Scientific Research Publishing

Advances in Microbiology, 2026, 16(4), 151-172
https://www.scirp.org/journal/aim
ISSN Online: 2165-3410
ISSN Print: 2165-3402

Insights into the Genotypic and Phenotypic Features of *Streptomyces* sp. PSAA01: Exploring Its Vast Promising Potential

Prasenjit Das*#, Biraj Sarkar*, Dipanwita Patra*, Sukhendu Mandal#

Laboratory of Molecular Bacteriology, Department of Microbiology, University of Calcutta, Kolkata, India
Email: #prasenjit.microbio@gmail.com, #smmicrobio@caluniv.ac.in

How to cite this paper: Das, P., Sarkar, B., Patra, D. and Mandal, S. (2026) Insights into the Genotypic and Phenotypic Features of *Streptomyces* sp. PSAA01: Exploring Its Vast Promising Potential. *Advances in Microbiology*, **16**, 151-172.
https://doi.org/10.4236/aim.2026.164008

Received: January 12, 2026
Accepted: March 28, 2026
Published: March 31, 2026

Abstract

Streptomyces sp. are actinobacteria, which are generally known to be prolific producers of antibiotics and various antimicrobial compounds. More than 90% of the reported *Streptomyces* sp. are known to be isolated from soil samples. Exploring the genomes of such actinobacteria has sparked interest in finding bioactive secondary metabolites after the remarkable breadth of biosynthetic gene clusters (BGCs) for antibacterial compounds was discovered. *Streptomyces* sp. PSAA01 was obtained from soil samples taken in Manas National Park, Assam, India. It was discovered that this strain and *S. yatensis* DSM 41771^T are closely related. The strain PSAA01 has been found to harbour diverse biosynthetic gene clusters (BGCs), which might be responsible for the production of many significant bioactive secondary metabolites.

Keywords

Actinobacteria, Antimicrobial, Whole-Genome, Genomics, Biosynthetic Gene Clusters (BGCs)

1. Introduction

Actinobacteria are Gram-positive bacteria with characteristic high GC content and are known to constitute a cosmopolitan phylum that includes both rod-shaped and filamentous bacteria [1]-[3]. Actinobacteria are known to be ubiquitous in terrestrial, freshwater, and marine ecosystems and exhibit the ability for both solitary inhabitation and symbioses with other microbes and/or higher-order organisms [3] [4]. The filamentous actinobacteria, particularly those which be-

*These authors have contributed equally to this work.
#Corresponding author.

long to the family Actinomycetaceae, are prolific producers of a wide range of secondary metabolites, which include around two-thirds of all known antibiotics, various antifungal, anticancer, and immunosuppressive agents, as well as numerous industrially relevant enzymes [3]. Among all the reported genera belonging to Actinomycetaceae, the most explored and well-studied genus is *Streptomyces* [3]. It is an incredibly diverse genus constituted of around 1231 species, which are included in the List of Prokaryotic names with Standing in Nomenclature, *i.e.*, LPSN [5] with a validly published correct name under ICNP (International Code of Nomenclature of Prokaryotes), as accessed on 16th July 2024. Aerobic soil bacteria make up the majority of the *Streptomyces* species that have been reported [3] and are supposed to constitute almost 90% of all the actinobacteria found in soil [6]. Species belonging to the genus *Streptomyces* display a distinct filamentous life cycle and reproduce by producing spores [3]. The commencement of sporulation is believed to be associated with the production of numerous bioactive secondary metabolites like antibiotics, anti-fungals, antivirals, antitumor, and/or insecticidal compounds [3] [7] [8]. The genus *Streptomyces* produces around 50% - 55% of the total antibiotics produced by all microorganisms [9] [10] and approximately 80% of all the antibiotics produced by actinobacteria [11]-[14].

Unfortunately, the rapid emergence of resistant bacteria, which has been occurring worldwide, compromises the effectiveness of once-significant antibiotics. This catastrophe is attributed to the overuse and misuse of antibiotics, along with the scarcity of new, efficient drug discoveries due to a lack of economic incentives and challenging regulatory requirements [15].

Amidst this crisis, there is a recent growing interest in finding new natural products (NPs) with diverse antimicrobial or antibiotic properties [3]. The discovery of an unforeseen number of biosynthetic gene clusters (BGCs) for antibiotic-like substances from the genomes of actinomycetes has acted as a stimulus among microbiologists to explore the genomes of actinomycetes to discover bioactive secondary metabolites [3] [9] [16]-[18].

