



**HAL**  
open science

## Comparative Genomics among Closely Related Streptomyces Strains Revealed Specialized Metabolite Biosynthetic Gene Cluster Diversity

Cláudia M Vicente, Annabelle Thibessard, Jean-Noël Lorenzi, Mabrouka Benhadj, Laurence Hotel, Djamila Gacemi-Kirane, Olivier Lespinet, Pierre Leblond, Bertrand Aigle

► **To cite this version:**

Cláudia M Vicente, Annabelle Thibessard, Jean-Noël Lorenzi, Mabrouka Benhadj, Laurence Hotel, et al.. Comparative Genomics among Closely Related Streptomyces Strains Revealed Specialized Metabolite Biosynthetic Gene Cluster Diversity. *Antibiotics*, 2018, 7 (4), pp.86. 10.3390/antibiotics7040086 . hal-01887213

**HAL Id: hal-01887213**

**<https://hal.univ-lorraine.fr/hal-01887213>**

Submitted on 3 Oct 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

Communication

# Comparative Genomics among Closely Related *Streptomyces* Strains Revealed Specialized Metabolite Biosynthetic Gene Cluster Diversity

Cláudia M. Vicente <sup>1</sup>, Annabelle Thibessard <sup>1</sup>, Jean-Noël Lorenzi <sup>2</sup>, Mabrouka Benhadj <sup>1,3</sup> , Laurence Hôtel <sup>1</sup>, Djamila Gacemi-Kirane <sup>4</sup>, Olivier Lespinet <sup>2</sup>, Pierre Leblond <sup>1,\*</sup> and Bertrand Aigle <sup>1,\*</sup>

- <sup>1</sup> Université de Lorraine, INRA, DynAMic, F-54000 Nancy, France; claudia.morgado-vicente@univ-lorraine.fr (C.M.V.); annabelle.thibessard@univ-lorraine.fr (A.T.); mabrouka.benhadj@gmail.com (M.B.); laurence.hotel@univ-lorraine.fr (L.H.)
- <sup>2</sup> Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, University Paris-Sud, University Paris-Saclay, Gif-sur-Yvette CEDEX, France; jean-noel.lorenzi@i2bc.paris-saclay.fr (J.-N.L.); olivier.lespinet@i2bc.paris-saclay.fr (O.L.)
- <sup>3</sup> Biomolécules and Application Laboratory, Faculty of Exact Sciences and Natural and Life Sciences, University of Tebessa, Tebessa 12002, Algeria
- <sup>4</sup> Department of Biochemistry, Faculty of Science, University Badji Mokhtar Annaba, Annaba 23000, Algeria; dj\_gacemi@yahoo.fr
- \* Correspondence: pierre.leblond@univ-lorraine.fr (P.L.); bertrand.aigle@univ-lorraine.fr (B.A.); Tel.: +33-372-745-143 (P.L.); +33-372-745-141 (B.A.)

Received: 12 September 2018; Accepted: 28 September 2018; Published: 2 October 2018



**Abstract:** Specialized metabolites are of great interest due to their possible industrial and clinical applications. The increasing number of antimicrobial resistant infectious agents is a major health threat and therefore, the discovery of chemical diversity and new antimicrobials is crucial. Extensive genomic data from *Streptomyces* spp. confirm their production potential and great importance. Genome sequencing of the same species strains indicates that specialized metabolite biosynthetic gene cluster (SMBGC) diversity is not exhausted, and instead, a pool of novel specialized metabolites still exists. Here, we analyze the genome sequence data from six phylogenetically close *Streptomyces* strains. The results reveal that the closer strains are phylogenetically, the number of shared gene clusters is higher. Eight specialized metabolites comprise the core metabolome, although some strains have only six core gene clusters. The number of conserved gene clusters common between the isolated strains and their closest phylogenetic counterparts varies from nine to 23 SMBGCs. However, the analysis of these phylogenetic relationships is not affected by the acquisition of gene clusters, probably by horizontal gene transfer events, as each strain also harbors strain-specific SMBGCs. Between one and 15 strain-specific gene clusters were identified, of which up to six gene clusters in a single strain are unknown and have no identifiable orthologs in other species, attesting to the existing SMBGC novelty at the strain level.

**Keywords:** *Streptomyces*; strain; specialized metabolites; biosynthetic gene cluster

## 1. Introduction

Microorganisms are unmatched in their capacity to produce chemically diverse specialized metabolites [1]. *Streptomyces* in particular, which are ubiquitous filamentous bacteria with a differentiating mycelial life cycle, have a significant role in biotechnology industry. They are employed in the production of more than half of all antibiotics used in human and veterinary medicine,

as well as a large assortment of other high-value molecules such as immune suppressants, anticancer, and anti-parasitic molecules [2,3]. The increasing incidence of antibiotic resistant infections and the lack of new antimicrobial agents have driven the race for natural product discovery. The significant advances made in high-throughput DNA sequencing, along with the relatively low cost of whole genome sequencing, have greatly contributed to this effort [4]. Strategies to address the antimicrobial shortage include both the intensive screening of uncultivable microorganisms through metagenomics and environmental DNA cloning for heterologous expression [5,6], as well as further genome mining of cultivable isolates. Consequently, correlation data between species and specialized metabolomes are accumulating progressively.

*Streptomyces* spp. are amongst the best-characterized bacteria, and the vastly increasing genomic data that is available has revealed the presence of thousands of specialized metabolite biosynthetic gene clusters (SMBGCs), highlighting their enormous potential for natural product biosynthesis [7,8]. Estimates say that no more than 10% have been identified to date, and furthermore, the majority of these SMBGCs appear to be “silent” [9]. A recent bioinformatic analysis of six *Streptomyces albus* strains led to the identification of strain-specific SMBGCs, speaking to the chemical diversity at the strain level [10]. It has also been shown that *Streptomyces* species with identical 16S rRNA sequences can have distinct specialized metabolomes [11]. Moreover, studies in actinobacteria *Salinispora* demonstrated an important biosynthetic pathway diversity in marine populations, providing insight into new chemical diversity-generating mechanisms [12].

A *Streptomyces* strain collection that included isolates with interesting biological activities against both Gram-positive and Gram-negative indicator strains was recently characterized [13]. Here, we report the genomic analysis of a small group of these environmental *Streptomyces* strains, including those antimicrobial-producing isolates and the associated SMBGC diversity that was encountered. This analysis highlights that the use of closely related strains to identify new SMBGCs in *Streptomyces* constitutes a promising approach for the identification of novel specialized biosynthetic pathways.

