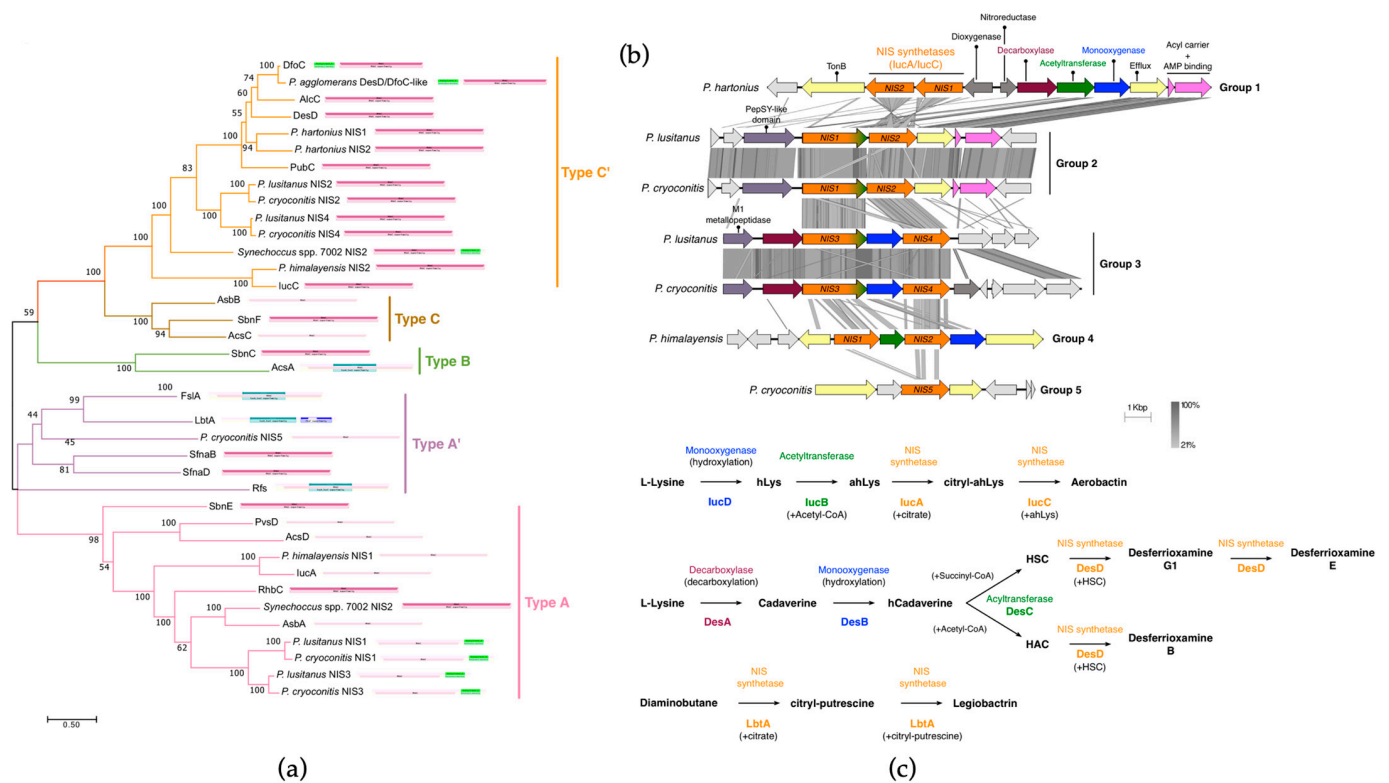

Figure 4. Neighbor-joining phylogenetic tree (**a**) of structurally characterized NIS synthetases and those found in the siderophore BGCs of *P. cryoconitis*, *P. lusitanus* and their close relatives [35]. The scale bar represents the number of amino acid substitutions per site. Tree branches were colored based on the NIS synthetases described by Carrol et al. (2018). Panel (**b**) presents the comparative genomic results of the siderophore BGCs via tBlastx analysis, where grey shading highlights regions of shared similarity, as indicated in the legend. Panel (**c**) offers a schematic representation of the biosynthetic pathway of known NISs, featuring some of the enzymes identified in the BGCs of *Pedobacter* spp.