This study delineates the polyphasic taxonomical identification of a *Streptomyces* sp. PSAA01, which was isolated from a soil sample of Manas National Park, Assam, India. The genome of strain PSAA01 was explored in search of various biosynthetic gene clusters (BGCs).

2. Materials and Methods

2.1. Isolation and Maintenance

The strain PSAA01 is a soil isolate obtained from a soil sample of Manas National Park, Assam, India. Initially, to selectively isolate actinobacteria from the soil sample, the sample was subjected to $CaCO_3$ treatment at an ambient temperature for around 168 hours, followed by heating the soil sample at 65˚C for 2 hours [19]. Following the preparatory steps for the selective separation of actinobacteria from the soil sample, 1 g of the soil was dissolved in 1 ml of 0.9% NaCl solution and diluted up to 10^{-6} times. Later, 100 µL of the final dilution sample was spread on

Starch casein medium [20] supplemented with 50 µg/mL of two antifungal compounds, *i.e.*, nystatin and cycloheximide, and incubated for 3 to 4 days at 30˚C until single colonies appeared on the selection plate. Later, the colonies were individually picked with a sterile inoculation loop and streaked again on sterile Starch casein agar plates for purification.

2.2. Biochemical Characterization

The carbon utilization behavior of the isolated strain PSAA01 was assessed using 1% (w/v) of different carbon sources, such as glucose, galactose, ribose, arabinose, maltose, xylose, rhamnose, melibiose, raffinose, adonitol, inositol, melizitose, pyruvate, and dulcitol in ISP-9 medium [21] [22]. The tests performed were as previously reported for the hydrolysis of starch, cellulose breakdown, tributyrin hydrolysis, nitrate reduction, gelatin liquefaction, IMVIC test, and H_2S generation [23]. Using conventional techniques, the production of urease, catalase, and indole acetic acid (IAA) was evaluated [23]. This isolate was shown to hydrolyze xanthine, hypoxanthine, xylan, tyrosine, casein, and arbutin, which was assessed by the standard method [23]. The standard Hicarbo Kit from Himedia, India, which has 35 distinct carbon sources, was used to ferment the carbohydrates.

2.3. Genomic DNA Isolation and Whole-Genome Sequencing

The genomic DNA of PSAA01 was isolated from freshly grown 5 mL of ISP-2 broth. The genomic DNA was extracted by the phenol: chloroform method according to the standard protocol [24]. Using the Illumina NovaSeq 6000 platform (Neuberg Diagnostics Pvt Ltd., Ahmedabad, India), the paired-end libraries were produced and sequenced.

The complete 16S rDNA sequence of the isolate PSA001 was obtained from the draft genome sequence of the strain and was compared with the 16S rDNA sequences of the closest type strains using EzBioCloud [25]. Using MEGA 6 software and the ClustalW program, multiple sequence alignment was carried out with the 16S rDNA sequences [26]. The phylogenetic tree based on the 16S rDNA sequences was constructed using the Maximum-Likelihood (ML) method [27]. Phylogenetic analysis was performed with the Jukes-Cantor model [28], gamma distributed with invariant sites (G + I) for the ML method [29]. Bootstrap analysis (1000 resampled datasets) was used to evaluate the topology of the trees.

For the multi-locus sequence analysis (MLSA) of PSAA01, four housekeeping genes, *i.e.*, *atpD* (F_0 - F_1 ATP synthase subunit beta), *gyrB* (DNA gyrase subunit B), *recA* (recombinase RecA), and *rpoB* (DNA-directed RNA polymerase subunit beta) were obtained from its draft genome sequence. The four mentioned housekeeping gene sequences of the closest relatives (based on 16S rDNA nucleotide sequence data) were directly downloaded from GenBank or retrieved from respective draft or complete genome sequences. Before being subjected to further analysis, the relevant genes' nucleotide sequences were manually cut at the same position after being aligned using MEGA6 software. The four housekeeping gene nu-

cleotide sequences were then concatenated head to tail as *atpD-gyrB-recA-rpoB* (1074, 1724, 1000, and 2774 nucleotides, respectively) for PSAA01 and its closest relatives. Pairwise distance was calculated using the Kimura-2 parameter model [30]. MEGA 6 software was used to determine the optimal model for maximum likelihood phylogenetic analysis, which was subsequently applied to the analysis. The best model for maximum likelihood was found to be the General Time Reversible model with Gamma distributed with invariant sites (GTR + G + I) [31]. The bootstrap technique (using 1000 similar datasets) was employed to assess the tree topology.