## 2. Results and Discussion

### 2.1. Analyzed Strains are Phylogenetically Diverse

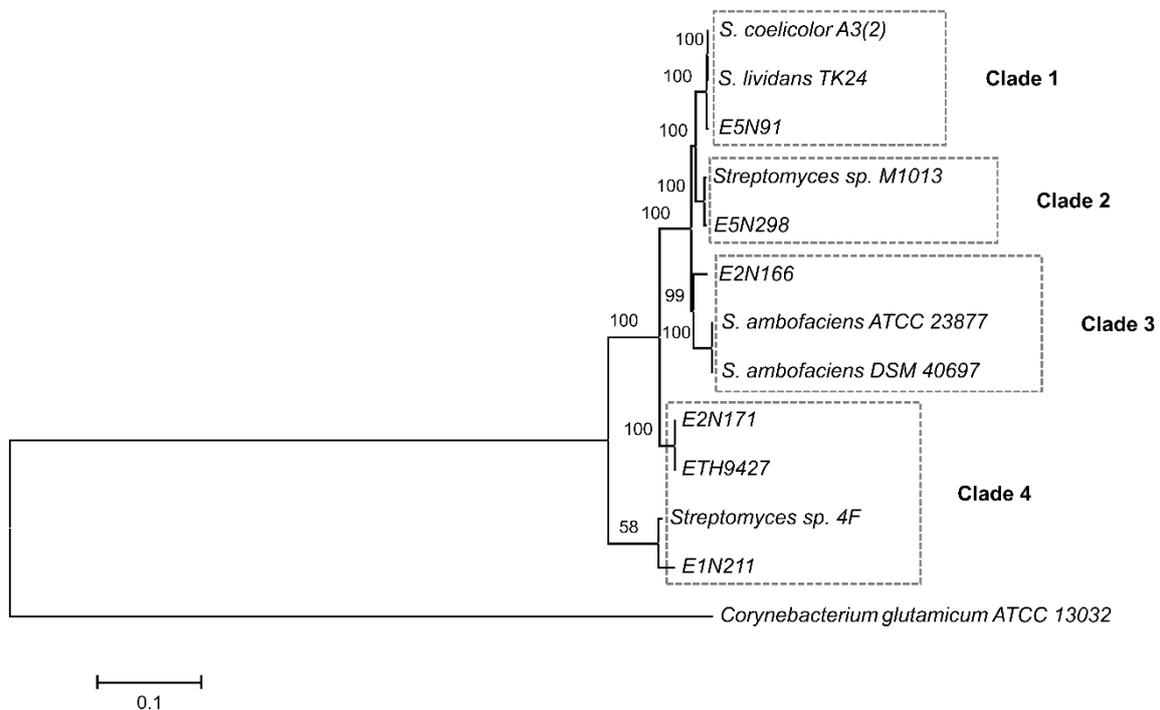
The genomes of five environmental strains from a previously constructed *Streptomyces* collection [13], isolates E5N91, E2N166, E2N171, E1N211, E5N298, were sequenced and examined. Draft genome sequences have 7.0–8.5 Mb with an average G+C content of 71.8%, which is in good accordance with the genus (Table 1).

**Table 1.** Strains analyzed in this work.

Strain	G + C Content (%)	Genome Size (Mb)
E5N91	71.9	8.51
E2N166	70.2	7.91
E2N171	72.3	7.00
E1N211	72.1	7.32
E5N298	71.9	7.87
<i>Streptomyces</i> sp. ETH9427	72.1	7.75

A comparative genomics analysis of the nucleotide sequences, including the taxonomically close *Streptomyces* sp. ETH9427 recently sequenced [14], was carried out through a multilocus sequence typing (MLST) using five genes: *atpD* (ATP synthase,  $\beta$  subunit), *gyrB* (DNA gyrase, subunit B), *recA* (recombination protein), *rpoB* (RNA polymerase,  $\beta$  subunit), and *trpB* (tryptophan synthase,  $\beta$  subunit). It highlighted the taxonomic relationships between the strains and with other model *Streptomyces* species (Figure 1). Four clades are distinguishable in the constructed phylogenetic tree: the first clade with strain E5N91, *Streptomyces coelicolor* A3(2) [7], and *S. lividans* TK24 [15], the second containing strain E5N298 and the recently described *Streptomyces* sp. M1013 [16], clade 3, including strain E2N166

and *S. ambofaciens* strains ATCC 23877 [17] and DSM 40697 [18], and finally strains E2N171, E1N211, and ETH9427 grouped with *Streptomyces* sp. 4F (accession number CP013142.1) in clade 4 (Figure 1).



**Figure 1.** Phylogenetic tree of sequenced isolated and reference strains. Constructed based on the concatenated sequence alignment of five loci (*atpB-gyrB-recA-rpoB-trpB*), using inference by neighbor-joining with 100 bootstrap replicates. Only positions with a minimum 92% site coverage were used and the final dataset has 9141 positions in total. The scale bar indicates 10% estimated sequence divergence. The tree is rooted on *Corynebacterium glutamicum* ATCC 13032.

Taxonomic affiliations were further assessed by calculating the average nucleotide identity (ANI) using JSpeciesWS [19], and are resumed in Table 2. The ANI in clade 1 between strain E5N91 and *S. coelicolor* and with *S. lividans* is 97.32% (with 75.5% of the genome nucleotides aligned) and 97.65% (77.0% of nucleotides aligned), respectively. In clade 2, the ANI value is 98.19% (80.9% nucleotides aligned) between strain E5N298 and *Streptomyces* sp. M1013, and in clade 3 between strain E2N166, *S. ambofaciens* ATCC 23877, and *S. ambofaciens* DSM 40697, the ANI values are 88.86% (65.3% alignment) and 88.75% (with 66.0% nucleotides aligned), respectively. The values in clade 4 are as follows: the E2N171 strain's ANI value is 98.98% with strain E1N211 (83.5% of aligned nucleotides), 99.14% (84.8% alignment) with strain ETH9427, and 94.56% (76.3% of nucleotides aligned) with *Streptomyces* sp. 4F; the ANI value of strain E1N211 is 99.90% (86.4% of aligned nucleotides) with strain ETH9427, and 94.41% (71.1% alignment) with *Streptomyces* sp. 4F; and the ETH9427 strain's ANI value is 94.13% (70.7% of nucleotides aligned) with *Streptomyces* sp. 4F. These results suggest that strain E5N91 belongs to a species that is taxonomically closely related to *S. coelicolor* and *S. lividans* in clade 1, as do strain E5N298 and *Streptomyces* sp. M1013 in clade 2. Moreover, the ANI value among the six studied strains varies between 81.40–99.90%, except in those from clade 4, where the ANI value is 99.00–99.90%, well above the species boundary [20] and indicating that strains E2N171, E1N211, and ETH9427 belong to the same species. The closest relative to these strains is *Streptomyces* sp. 4F, and as such, all are clustered in clade 4. However, these strains and *Streptomyces* sp. 4F belong to distinct but related species, as seen by the ANI values between 94.13–94.56% and a low branch bootstrap in the phylogenetic tree (Figure 1). Still, these strains have highly similar 16S ribosomal DNA sequences (dissimilarities between 0.1–1.1%), which would mistakenly suggest that most of the strains belong to the same species. These results indicate that 16S single-gene analysis is not the most accurate tool for streptomycetes