Classical enzymes of type A and type C' include IucA and IucC, respectively. They play a pivotal role in the biosynthesis of aerobactin by sequentially coupling N^6-acetyl-N^6-hydroxylysine to the primary carboxylates of citrate (Figure 4c; [53]). Generally, type A

enzymes are responsible for the condensation of citric acid or its derivatives with mono-amines or amides, forming the core structure of siderophores. They display a particular affinity for citric acid and less complex amine or amide substrates. Type C' synthetases perform the condensation of these intricate building blocks and are often engaged in macro-cyclization reactions, leading to the production of larger and more diverse siderophores [35]. The BGC of *P. himalayensis* (Group 4; Figure 4) was identified to possess both types, with NIS1 and NIS2 following in the same clade as IucA and IucC, respectively. On the other hand, the *P. hartonius* type C' enzymes lacked the cognate type A enzymes and were in the same DfoC clade, despite not having an acetyltransferase (AcD) domain. DfoC-like synthetases feature an N-terminal AcD domain fused to the NIS domain (IucAC) and have been found in the genomes of Bacteroidota and Proteobacteria phyla [54,55]. Interestingly, the type A enzymes of Group 2 and Group 3 (NIS1 and NIS3 of *P. lusitanus* and *P. cryoconitis*) also possess these two domains, albeit in an inverse arrangement (N-terminal IucAC and a C-terminal AcD domains, herein referred to as IucAC/AcD). While this configuration of domains (IucAC/AcD) were previously detected in the siderophore BGCs of specific Cyanobacteria, namely, *Anabaena variabilis* and *Synechococcus* spp. [55], there are few studies focused on these synthetases so far. Finally, the NIS synthetase found in the G5 BGC of *P. cryoconitis* is classified as type A' (NIS5; Figure 4a), which is a type A subgroup with unique enantioselective substrates [35].

P. cryoconitis and *P. lusitanus* May Produce a Mixture of Unknown Siderophores

The siderophore BGC of *P. himalayensis* (Group 4 cluster; Figure 4b) showed a 22% homology with aerobactin BGC from the MIBiG database (Figure S2) and, as previously mentioned, its NIS synthetases are affiliated with IucA and IucC. This cluster also encodes two other enzymes required for aerobactin production, namely an acetyltransferase and a monooxygenase (Figure 4b,c), suggesting that this species may produce aerobactin or a very similar siderophore.

The siderophore BGC found in *P. hartonius* (G1; Figure 4b) represents one of the previously mentioned clusters with similarity to desferrioxamine E BGC (Figures 4 and S2). Apart from NIS synthetases, this cluster also contains genes encoding a decarboxylase, a monooxygenase and an acetyltransferase, which are the other enzymes required for desferrioxamine E production (Figure 4b,c). However, a closer analysis reveals additional genes encoding a dioxygenase and a nitroreductase (Figure 4b), suggesting that the siderophore produced by this species is distinct from desferrioxamine E after all. The two clusters of *P. lusitanus* and *P. cryoconitis* that form Group 3 also bear similarity to the desferrioxamine E BGC (Figure S2). These clusters feature a decarboxylase and a monooxygenase gene, while the bifunctional IucAC/AcD synthetase can account for the acetyltransferase function. Given the lack of characterized siderophores that are produced by these type of enzymes, it is likely that these clusters encode the production of a new type of siderophore instead of desferrioxamine E. Interestingly, similar BGCs were identified in the genomes of other *Pedobacter* spp., indicating that their products are widespread SMs within the genus (Figure S3).

The third siderophore BGC identified in *P. cryoconitis* (Group 5) contains an orphan type A' NIS synthetase flanked by export-related genes (Figure 4b). A cluster with analogous proteins is responsible for the biosynthesis of legiobactin by *Legionella pneumophila* [56,57] (Figure 4c). However, without characterization, it remains uncertain whether *P. cryoconitis* can produce a compound similar to this siderophore.

In summary, *P. lusitanus* and *P. cryoconitis* seem to be the species that produce more siderophores (at least two each), justifying further investigation into their structure and function. Given the habitat of both species, they hold potential for applications in bioremediation and agriculture biotechnology. Siderophores can be used for the treatment of metal-contaminated samples or as plant growth promoters to increase crop yield in the context of global climate change [58–61].