The closest type strain genomes were aligned and compared using the BLAST Ring Image Generator (BRIG) to create a circular map of each genome [32], with PSAA01 as the reference strain. The National Center for Biotechnology Information (NCBI) and KEGG GENOME Database are public databases that provide access to and downloads of the genome sequences of all known type strains of *Streptomyces* (a total of 30 strains, including strain PSA001). Whole genome-based taxonomic analyses were conducted using the Type Strain Genome Server (TYGS) (https://tygs.dsmz.de) [33]. FastME was used to build the phylogenomic tree from the genome blast distance phylogeny (GBDP). First, pairwise genome comparisons were performed using the GBDP, and inter-genomic distances were inferred under the algorithm "trimming" and distance formula d5. The inference of branch supports was derived from 100 pseudo-bootstrap replicates [33]. The genome was annotated according to the RAST protocol [34].

The average nucleotide identity (ANI) values between strain PSAA01 and its closest type strains were calculated in EzGenome (https://ezbiocloud.net/tools/ani) [25]; the average amino acid identity (AAI) was calculated using the online calculator (http://enve-omics.ce.gatech.edu/aai/) [35], and the genome-to-genome distances (GGDs)/digital DNA-DNA hybridization (dDDH) values were calculated using the Genome-to-Genome Distance Calculator (GGDC; http://ggdc.dsmz.de) [36].

2.4. Biosynthetic Gene Clusters (BGCs): Characterization and Comparative Study of Their Distribution

Secondary metabolite biosynthetic gene clusters (BGCs) from the genome sequences of PSAA01 and its 10 closest type strains were analyzed using antiSMASH version 6.1.0 [37] to examine the distribution of BGCs among the studied organisms. The structures of the secondary metabolites encoded by the genome of PSAA01 were predicted using the PRISM 4.4.5 webserver [38].

2.5. Identification of BGCs, PKS-KS Domains, and NRPS C Domains

Antibiotics and secondary metabolite analysis shells were used to identify the BGCs found in the genome of *Streptomyces* (antiSMASH v7.1.0) (https://antismash.secondarymetabolites.org) [39]. After retrieving the relevant sequences from each identified BGC, phylogenetic analyses were performed for the KS and C domains of the PKS and NRPS genes, respectively.

3. Results and Discussion

The strain PSAA01 was isolated from soil samples taken from the Manas National Park in Assam, India. From the biochemical analysis of strain PSAA01 (Table S1), it has been found that the strain is capable of utilizing L-rhamnose, D-sucrose, inositol, melezitose, maltose, D-galactose, raffinose, adonitol, pyruvate, and melibiose; but is unable to use D-ribose, glucose, arabinose, xylose, and dulcitol; whereas, *S. yatensis* is reported to utilize glucose and several pentose sugars such as ribose, arabinose, and xylose. These differences in pentose sugar utilization further distinguish PSAA01 from *S. yatensis* [40]. According to the species description of *S. yatensis*, the type strain shows positive reactions for gelatin liquefaction, casein hydrolysis, and tyrosine degradation while PSAA01 shows negative results in caseinase and gelatinase activity. Strain PSAA01 utilized L-alanine, L-arginine, and L-asparagine. These reactions are consistent with general metabolic characteristics of the genus *Streptomyces* and do not strongly differentiate PSAA01 from *S. yatensis*. The strain PSAA01 can hydrolyze xylan and hypoxanthine, but cannot hydrolyze tyrosine and xanthine. PSAA01 was catalase positive and capable of nitrate reduction, consistent with reported characteristics of *S. yatensis*. It was negative for H_2S production, methyl red, Voges-Proskauer, and indole tests, which also aligns with general *streptomycete* properties.

Table 1. Shows the respective accession numbers, genome size (bp), and contigs used for the genomic analyses.