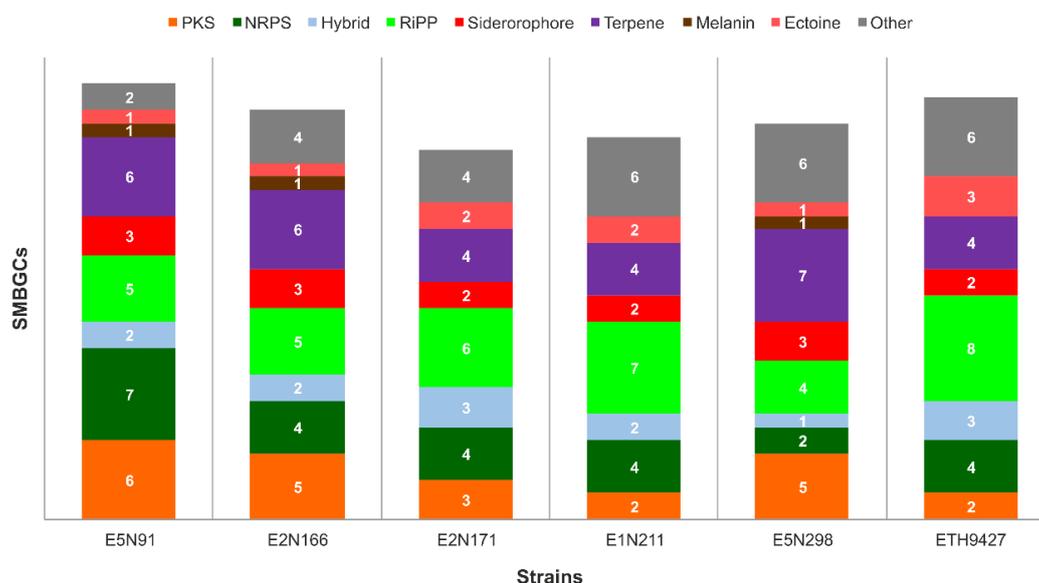
phylogeny analysis and inferring intraspecific genetic diversity, as already seen previously [21,22]. Instead, taxonomic relationships are better resolved using genome comparison indexes, such as ANI, and multi-gene analysis, such as MLST [23,24], which in some cases can even match whole-genome comparisons [25].

**Table 2.** Average nucleotide identity values of the compared strains.

	SLI	E5N91	M1013	E5N298	ATCC	DSM	E2N166	E2N171	E1N211	ETH9427	4F
SCO	99.0	97.3	91.5	91.3	86.9	86.9	87.5	81.7	81.8	81.9	81.8
SLI		97.7	91.8	91.6	87.0	87.0	87.7	81.6	81.7	81.8	81.7
E5N91			91.2	91.0	86.6	86.6	87.4	81.4	81.4	81.6	81.3
M1013				98.2	87.1	87.1	87.9	81.5	81.5	81.7	81.4
E5N298					87.2	87.2	87.9	81.7	81.7	81.8	81.6
ATCC						99.0	88.9	81.8	81.7	81.8	81.9
DSM							88.8	81.7	81.6	81.7	81.9
E2N166								82.1	82.1	82.2	82.0
E2N171									99.0	99.1	94.6
E1N211										99.9	94.4
ETH9427											94.1

## 2.2. Specialized Metabolism Diversity

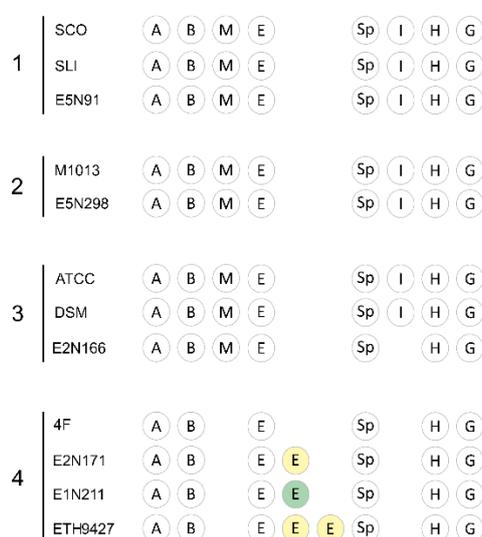
The identification of SMBGCs was performed using antiSMASH 3.0 [26] and then edited to take into account previously published experimental data. A total of 28–33 SMBGCs was predicted in the isolated strains (Figure 2), with strain E2N171 harboring the least (28) and strain E5N91 harboring the most (33) gene clusters, expectantly coinciding with the smallest and largest genome, respectively. All six strains harbor SMBGCs belonging to the principal classes such as polyketide synthases (PKS), non-ribosomal synthetases (NRPS), and PKS–NRPS hybrids. Interestingly, the analysis revealed that a single strain within its clade can contain up to 48% strain-specific SMBGCs, suggesting that strain-level search for new SMBGCs diversity is not exhausted, and genome sequencing of several strains from a single species can lead to the identification of new gene clusters, as already seen previously [10,12].



**Figure 2.** Specialized metabolite biosynthetic gene clusters identified in analyzed strains. Genome sequences were analyzed with antiSMASH 3.0 for gene cluster identification. The “Hybrid” category refers only to PKS–NRPS hybrid gene clusters. Other hybrid clusters that were detected were separated into their constituent parts to better visualize cluster diversity. Clusters encoding bacteriocins, lantipeptides, and lassopeptides are grouped in the category “RiPP”, and the category “Other” includes other KS pathways and less common clusters such as phenazines and butyrolactones, among others.