3.6. P. cryoconitis and P. lusitanus Terpenes: Carotenoids and Squalene

In the *P. cryoconitis* and *P. lusitanus* related species group, a total of six BGCs encoding the biosynthesis of terpenes were identified (Figure 5a). In total, three BGCs were identified in *P. hartonius*, two in *P. cryoconitis*, one *in P. lusitanus* and none in *P. himalayensis*. The *P. lusitanus* cluster was classified via antiSMASH as a NRPS-terpene hybrid due to the presence of an NRPS within the adjacent 10 kb region of the terpene cluster. However, the association between these biosynthetic elements might not necessarily indicate the production of a hybrid compound. The comparative genomics analysis of the terpene BGCs resulted in their classification into Group 1 (G1) and Group 2 (G2) clusters (Figure 5a).
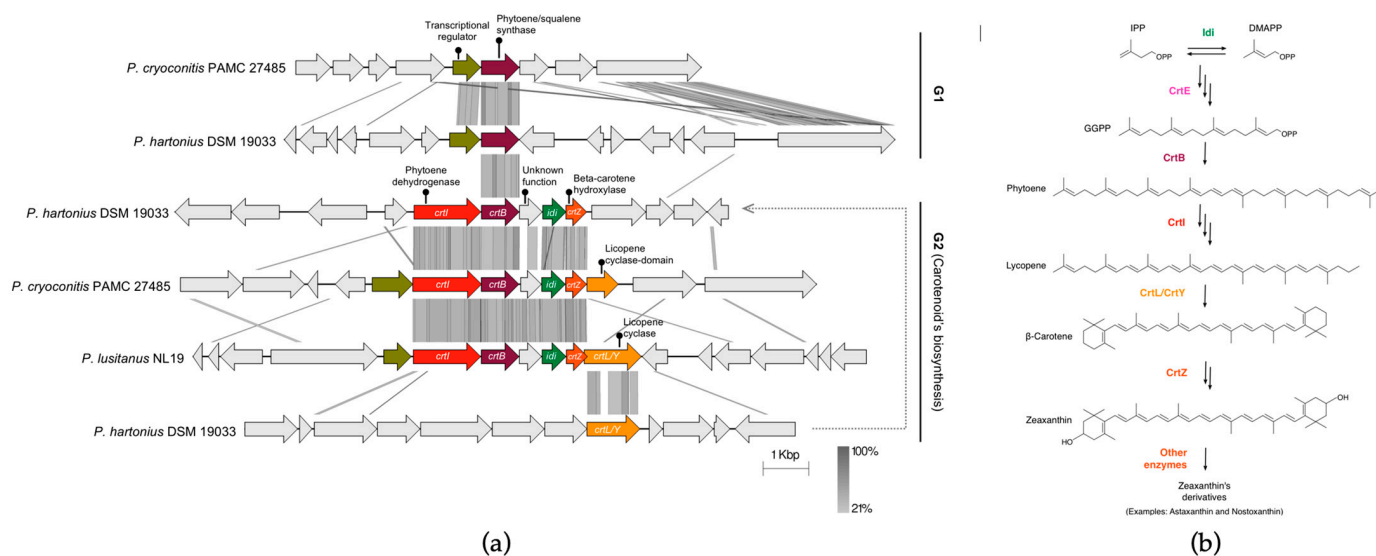

Figure 5. Comparative genomics of the six terpene BGCs found in *P. cryoconitis*, *P. lusitanus* and *P. hartonius* based on tBlastx analysis (**a**), where grey shadowing indicates regions of shared similarity according to the legend; (**b**) some biosynthetic pathways of squalene and known carotenoids, involving the production of beta-carotene, zeaxanthin and some derivatives.

The common traits among G1 clusters comprise a phytoene/squalene synthase and a transcriptional regulator gene. Typically, in bacteria, these orphan synthases correspond to squalene synthases (SQSs; Figure 5b), which participate in the biosynthesis of squalene, a precursor of triterpenoids and steroids [62]. Squalene possesses notable properties, such as antioxidant, anticancer and anti-inflammatory [62,63]. Although various organisms from the three domains of life produce squalene, deep-sea shark liver oil has predominantly served as the natural source of this compound for several years [62]. Thus, due to the increased demand, sustainable alternatives are being sought, such as harnessing the biomass of microorganisms. Thus, *P. hartonius* and *P. cryoconitis* may have biotechnological potential in the production of bioactive squalene.

In addition to phytoene/squalene synthases, the G2 clusters also include other essential genes for the carotenoid biosynthetic pathway (Figure 5b), notably the lycopene cyclase (*crtL/Y*), the phytoene dehydrogenase (*crtI*), the beta-carotene hydroxylase (*crtZ*) and the isopentenyl pyrophosphate isomerase (*idi*) gene (Figure 5a). Thus, the phytoene/squalene genes of these clusters were classified as *crtB* that encode the phytoene synthase, which converts geranylgeranyl pyrophosphate (GGPP) molecules into phytoene, the precursor of carotenoid biosynthesis [64] (Figure 5b). In the case of *P. hartonius*, the carotenogenesis genes were identified as separate BGCs via antiSMASH, as the *crtL/Y* gene was not found in close proximity to the other genes (Figure 5a). Carotenoids are pigments, and as such, the G2 clusters are likely the biosynthetic clusters responsible for the synthesis of the light/sand yellow pigment observed in colonies of *P. lusitanus*, *P. cryoconitis* and *P. hartonius* [65–67]. The demand for carotenoids in cosmetics and human healthcare has steadily increased,

prompting researchers and R&D companies to seek new alternatives [68–70]. Thus, these *Pedobacter* species and their BGCs deserve further investigation in this context.