Strain	Accession No.	Genome size (bp)	Contigs
Streptomyces sp. PSAA01	NZ_JAKKUU000000000	9,224,189	271
S. melanosporofaciens DSM 40318^T	NZ_FNST00000000	10,769,732	2
S. antimycoticus NBRC 12839^T	NZ_BJHV00000000	11,174,199	3
S. yatensis DSM 41771^T	NZ_CP072941	10,360,357	1
S. rhizosphaericus DSM 41760^T	NZ_JAGMTS000000000	11,044,339	5
S. indonesiensis DSM 41759^T	NZ_JAGSHY000000000	11,666,822	2
S. cangkringensis DSM 41769^T	NZ_JAGMTV000000000	11,723,082	4
S. hygroscopicus subsp. *hygroscopicus* NBRC 13472^T	NZ_BBOX00000000	9,464,604	680
S. antioxidans MUSC 164^T	NZ_LAKD00000000	9,118,065	282
S. sioyaensis DSM 40032^T	NZ_SDIF00000000	7,847,945	289
S. decoyicus NRRL 2666^T	NZ_CP082301	8,632,952	1
S. asiaticus DSM 41761^T	NZ_JAGSHX000000000	11,877,923	6
S. rimosus subsp. *rimosus* ATCC 10970^T	NZ_CP048261	9,643,891	2
S. himastatinicus ATCC 53653^T	NZ_ACEX00000000	11,030,030	783
S. lydicus ATCC 25470^T	NZ_RDTD00000000	7,935,716	20
S. chattanoogensis NRRL ISP-5002^T	NZ_LGKG00000000	9,129,105	217
Kitasatospora setae KM-6054^T	NC_016109	8,783,278	1