### 2.3. The Core and Conserved Specialized Metabolites

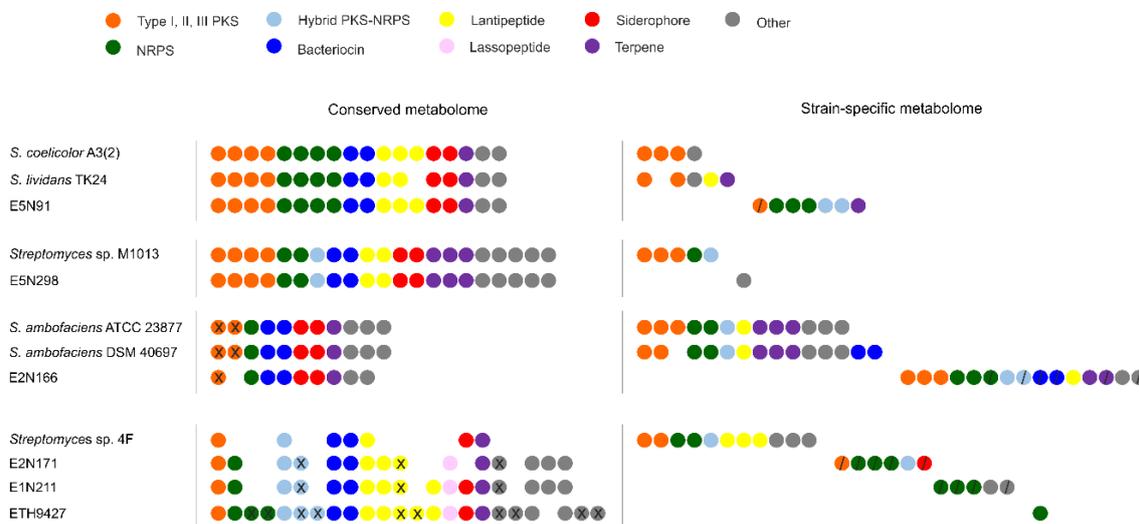
Some of the identified gene clusters are well known and usually found in *Streptomyces* spp. The core SMBGCs are gene clusters that are present in the majority of streptomycetes and in the examined strains were selected based on conservation, with at least 58% of the strains sharing all of the core SMBGCs that encode for: melanin [27], ectoine osmolytes [28], spore pigment [29], the siderophore desferrioxamine B [30], the terpenes albaflavenone [31], hopene [32], isorenieratene [33], and the volatile geosmin [34]. This set of eight biosynthetic gene clusters is present in all of the strains from clades 1, 2, and 3, although no isorenieratene gene cluster was found in strain E2N166 (Figure 3). Interestingly, none of the strains in clade 4 possess either melanin or isorenieratene encoding gene clusters. Furthermore, strains E2N171, E1N211, and ETH9427 have additional ectoine gene clusters: two in ETH9427, and one in each of the other two strains. These have different gene contents (Figure 3, yellow and green E clusters), and as expected, the phylogenetically closer E2N171 and ETH9427 strains have the same extra ectoine encoding gene cluster, that in the case of ETH9427 is doubled due to its location on the chromosomal terminal inverted repeats. This result is consistent with strains from the same species, which usually share the same core metabolome [35].



**Figure 3.** The core specialized metabolites produced by the strains. From the total number of SMBGCs that were identified, the core metabolome constitutes a set of gene clusters frequently conserved across *Streptomyces* species. The letters stand for the specialized metabolites albaflavenone (A), desferrioxamine B (B), melanin (M), ectoine (E), spore pigment (Sp), isorenieratene (I), hopene (H), and geosmin (G). Cluster disposition is not related to chromosomal position. The ectoine clusters in yellow and green represent different gene clusters.

In addition to the core metabolome, other SMBGCs are conserved within each clade (Figure 4). Strain E5N91 contains 18 conserved gene clusters out of 33 in total, which are also present in *S. coelicolor* in clade 1. Surprisingly, biosynthetic gene clusters for the characteristic metabolites coelichelin [36], coelibactin [37], calcium-dependent antibiotic [38], and coelimycin [39] are conserved in strain E5N91, whereas the actinorhodin [40] and undecylprodigiosin [41] gene clusters are not, even if they are present in phylogenetically closely related species. In clade 2, two-thirds of the gene clusters identified in strain E5N298 are conserved in *Streptomyces* sp. M1013 (21 out of 30 in total). Conversely, in clade 3, strain E2N166 has only nine gene clusters out of 31 in total conserved with *S. ambofaciens* strains ATCC 23877 and DSM 40697. These include the species-distinctive Type II PKS gene cluster encoding kinamycin [42], the NRPS biosynthetic gene cluster for coelichelin, two siderophore and two bacteriocin gene clusters, and a terpene-encoding gene cluster [43]. The E2N171, E1N211, and ETH9427 strains in clade 4 have a remarkably high number of conserved SMBGCs, with each strain showing 54 to

72% of the identified gene clusters that are present in one or both of the other two strains. However, only seven gene clusters are conserved in *Streptomyces* sp. 4F, which is consistent with its taxonomic relationship in clade 4. These results indicate that phylogenetic relationships are accompanied by specialized gene cluster maintenance.