4. Conclusions

The *Pedobacter* pangenome contains a substantial array of BGCs involved in the biosynthesis of compounds with promising biotechnological and commercial value, including RiPPS, terpenes, PKs, NRPs and siderophores. Additionally, some BGCs are expected to produce hybrid SMs. Most of the BGCs exhibit no significant similarities to their known counterparts in the MIBiG database, highlighting the biosynthetic diversity of *Pedobacter* spp. Our study revealed a plethora of SM clusters within the *Pedobacter* genomes, positioning this genus as a source of novel compounds ready for exploration, alongside other Gram-negative genera. Herein, *P. cryoconitis* and *P. lusitanus* emerged as excellent research models due to their abundant and diverse number of clusters that enable the production of various compounds such as siderophores, carotenoids and squalene, together with NRPs and lanthipeptides, all within a single strain. Given this profile, it would be interesting to harness the potential of these bacteria as plant probiotics, particularly considering the pressing challenge of promoting sustainable agriculture.

References

1. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79*, 629–661. [CrossRef] [PubMed]
2. van Santen, J.A.; Kautsar, S.A.; Medema, M.H.; Linington, R.G. Microbial Natural Product Databases: Moving Forward in the Multi-Omics Era. *Nat. Prod. Rep.* **2020**, *38*, 264–278. [CrossRef] [PubMed]
3. Adamek, M.; Spohn, M.; Stegmann, E.; Ziemert, N. Mining Bacterial Genomes for Secondary Metabolite Gene Clusters. In *Methods in Molecular Biology*; Humana Press Inc.: Totowa, NJ, USA, 2017; Volume 1520, pp. 23–47.
4. Ziemert, N.; Alanjary, M.; Weber, T. The Evolution of Genome Mining in Microbes-a Review. *Nat. Prod. Rep.* **2016**, *33*, 988–1005. [CrossRef] [PubMed]
5. Pham, J.V.; Yilma, M.A.; Feliz, A.; Majid, M.T.; Maffetone, N.; Walker, J.R.; Kim, E.; Cho, H.J.; Reynolds, J.M.; Song, M.C.; et al. A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Front. Microbiol.* **2019**, *10*, 1404. [CrossRef] [PubMed]
6. Shavandi, A.; Jalalvandi, E. Biofabrication of Bacterial Constructs: New Three-Dimensional Biomaterials. *Bioengineering* **2019**, *6*, 44. [CrossRef] [PubMed]
7. Gupta, P.L.; Rajput, M.; Oza, T.; Trivedi, U.; Sanghvi, G. Eminence of Microbial Products in Cosmetic Industry. *Nat. Prod. Bioprospect.* **2019**, *9*, 267–278. [CrossRef]

8. Hug, J.J.; Krug, D.; Müller, R. Bacteria as Genetically Programmable Producers of Bioactive Natural Products. *Nat. Rev. Chem.* **2020**, *4*, 172–193. [CrossRef]

9. Oren, A.; Garrity, G.M. Valid Publication of the Names of Forty-Two Phyla of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2021**, *71*, 005056. [CrossRef]

10. Steyn, P.L.; Segers, P.; Vancanneyt, M.; Sandra, P.; Kersters, K.; Joubert, J.J. Classification of Heparinolytic Bacteria into a New Genus, Pedobacter, Comprising Four Species: *Pedobacter heparinus* Comb. Nov., *Pedobacter piscium* Comb. Nov., *Pedobacter africanus* sp. Nov. and *Pedobacter saltans* sp. Nov. Proposal of the Family Sphingobac. *Int. J. Syst. Bacteriol.* **1998**, *48*, 165–177. [CrossRef]

11. Viana, A.T.; Caetano, T.; Covas, C.; Santos, T.; Mendo, S. Environmental Superbugs: The Case Study of *Pedobacter* spp. *Environ. Pollut.* **2018**, *241*, 1048–1055. [CrossRef]

12. Bjerketorp, J.; Levenfors, J.J.; Nord, C.; Guss, B.; Öberg, B.; Broberg, A. Selective Isolation of Multidrug-Resistant *Pedobacter* spp., Producers of Novel Antibacterial Peptides. *Front. Microbiol.* **2021**, *12*, 642829. [CrossRef] [PubMed]