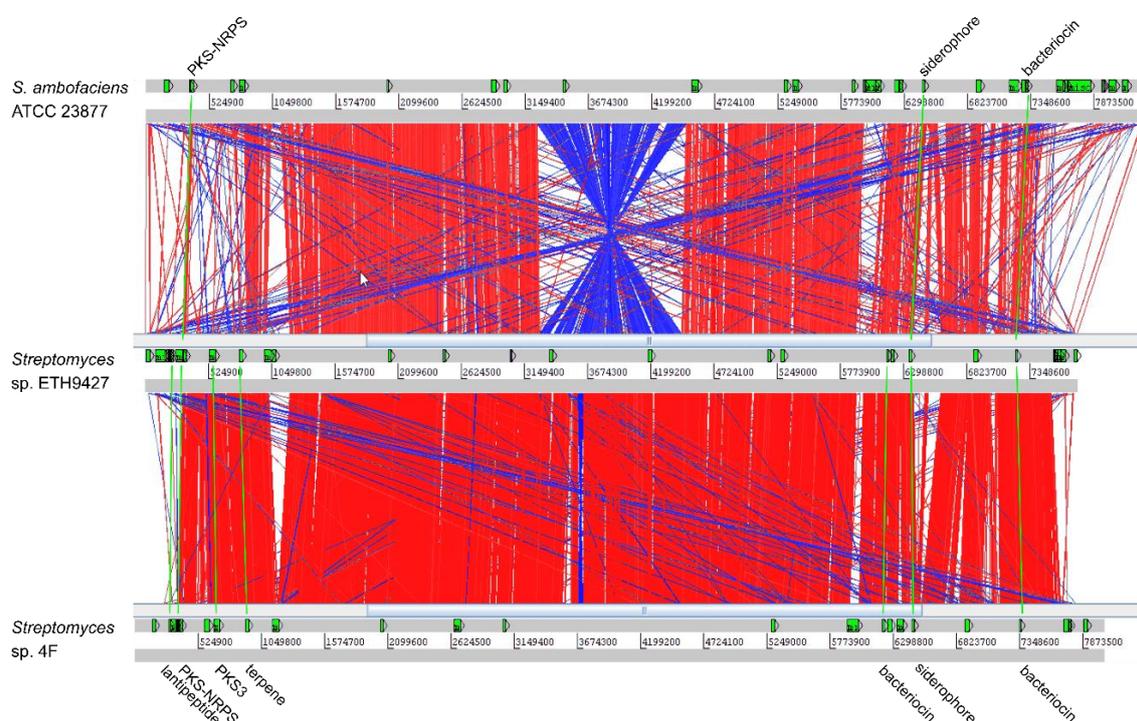
**Figure 4.** Conserved and strain-specific biosynthetic gene clusters. Within each clade, conserved and strain-specific gene clusters (circles) were identified. Original chromosomal organization is not considered, and clusters with the same position are identical. Colors represent the major classes of specialized metabolites: Type I, II, and III PKS (orange), NRPS (green), hybrid PKS–NRPS (light blue), bacteriocin (dark blue), lantipeptide (yellow), lasso peptide (pink), siderophore (red), terpene (purple), and others that include other KS pathways and less common metabolites such as phenazines, butyrolactones, and indoles (grey). Unknown clusters are indicated with “/” and duplicated clusters with “X”.

#### 2.4. Strain-Specific Biosynthetic Gene Clusters

Other gene clusters exist in each of the six examined strains beyond the core and conserved metabolome, the strain-specific SMBGCs. These were determined as gene clusters that are exclusive to a given strain and absent from its phylogenetically closest relatives, based on the nucleotide identity of putative biosynthetic genes. The putative products of these strain-specific biosynthetic pathways comprise all the major classes of specialized metabolites (Figure 4). Each strain harbors at least one strain-specific gene cluster. The fewest strain-specific gene clusters were encountered in strains E5N298 and ETH9427, with only one each. Strains E5N91 and E5N298, which are taxonomically close to their clade’s counterparts, dedicate almost opposite amounts of their specialized metabolome to strain-specific products (21% and 3%, respectively), perhaps as a result of horizontal gene transfer events and reflecting different levels of adaptation [44,45]. Conversely, other strains show a significantly higher number of strain-specific gene clusters. Strains E2N171 and E2N166 harbor six and 15 strain-specific gene clusters that represent 21% and 48% of their specialized metabolome, respectively.

The preliminary results of genome alignment and sequence comparison of the relatively close E2N166 and *S. ambofaciens* strains reveals large chromosome reorganizations events, such as a pericentral inversion, and most of the acquired strain-specific gene clusters are unexpectedly located in the central region of the chromosome. Moreover, genome alignment and comparison of the ETH9427 strain with its closest related strain *Streptomyces* sp. 4F and the reference species *S. ambofaciens* ATCC 23477 shows a relatively conserved chromosome organization, particularly between the first two strains (Figure 5). The identified SMBGCs also reflect the shared chromosomal regions, with seven gene clusters shared between ETH9427 and *Streptomyces* sp. 4F, most of which are also maintaining relative

positions, and only three gene clusters between ETH9427 and *S. ambofaciens* ATCC 23477 (Figure 5), supporting the notion that the distribution of SMBGCs is coherent with phylogenetic relationships [46].



**Figure 5.** Genome alignment and specialized metabolome comparison of *Streptomyces* sp. ETH9427. Alignment was performed with the Artemis Comparison Tool software (ACT) using an identity threshold of 77% and a score threshold of 500. Synteny regions are represented by red lines, inversion events are represented in blue, and breaks in synteny are seen as blank spaces. SMBGCs are marked in the horizontal panels (green arrows), and conserved gene clusters are linked (green lines).

While some of the identified strain-specific gene clusters have orthologues in other species, four out of the six strains that are used in this work (strains E5N91, E2N166, E2N171, and E1N211) have between one and six unknown biosynthetic gene clusters. These represent 3% to 19% of the metabolome, leading to a total of 16 SMBGCs that could possibly lead to novel chemistry. The strain-specific gene clusters identified in E2N171 and E1N211 (six NRPS, a type I PKS and a siderophore-like and another gene cluster classified as “other”) are of particular interest, as they can conceivably be responsible for the antimicrobial activities previously observed in Benhadj et al., namely against nosocomial infections responsible bacterial strains of *Escherichia coli*, *Streptococcus aureus* and *Pseudomonas aeruginosa*. These bioactivities are exclusive to E2N171 and E1N211 amongst the entire collection, and are absent from strains such as E5N298; therefore, the identified strain-specific SMBGCs in these two strains are good candidates for further detailed characterization [13].

### 3. Materials and Methods

#### 3.1. Strains and DNA Extraction

Five strains from a previously built *Streptomyces* collection were used in this work [13]. In brief, isolates were obtained from water samples of Lake Fetzara (Annaba, Algeria) on different agar media supplemented with antifungals. Additionally, a strain from the ETH collection (Eidgenössische Technische Hochschule Kultursammlungen, Zurich, Switzerland), *Streptomyces* sp. ETH9427 (accession numbers CP029624, CP029625, CP029626), was also analyzed. Genomic DNA was extracted from liquid cultures using a standard “salting out” protocol [47].