13. Kozuma, S.; Hirota-Takahata, Y.; Fukuda, D.; Kuraya, N.; Nakajima, M.; Ando, O. Screening and Biological Activities of Pedopeptins, Novel Inhibitors of LPS Produced by Soil Bacteria. *J. Antibiot.* **2014**, *67*, 237–242. [CrossRef] [PubMed]

14. Covas, C.; Almeida, B.; Esteves, A.C.; Lourenço, J.; Domingues, P.; Caetano, T.; Mendo, S. Peptone from Casein, an Antagonist of Nonribosomal Peptide Synthesis: A Case Study of Pedopeptins Produced by Pedobacter Lusitanus NL19. *N. Biotechnol.* **2021**, *60*, 62–71. [CrossRef] [PubMed]

15. Song, Y.-S.; Seo, D.-J.; Jung, W.-J. Identification, Purification, and Expression Patterns of Chitinase from *Psychrotolerant pedobacter* sp. PR-M6 and Antifungal Activity in vitro. *Microb. Pathog.* **2017**, *107*, 62–68. [CrossRef] [PubMed]

16. Nord, C.; Bjerketorp, J.; Levenfors, J.J.; Cao, S.; Strömstedt, A.A.; Guss, B.; Larsson, R.; Hughes, D.; Öberg, B.; Broberg, A. Isopedopeptins A–H: Cationic Cyclic Lipodepsipeptides from Pedobacter Cryoconitis UP508 Targeting WHO Top-Priority Carbapenem-Resistant Bacteria. *ACS Chem. Biol.* **2020**, *15*, 2937–2944. [CrossRef]

17. Tacconelli, E.; Carrara, E.; Savoldi, A.; Harbarth, S.; Mendelson, M.; Monnet, D.L.; Pulcini, C.; Kahlmeter, G.; Kluytmans, J.; Carmeli, Y.; et al. Discovery, Research, and Development of New Antibiotics: The WHO Priority List of Antibiotic-Resistant Bacteria and Tuberculosis. *Lancet Infect. Dis.* **2018**, *18*, 318–327. [CrossRef]

18. Caetano, T.; van der Donk, W.; Mendo, S. Bacteroidetes Can Be a Rich Source of Novel Lanthipeptides: The Case Study of Pedobacter Lusitanus. *Microbiol. Res.* **2020**, *235*, 126441. [CrossRef]

19. Repka, L.M.; Chekan, J.R.; Nair, S.K.; van der Donk, W.A. Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes. *Chem. Rev.* **2017**, *117*, 5457–5520. [CrossRef]

20. Blockus, S.; Sake, S.M.; Wetzke, M.; Grethe, C.; Graalmann, T.; Pils, M.; Le Goffic, R.; Galloux, M.; Prochnow, H.; Rox, K.; et al. Labyrinthopeptins as Virolytic Inhibitors of Respiratory Syncytial Virus Cell Entry. *Antivir. Res.* **2020**, *177*, 104774. [CrossRef]

21. Liolios, K.; Sikorski, J.; Lu, M.; Nolan, M.; Lapidus, A.; Lucas, S.; Hammon, N.; Deshpande, S.; Cheng, J.F.; Tapia, R.; et al. Complete Genome Sequence of the Gliding, Heparinolytic *Pedobacter saltans* Type Strain (113 T). *Stand. Genom. Sci.* **2011**, *5*, 30–40. [CrossRef]

22. Shaya, D.; Tocilj, A.; Li, Y.; Myette, J.; Venkataraman, G.; Sasisekharan, R.; Cygler, M. Crystal Structure of Heparinase II from Pedobacter Heparinus and Its Complex with a Disaccharide Product. *J. Biol. Chem.* **2006**, *281*, 15525–15535. [CrossRef] [PubMed]

23. Lay, C.-Y.; Bell, T.H.; Hamel, C.; Harker, K.N.; Mohr, R.; Greer, C.W.; Yergeau, É.; St-Arnaud, M. Canola Root–Associated Microbiomes in the Canadian Prairies. *Front. Microbiol.* **2018**, *9*, 01188. [CrossRef] [PubMed]

24. Lee, S.A.; Kim, H.S.; Kim, T.-W.; Sang, M.K.; Song, J.; Weon, H.-Y. Complete Genome Sequence of Plant Growth-Promoting Bacterium *Pedobacter ginsengisoli* T01R-27 Isolated from Tomato (*Solanum lycopersicum* L.) Rhizosphere. *Korean J. Microbiol.* **2019**, *55*, 425–427. [CrossRef]