### 3.2. Whole-Genome Sequencing and Annotation

Draft genomes were constructed by assembling reads from an Illumina Genome Analyzer (Illumina, San Diego, CA, USA), with pair-ends (c.300 bp) and/or mate-pairs (c.8 kb) DNA fragments libraries, and were processed as described in Thibessard et al. [14]. Sequences were trimmed to eliminate adaptor sequences, and contigs with a size smaller than 1000 pb were eliminated. Draft genomes of the strains E5N91, E2N166, E2N171, E1N211, and E5N298 were submitted into the NCBI GenBank database and given the following accession numbers: RAIE000000000, RAIF000000000, RAIH000000000, RAIJ000000000, and RAIG000000000, respectively. Whenever possible, sequence assembly and contig ordering were achieved using the CLC Main Workbench (Qiagen, Hilden, Germany), and coding sequence prediction and annotation were automatically performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [48]. Genome alignments were performed with Artemis Comparison Tool (ACT) [49]. Average nucleotide identity (ANI) was determined using the JSpeciesWS server with the BLAST algorithm (<http://jspecies.ribohost.com/jspeciesws/>) [19].

### 3.3. Phylogenetic Analysis and Biosynthetic Gene Cluster Identification

The phylogenetic relationships between strains were reconstructed with an MLST analysis [50], using DNA sequences from *atpB*, *gyrB*, *recA*, *rpoB*, and *trpB* genes. Several reference *Streptomyces* species were included in the analysis: *S. ambofaciens* ATCC 23877 (accession number CP012382.1), *S. ambofaciens* DSM 40697 (accession number CP012949.1), *S. coelicolor* A3(2) (accession number NC\_003888.3), *S. lividans* TK24 (accession number CP009124.1), *Streptomyces* sp. M1013 (accession number MQUH000000000.1), and *Streptomyces* sp. 4F (accession number CP013142.1). The actinobacteria *Corynebacterium glutamicum* ATCC 13032 (accession number BA000036.3) was used to root the phylogenetic tree. Nucleotide sequence alignments were performed with SeaView software [51] using the muscle algorithm, and a concatenated gene sequence (20592 nt in total) was used to construct a neighbor-joining phylogenetic tree with Kimura's 2-parameter distance correction and 100 bootstrap replicates with MEGA7 software [52]. Overall, there was a good separation and statistical support for most of the branches in the tree. Biosynthetic gene clusters were predicted using the antiSMASH 3.0 server (<https://antismash.secondarymetabolites.org/#!/start>) without ClusterFinder algorithm [26].

## 4. Conclusions

The efforts for *Streptomyces* environmental strains genome sequencing continue to show the genus SMBGC's diversity. Strain de-replication of isolates can be achieved using multi-gene sequence comparison instead of the traditional 16S rRNA, and the analysis of isolated strains reveals a pool for novel specialized metabolites in taxonomically closely related strains. Encountered gene cluster diversity supports phylogenetic relationships, where closely related strains share a higher number of SMBGCs. Horizontal gene transfer as a source of genetic diversity does not distort phylogenetic examination, as these are rare events, and the presence or absence of a SMBGC is not enough to blur phylogenetic relationships. Strain-level analysis has already been demonstrated to lead to the identification of new biosynthetic pathways. Though the strain-specific gene clusters identified here require further characterization, they nonetheless reflect that chemical richness can be found at this level, and imply that the deep sequencing of strains belonging to same species will continue to produce originality.

**Author Contributions:** Conceptualization, C.M.V., P.L. and B.A.; methodology, C.M.V.; validation, P.L. and B.A.; formal analysis, C.M.V.; investigation, C.M.V., L.H. and J.-N.L.; resources, M.H., D.K.-G.; data curation, C.M.V., P.L., A.T. and O.L.; writing—original draft preparation, C.M.V.; writing—review and editing, B.A. and P.L.; visualization, all authors; supervision, P.L. and B.A.; project administration, B.A.; funding acquisition, B.A. and P.L.

**Funding:** This work was funded by the French National Research Agency through two programs (ANR Streptoflux ANR-07-BLAN-0096, ANR MiGenIs ANR-13-BSV6-0006), the French National Institute for Agricultural

Research (INRA), Région Lorraine and the Laboratory Excellence ARBRE (ANR-11-LABX-0002-01). CMV was also supported by the AgreenSkillsPlus Program (FP7-609398.0000) and the Région Grand Est.

**Acknowledgments:** The authors acknowledge the support of the DynAMic laboratory by the Impact Biomolecules project of the “Lorraine Université d’Excellence (Investissements d’avenir—ANR)”.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Bérdy, J. Bioactive microbial metabolites. *J. Antibiot.* **2005**, *58*, 1–26. [[CrossRef](#)] [[PubMed](#)]
2. Worrall, J.A.R.; Vijgenboom, E. Copper mining in *Streptomyces*: Enzymes, natural products and development. *Nat. Prod. Rep.* **2010**, *27*, 742–756. [[CrossRef](#)] [[PubMed](#)]
3. 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](#)]
4. Corre, C.; Challis, G.L. New natural product biosynthetic chemistry discovered by genome mining. *Nat. Prod. Rep.* **2009**, *26*, 977–986. [[CrossRef](#)] [[PubMed](#)]
5. Lewis, K. Platforms for antibiotic discovery. *Nat. Rev. Drug. Discov.* **2013**, *12*, 371–387. [[CrossRef](#)] [[PubMed](#)]
6. Rebets, Y.; Kormanec, J.; Luzhetskyy, A.; Bernaerts, K.; Anné, J. Cloning and Expression of Metagenomic DNA in *Streptomyces lividans* and Subsequent Fermentation for Optimized Production. *Methods Mol. Biol.* **2017**, *1539*, 99–144. [[CrossRef](#)] [[PubMed](#)]
7. Bentley, S.D.; Chater, K.F.; Cerdeño-Tárraga, A.-M.; Challis, G.L.; Thomson, N.R.; James, K.D.; Harris, D.E.; Quail, M.A.; Kieser, H.; Harper, D.; et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **2002**, *417*, 141–147. [[CrossRef](#)] [[PubMed](#)]
8. Medema, M.H.; Fischbach, M.A. Computational approaches to natural product discovery. *Nat. Chem. Biol.* **2015**, *11*, 639–648. [[CrossRef](#)] [[PubMed](#)]
9. Nett, M.; Ikeda, H.; Moore, B.S. Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat. Prod. Rep.* **2009**, *26*, 1362–1384. [[CrossRef](#)] [[PubMed](#)]
10. Seipke, R.F. Strain-level diversity of secondary metabolism in *Streptomyces albus*. *PLoS ONE* **2015**, *10*, e0116457. [[CrossRef](#)] [[PubMed](#)]
11. Antony-Babu, S.; Stien, D.; Eparvier, V.; Parrot, D.; Tomasi, S.; Suzuki, M.T. Multiple *Streptomyces* species with distinct secondary metabolomes have identical 16S rRNA gene sequences. *Sci. Rep.* **2017**, *7*, 11089. [[CrossRef](#)] [[PubMed](#)]
12. Ziemert, N.; Lechner, A.; Wietz, M.; Millán-Aguíñaga, N.; Chavarria, K.L.; Jensen, P.R. Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E1130–E1139. [[CrossRef](#)] [[PubMed](#)]
13. Benhadj, M.; Gacemi-Kirane, D.; Toussaint, M.; Hotel, L.; Bontemps, C.; Duval, R.E.; Aigle, B.; Leblond, P. Diversity and antimicrobial activities of *Streptomyces* isolates from Fetzara Lake, north eastern Algeria. *Ann. Biol. Clin.* **2018**, *76*, 81–95. [[CrossRef](#)] [[PubMed](#)]
14. Thibessard, A.; Vicente, C.M.; Bertrand, C.; Aigle, B.; Leblond, P. Draft genome of *Streptomyces* sp. ETH9427, whole genome shotgun sequence. *Genome Announc.* **2018**, in press.
15. Rückert, C.; Albersmeier, A.; Busche, T.; Jaenicke, S.; Winkler, A.; Friðjónsson, Ó.H.; Hreggviðsson, G.Ó.; Lambert, C.; Badcock, D.; Bernaerts, K.; et al. Complete genome sequence of *Streptomyces lividans* TK24. *J. Biotechnol.* **2015**, *199*, 21–22. [[CrossRef](#)] [[PubMed](#)]
16. Haas, D.; Gerbaud, C.; Sahin, N.; Pernodet, J.-L.; Lautru, S. Draft Genome Sequence of *Streptomyces* sp. M1013, a Close Relative of *Streptomyces ambofaciens* and *Streptomyces coelicolor*. *Genome Announc.* **2017**, *5*. [[CrossRef](#)] [[PubMed](#)]
17. Thibessard, A.; Haas, D.; Gerbaud, C.; Aigle, B.; Lautru, S.; Pernodet, J.-L.; Leblond, P. Complete genome sequence of *Streptomyces ambofaciens* ATCC 23877, the spiramycin producer. *J. Biotechnol.* **2015**, *214*, 117–118. [[CrossRef](#)] [[PubMed](#)]
18. Thibessard, A.; Leblond, P. Complete Genome Sequence of *Streptomyces ambofaciens* DSM 40697, a Paradigm for Genome Plasticity Studies. *Genome Announc.* **2016**, *4*. [[CrossRef](#)] [[PubMed](#)]

19. Richter, M.; Rosselló-Móra, R.; Oliver Glöckner, F.; Peplies, J. JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* **2016**, *32*, 929–931. [[CrossRef](#)] [[PubMed](#)]
20. Rosselló-Móra, R.; Amann, R. Past and future species definitions for Bacteria and Archaea. *Syst. Appl. Microbiol.* **2015**, *38*, 209–216. [[CrossRef](#)] [[PubMed](#)]
21. Alam, M.T.; Merlo, M.E.; Takano, E.; Breitling, R. Genome-based phylogenetic analysis of *Streptomyces* and its relatives. *Mol. Phylogenet. Evol.* **2010**, *54*, 763–772. [[CrossRef](#)] [[PubMed](#)]
22. Rong, X.; Liu, N.; Ruan, J.; Huang, Y. Multilocus sequence analysis of *Streptomyces griseus* isolates delineating intraspecific diversity in terms of both taxonomy and biosynthetic potential. *Antonie. Van. Leeuwenhoek* **2010**, *98*, 237–248. [[CrossRef](#)] [[PubMed](#)]
23. Rong, X.; Huang, Y. Taxonomic evaluation of the *Streptomyces hygroscopicus* clade using multilocus sequence analysis and DNA-DNA hybridization, validating the MLSA scheme for systematics of the whole genus. *Syst. Appl. Microbiol.* **2012**, *35*, 7–18. [[CrossRef](#)] [[PubMed](#)]
24. Labeda, D.P.; Dunlap, C.A.; Rong, X.; Huang, Y.; Doroghazi, J.R.; Ju, K.-S.; Metcalf, W.W. Phylogenetic relationships in the family Streptomycetaceae using multi-locus sequence analysis. *Antonie Van Leeuwenhoek* **2017**, *110*, 563–583. [[CrossRef](#)] [[PubMed](#)]
25. Jiménez, G.; Urdiain, M.; Cifuentes, A.; López-López, A.; Blanch, A.R.; Tamames, J.; Kämpfer, P.; Kolstø, A.-B.; Ramón, D.; Martínez, J.F.; et al. Description of *Bacillus toyonensis* sp. nov., a novel species of the *Bacillus cereus* group, and pairwise genome comparisons of the species of the group by means of ANI calculations. *Syst. Appl. Microbiol.* **2013**, *36*, 383–391. [[CrossRef](#)] [[PubMed](#)]
26. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H.U.; Brucoleri, R.; Lee, S.Y.; Fischbach, M.A.; Müller, R.; Wohlleben, W.; et al. AntiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **2015**, *43*, W237–W243. [[CrossRef](#)] [[PubMed](#)]
27. Bernan, V.; Filpula, D.; Herber, W.; Bibb, M.; Katz, E. The nucleotide sequence of the tyrosinase gene from *Streptomyces antibioticus* and characterization of the gene product. *Gene* **1985**, *37*, 101–110. [[CrossRef](#)]
28. Malin, G.; Lapidot, A. Induction of synthesis of tetrahydropyrimidine derivatives in *Streptomyces* strains and their effect on *Escherichia coli* in response to osmotic and heat stress. *J. Bacteriol.* **1996**, *178*, 385–395. [[CrossRef](#)] [[PubMed](#)]
29. Davis, N.K.; Chater, K.F. Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. *Mol. Microbiol.* **1990**, *4*, 1679–1691. [[CrossRef](#)] [[PubMed](#)]
30. Neilands, J.B. Microbial iron compounds. *Annu. Rev. Biochem.* **1981**, *50*, 715–731. [[CrossRef](#)] [[PubMed](#)]
31. Gürtler, H.; Pedersen, R.; Anthoni, U.; Christophersen, C.; Nielsen, P.H.; Wellington, E.M.; Pedersen, C.; Bock, K. Albaflavenone, a sesquiterpene ketone with a zizaene skeleton produced by a streptomycete with a new rope morphology. *J. Antibiot.* **1994**, *47*, 434–439. [[CrossRef](#)] [[PubMed](#)]
32. Poralla, K.; Muth, G.; Härtner, T. Hopanoids are formed during transition from substrate to aerial hyphae in *Streptomyces coelicolor* A3(2). *FEMS Microbiol. Lett.* **2000**, *189*, 93–95. [[CrossRef](#)] [[PubMed](#)]
33. Krügel, H.; Krubasik, P.; Weber, K.; Saluz, H.P.; Sandmann, G. Functional analysis of genes from *Streptomyces griseus* involved in the synthesis of isorenieratene, a carotenoid with aromatic end groups, revealed a novel type of carotenoid desaturase. *Biochim. Biophys. Acta* **1999**, *1439*, 57–64. [[CrossRef](#)]
34. Gerber, N.N.; Lechevalier, H.A. Geosmin, an earthy-smelling substance isolated from actinomycetes. *Appl. Microbiol.* **1965**, *13*, 935–938. [[PubMed](#)]
35. Choudoir, M.J.; Pepe-Ranne, C.; Buckley, D.H. Diversification of Secondary Metabolite Biosynthetic Gene Clusters Coincides with Lineage Divergence in *Streptomyces*. *Antibiotics* **2018**, *7*, 12. [[CrossRef](#)]
36. Lautru, S.; Deeth, R.J.; Bailey, L.M.; Challis, G.L. Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat. Chem. Biol.* **2005**, *1*, 265–269. [[CrossRef](#)] [[PubMed](#)]
37. Hesketh, A.; Kock, H.; Mootien, S.; Bibb, M. The role of *absC*, a novel regulatory gene for secondary metabolism, in zinc-dependent antibiotic production in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.* **2009**, *74*, 1427–1444. [[CrossRef](#)] [[PubMed](#)]
38. Chong, P.P.; Podmore, S.M.; Kieser, H.M.; Redenbach, M.; Turgay, K.; Marahiel, M.; Hopwood, D.A.; Smith, C.P. Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of *Streptomyces coelicolor* A3(2). *Microbiology* **1998**, *144 Pt 1*, 193–199. [[CrossRef](#)] [[PubMed](#)]