25. Morais, M.C.; Mucha, Â.; Ferreira, H.; Gonçalves, B.; Bacelar, E.; Marques, G. Comparative Study of Plant Growth-promoting Bacteria on the Physiology, Growth and Fruit Quality of Strawberry. *J. Sci. Food Agric.* **2019**, *99*, 5341–5349. [CrossRef]

26. Yin, C.; Casa Vargas, J.M.; Schlatter, D.C.; Hagerty, C.H.; Hulbert, S.H.; Paulitz, T.C. Rhizosphere Community Selection Reveals Bacteria Associated with Reduced Root Disease. *Microbiome* **2021**, *9*, 86. [CrossRef]

27. Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S.Y.; Medema, M.H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47*, W81–W87. [CrossRef]

28. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Jones, S.J.; Marra, M.A. Circos: An Information Aesthetic for Comparative Genomics. *Genome Res.* **2009**, *9*, 1639–1645. [CrossRef]

29. Navarro-Muñoz, J.C.; Selem-Mojica, N.; Mullowney, M.W.; Kautsar, S.; Tryon, J.H.; Parkinson, E.I.; Santos, E.L.C.D.L.; Yeong, M.; Cruz-Morales, P.; Abubucker, S.; et al. A Computational Framework for Systematic Exploration of Biosynthetic Diversity from Large-Scale Genomic Data. *bioRxiv* **2018**, 445270. [CrossRef]

30. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H.U.; Bruccoleri, R.; Lee, S.Y.; Fischbach, M.A.; Müller, R.; Wohlleben, W.; et al. AntiSMASH 3.0-A Comprehensive Resource for the Genome Mining of Biosynthetic Gene Clusters. *Nucleic Acids Res.* **2015**, *43*, W237–W243. [CrossRef]

31. Metsalu, T.; Vilo, J. ClustVis: A Web Tool for Visualizing Clustering of Multivariate Data Using Principal Component Analysis and Heatmap. *Nucleic Acids Res.* **2015**, *43*, W566–W570. [CrossRef]

32. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]

33. Saitou, N.; Nei, M. The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [PubMed]

34. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The CLUSTAL X Windows Interface: Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools. *Nucleic Acids Res.* **1997**, *25*, 4876–4882. [CrossRef]

35. Carroll, C.S.; Moore, M.M.; Carroll, C.S.; Moore, M.M.; Carroll, C.S.; Moore, M.M. Ironing out Siderophore Biosynthesis: A Review of Non-Ribosomal Peptide Synthetase (NRPS)-Independent Siderophore Synthetases. *Crit. Rev. Biochem. Mol. Biol.* **2018**, *53*, 356–381. [CrossRef] [PubMed]

36. Agarwala, R.; Barrett, T.; Beck, J.; Benson, D.A.; Bollin, C.; Bolton, E.; Bourexis, D.; Brister, J.R.; Bryant, S.H.; Canese, K.; et al. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2018**, *46*, D8–D13. [CrossRef]

37. Mitchell, A.L.; Attwood, T.K.; Babbitt, P.C.; Blum, M.; Bork, P.; Bridge, A.; Brown, S.D.; Chang, H.; El-gebali, S.; Fraser, M.I.; et al. InterPro in 2019: Improving Coverage, Classification and Access to Protein Sequence Annotations. *Nucleic Acids Res.* **2019**, *47*, 351–360. [CrossRef]

38. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The Conserved Domain Database in 2020. *Nucleic Acids Res.* **2019**, *48*, D265–D268. [CrossRef]

39. Carver, T.; Harris, S.R.; Berriman, M.; Parkhill, J.; McQuillan, J.A. Artemis: An Integrated Platform for Visualization and Analysis of High-Throughput Sequence-Based Experimental Data. *Bioinformatics* **2012**, *28*, 464–469. [CrossRef]

40. Sullivan, M.J.; Petty, N.K.; Beatson, S.A. Easyfig: A Genome Comparison Visualizer. *Bioinformatics* **2011**, *27*, 1009–1010. [CrossRef]

41. Medema, M.H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M.A.; Weber, T.; Takano, E.; Breitling, R. AntiSMASH: Rapid Identification, Annotation and Analysis of Secondary Metabolite Biosynthesis Gene Clusters in Bacterial and Fungal Genome Sequences. *Nucleic Acids Res.* **2011**, *39*, W339–W346. [CrossRef]

42. Letzel, A.C.; Pidot, S.J.; Hertweck, C. Genome Mining for Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs) in Anaerobic Bacteria. *BMC Genom.* **2014**, *15*, 983. [CrossRef]

43. Gregory, K.; Salvador, L.A.; Akbar, S.; Adaikpoh, B.I.; Cole Stevens, D. Survey of Biosynthetic Gene Clusters from Sequenced Myxobacteria Reveals Unexplored Biosynthetic Potential. *Microorganisms* **2019**, *7*, 181. [CrossRef] [PubMed]

44. Becerril, A.; Álvarez, S.; Braña, A.F.; Rico, S.; Díaz, M.; Santamaría, R.I.; Salas, J.A.; Méndez, C. Uncovering Production of Specialized Metabolites by Streptomyces Argillaceus: Activation of Cryptic Biosynthesis Gene Clusters Using Nutritional and Genetic Approaches. *PLoS ONE* **2018**, *13*, e0198145. [CrossRef] [PubMed]

45. Undabarrena, A.; Valencia, R.; Cumsille, A.; Zamora-Leiva, L.; Castro-Nallar, E.; Barona-Gomez, F.; Cámara, B. Rhodococcus Comparative Genomics Reveals a Phylogenomic-Dependent Non-Ribosomal Peptide Synthetase Distribution: Insights into Biosynthetic Gene Cluster Connection to an Orphan Metabolite. *Microb. Genom.* **2021**, *7*. [CrossRef]

46. Saati-Santamaría, Z.; Selem-Mojica, N.; Peral-Aranega, E.; Rivas, R.; García-Fraile, P. Unveiling the Genomic Potential of Pseudomonas Type Strains for Discovering New Natural Products. *Microb. Genom.* **2022**, *8*. [CrossRef]

47. Wei, B.; Du, A.; Zhou, Z.; Lai, C.; Yu, W.; Yu, J.; Yu, Y.; Chen, J.; Zhang, H.; Xu, X.; et al. An Atlas of Bacterial Secondary Metabolite Biosynthesis Gene Clusters. *Environ. Microbiol.* **2021**, *23*, 6981–6992. [CrossRef]

48. Schöner, T.A.; Gassel, S.; Osawa, A.; Tobias, N.J.; Okuno, Y.; Sakakibara, Y.; Shindo, K.; Sandmann, G.; Bode, H.B. Aryl Polyenes, a Highly Abundant Class of Bacterial Natural Products, Are Functionally Related to Antioxidative Carotenoids. *ChemBioChem* **2016**, *17*, 247–253. [CrossRef] [PubMed]

49. Cimermancic, P.; Medema, M.H.; Claesen, J.; Kurita, K.; Wieland Brown, L.C.; Mavrommatis, K.; Pati, A.; Godfrey, P.A.; Koehrsen, M.; Clardy, J.; et al. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158*, 412–421. [CrossRef]

50. Miller, B.R.; Drake, E.J.; Shi, C.; Aldrich, C.C.; Gulick, A.M. Structures of a Nonribosomal Peptide Synthetase Module Bound to MbtH-like Proteins Support a Highly Dynamic Domain Architecture. *J. Biol. Chem.* **2016**, *291*, 22559–22571. [CrossRef]

51. Tao, L.; Yao, H.; Kasai, H.; Misawa, N.; Cheng, Q. A Carotenoid Synthesis Gene Cluster from *Algoriphagus* sp. KK10202C with a Novel Fusion-Type Lycopene β-Cyclase Gene. *Mol. Genet. Genom.* **2006**, *276*, 79–86. [CrossRef]

52. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Resarch* **2003**, *13*, 2498–2504. [CrossRef] [PubMed]

53. Mydy, L.S.; Bailey, D.C.; Patel, K.D.; Rice, M.R.; Gulick, A.M. The Siderophore Synthetase IucA of the Aerobactin Biosynthetic Pathway Uses an Ordered Mechanism. *Biochemistry* **2020**, *59*, 2143–2153. [CrossRef] [PubMed]

54. Salomone-Stagni, M.; Bartho, J.D.; Polsinelli, I.; Bellini, D.; Walsh, M.A.; Demitri, N.; Benini, S. A Complete Structural Characterization of the Desferrioxamine E Biosynthetic Pathway from the Fire Blight Pathogen Erwinia Amylovora. *J. Struct. Biol.* **2018**, *202*, 236–249. [CrossRef] [PubMed]

55. Burrell, M.; Hanfrey, C.C.; Kinch, L.N.; Elliott, K.A.; Michael, A.J. Evolution of a Novel Lysine Decarboxylase in Siderophore Biosynthesis. *Mol. Microbiol.* **2012**, *86*, 485–499. [CrossRef] [PubMed]