39. Takano, E.; Kinoshita, H.; Mersinias, V.; Bucca, G.; Hotchkiss, G.; Nihira, T.; Smith, C.P.; Bibb, M.; Wohlleben, W.; Chater, K. A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol. Microbiol.* **2005**, *56*, 465–479. [[CrossRef](#)] [[PubMed](#)]
40. Wright, L.F.; Hopwood, D.A. Actinorhodin is a chromosomally-determined antibiotic in *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* **1976**, *96*, 289–297. [[CrossRef](#)] [[PubMed](#)]
41. Rudd, B.A.; Hopwood, D.A. A pigmented mycelial antibiotic in *Streptomyces coelicolor*: Control by a chromosomal gene cluster. *J. Gen. Microbiol.* **1980**, *119*, 333–340. [[CrossRef](#)] [[PubMed](#)]
42. Aigle, B.; Pang, X.; Decaris, B.; Leblond, P. Involvement of AlpV, a new member of the *Streptomyces* antibiotic regulatory protein family, in regulation of the duplicated type II polyketide synthase alp gene cluster in *Streptomyces ambofaciens*. *J. Bacteriol.* **2005**, *187*, 2491–2500. [[CrossRef](#)] [[PubMed](#)]
43. Aigle, B.; Lautru, S.; Spittler, D.; Dickschat, J.S.; Challis, G.L.; Leblond, P.; Pernodet, J.-L. Genome mining of *Streptomyces ambofaciens*. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 251–263. [[CrossRef](#)] [[PubMed](#)]
44. Jensen, P.R.; Williams, P.G.; Oh, D.-C.; Zeigler, L.; Fenical, W. Species-specific secondary metabolite production in marine actinomycetes of the genus *Salinispora*. *Appl. Environ. Microbiol.* **2007**, *73*, 1146–1152. [[CrossRef](#)] [[PubMed](#)]
45. Penn, K.; Jenkins, C.; Nett, M.; Udworthy, D.W.; Gontang, E.A.; McGlinchey, R.P.; Foster, B.; Lapidus, A.; Podell, S.; Allen, E.E.; et al. Genomic islands link secondary metabolism to functional adaptation in marine actinobacteria. *ISME J.* **2009**, *3*, 1193–1203. [[CrossRef](#)] [[PubMed](#)]
46. Joynt, R.; Seipke, R.F. A phylogenetic and evolutionary analysis of antimycin biosynthesis. *Microbiology* **2018**, *164*, 28–39. [[CrossRef](#)] [[PubMed](#)]
47. Kieser, T.; Bibb, M.; Buttner, M.; Chater, K.; Hopwood, D.A. *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, UK, 2000.
48. Tatusova, T.; DiCuccio, M.; Badretdin, A.; Chetvernin, V.; Nawrocki, E.P.; Zaslavsky, L.; Lomsadze, A.; Pruitt, K.D.; Borodovsky, M.; Ostell, J. NCBI prokaryotic genome annotation pipeline. *Nucleic. Acids. Res.* **2016**, *44*, 6614–6624. [[CrossRef](#)] [[PubMed](#)]
49. Carver, T.; Berriman, M.; Tivey, A.; Patel, C.; Böhme, U.; Barrell, B.G.; Parkhill, J.; Rajandream, M.-A. Artemis and ACT: Viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* **2008**, *24*, 2672–2676. [[CrossRef](#)] [[PubMed](#)]
50. Guo, Y.; Zheng, W.; Rong, X.; Huang, Y. A multilocus phylogeny of the *Streptomyces griseus* 16S rRNA gene clade: Use of multilocus sequence analysis for streptomycete systematics. *Int. J. Syst. Evol. Microbiol.* **2008**, *58*, 149–159. [[CrossRef](#)] [[PubMed](#)]
51. Gouy, M.; Guindon, S.; Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **2010**, *27*, 221–224. [[CrossRef](#)] [[PubMed](#)]
52. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [[CrossRef](#)] [[PubMed](#)]