56. Burnside, D.M.; Wu, Y.; Shafaie, S.; Cianciotto, N.P. The Legionella Pneumophila Siderophore Legiobactin Is a Polycarboxylate That Is Identical in Structure to Rhizoferrin. *Infect. Immun.* **2015**, *83*, 3937–3945. [CrossRef] [PubMed]

57. Allard, K.A.; Viswanathan, V.K.; Cianciotto, N.P. LbtA and LbtB Are Required for Production of the Legionella Pneumophila Siderophore Legiobactin. *J. Bacteriol.* **2006**, *188*, 1351–1363. [CrossRef] [PubMed]

58. Scavino, A.F.; Pedraza, R.O. The Role of Siderophores in Plant Growth-Promoting Bacteria. In *Bacteria in Agrobiology: Crop Productivity*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 265–285. ISBN 9783642372414.

59. Hesse, E.; O'Brien, S.; Tromas, N.; Bayer, F.; Luján, A.M.; Veen, E.M.; Hodgson, D.J.; Buckling, A. Ecological Selection of Siderophore-producing Microbial Taxa in Response to Heavy Metal Contamination. *Ecol. Lett.* **2018**, *21*, 117–127. [CrossRef]

60. Prema, P.; Selvarani, M. Microbial Siderophore as a Potent Biocontrol Agent for Plant Pathogens. *Int. J. Sci. Res.* **2012**, *2*, 521–523. [CrossRef]

61. Saha, M.; Sarkar, S.; Sarkar, B.; Sharma, B.K.; Bhattacharjee, S.; Tribedi, P. Microbial Siderophores and Their Potential Applications: A Review. *Environ. Sci. Pollut. Res.* **2016**, *23*, 3984–3999. [CrossRef]

62. Gohil, N.; Bhattacharjee, G.; Khambhati, K.; Braddick, D.; Singh, V. Engineering Strategies in Microorganisms for the Enhanced Production of Squalene: Advances, Challenges and Opportunities. *Front. Bioeng. Biotechnol.* **2019**, *7*, 50. [CrossRef]

63. Xu, W.; Ma, X.; Wang, Y. Production of Squalene by Microbes: An Update. *World J. Microbiol. Biotechnol.* **2016**, *32*, 195. [CrossRef] [PubMed]

64. Sandmann, G. New Functional Assignment of the Carotenogenic Genes CrtB and CrtE with Constructs of These Genes from Erwinia Species. *FEMS Microbiol. Lett.* **1992**, *90*, 253–257. [CrossRef]

65. Muurholm, S.; Cousin, S.; Pauker, O.; Brambilla, E.; Stackebrandt, E. *Pedobacter duraquae* sp. Nov., *Pedobacter westerhofensis* sp. Nov., *Pedobacter metabolipauper* sp. Nov., *Pedobacter hartonius* sp. Nov. and *Pedobacter steynii* sp. Nov., Isolated from a Hard-Water Rivulet. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 2221–2227. [CrossRef]

66. Margesin, R.; Spro, C.; Schumann, P.; Schinner, F. *Pedobacter cryoconitis* sp. Nov., a Facultative Psychrophile from Alpine Glacier Cryoconite. *Int. J. Syst. Evol. Microbiol.* **2003**, *53*, 1291–1296. [CrossRef] [PubMed]

67. Covas, C.; Caetano, T.; Cruz, A.; Santos, T.; Dias, L.; Klein, G.; Abdulmawjood, A.; Rodríguez-Alcalá, L.M.; Pimentel, L.L.; Gomes, A.; et al. *Pedobacter lusitanus* sp. Nov., Isolated from Sludge of a Deactivated Uranium Mine. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 1339–1348. [CrossRef] [PubMed]

68. Torregrosa-Crespo, J.; Montero, Z.; Fuentes, J.L.; García-Galbis, M.R.; Garbayo, I.; Vílchez, C.; Martínez-Espinosa, R.M. Exploring the Valuable Carotenoids for the Large-Scale Production by Marine Microorganisms. *Mar. Drugs* **2018**, *16*, 203. [CrossRef] [PubMed]

69. Bhatt, T.; Patel, K. Carotenoids: Potent to Prevent Diseases Review. *Nat. Prod. Bioprospect.* **2020**, *10*, 109–117. [CrossRef]

70. Zakynthinos, G.; Varzakas, T. Carotenoids: From Plants to Food Industry. *Curr. Res. Nutr. Food Sci.* **2016**, *4*, 38–51. [CrossRef]