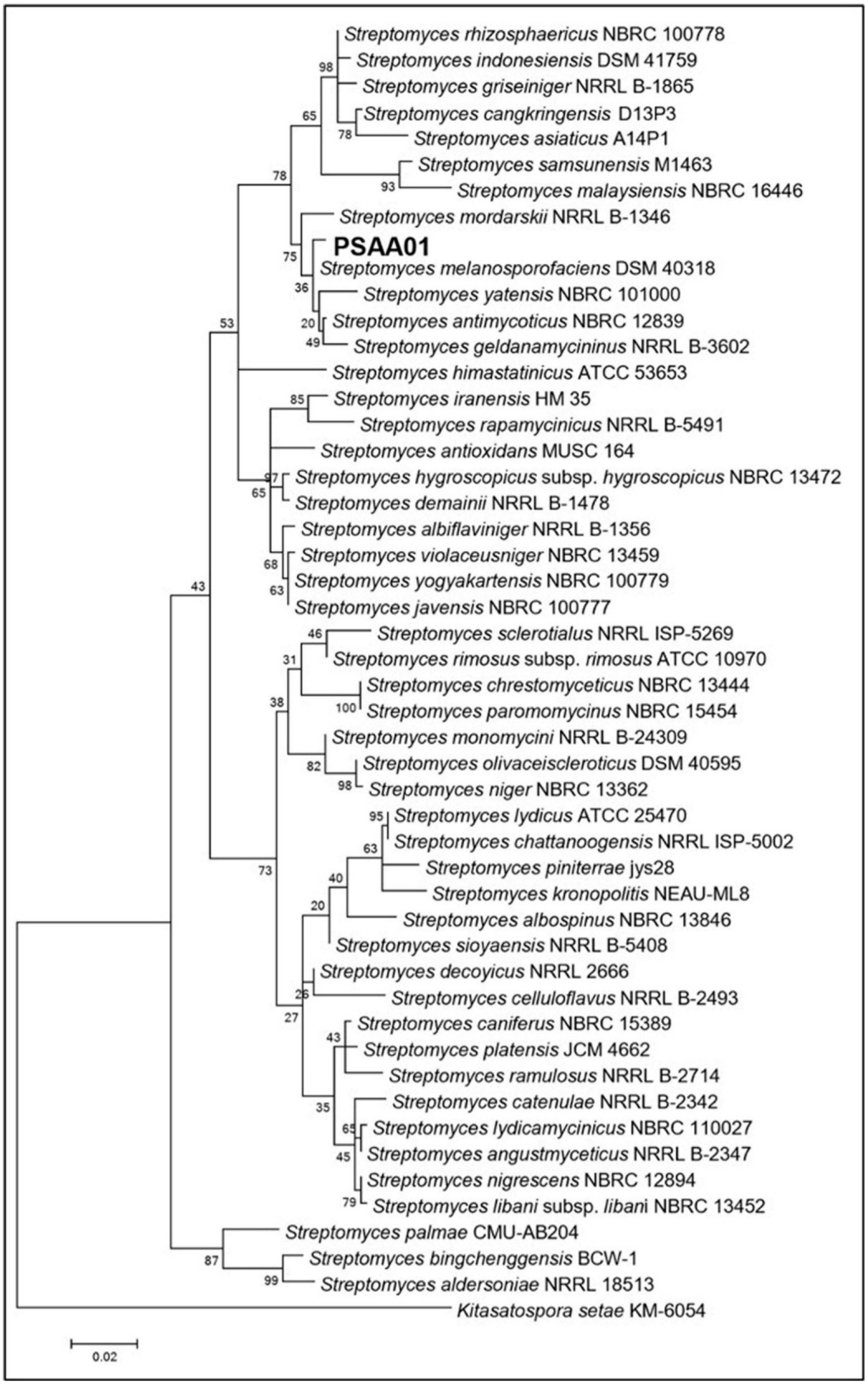

Figure 1. Shows the maximum-likelihood (ML) phylogenetic tree based on the 16S rRNA sequence of strain PSAA01 and its closest type strains. Only bootstrap values >95% (expressed as % of 1000 replications) are shown at the nodes. The bar at the bottom indicates the scale of the branch lengths.

The draft genome sequence of the strain PSAA01 was obtained in 271 contigs, with an estimated genome size of 9,224,189 bp and 71.2% GC content (Table 1). The draft sequence of the strain is submitted to the NCBI under the accession number NZ_JAKKUU000000000. The 1533 bp long complete *16S rRNA* gene sequence of the strain PSAA01 showed the highest similarity of 99.72% with *Streptomyces melanosporofaciens* DSM 40138^T according to EZBioCloud. Yet, according to the maximum-likelihood (ML) tree based on the *16S rRNA* gene sequences of strain PSAA01 and its closest type strains (Figure 1), it was found that strain PSAA01 was placed in a distinct clade. The whole genome sequences of the closest strains or type strains based on the 16S rDNA sequence similarities were obtained from the NCBI GenBank, along with an outgroup strain for phylogenetic analyses (Table 2).

For the visualization of circular genome comparison, a BLASTN-based ring image was generated by BRIG version 0.95 [32] for available genomes of the closest type strains with *Streptomyces* sp. PSAA01 as the reference strain (Figure 2).

Table 2. Comparative genotypic analysis of *Streptomyces* sp. PSAA01 and its closest related type strains. [a]**ANI values** (determined on https://ezbiocloud.net/tools/ani); [b]**AAI values** (determined on http://enve-omics.ce.gatech.edu/aai/); [c]**Digital DNA-DNA hybridization (dDDH)** calculated with the Genome-to-Genome Distance Calculator 3.0 (GGDC 3.0) (available at https://ggdc.dsmz.de/); [d]**MLSA** distances (*i.e.*, the pairwise distances of the housekeeping genes concatenated in the order *atpD-gyrB-recA-rpoB*) calculated in MEGA 6 software using the Kimura-2 parameter model for distance calculation.

Strains	ANI value[a]	AAI value[b]	dDDH[c]	MLSA distance[d]
S. melanosporofaciens DSM 40318^T	95.75	94.93	63.90	0.009
S. antimycoticus NBRC 12839^T	95.52	94.41	62.90	0.009
S. yatensis DSM 41771^T	97.21	96.53	73.80	0.005
S. rhizosphaericus DSM 41760^T	91.42	90.07	44.00	0.029
S. indonesiensis DSM 41759^T	91.41	90.01	43.80	0.028
S. cangkringensis DSM 41769^T	91.32	90.01	43.70	0.029
S. hygroscopicus subsp. *hygroscopicus* NBRC 13472^T	89.90	87.13	40.40	0.045
S. antioxidans MUSC 164^T	90.44	88.91	50.30	0.031
S. sioyaensis DSM 40032^T	78.76	71.20	22.60	0.094
S. decoyicus NRRL 2666^T	78.76	71.15	23.10	0.096
S. asiaticus DSM 41761^T	91.33	89.95	43.70	0.028
S. rimosus subsp. *rimosus* ATCC 10970^T	78.61	70.66	23.00	0.099
S. himastatinicus ATCC 53653^T	84.79	80.64	29.10	0.052
S. lydicus ATCC 25470^T	78.81	71.54	23.10	0.097
S. chattanoogensis NRRL ISP-5002^T	78.75	70.47	22.70	0.098

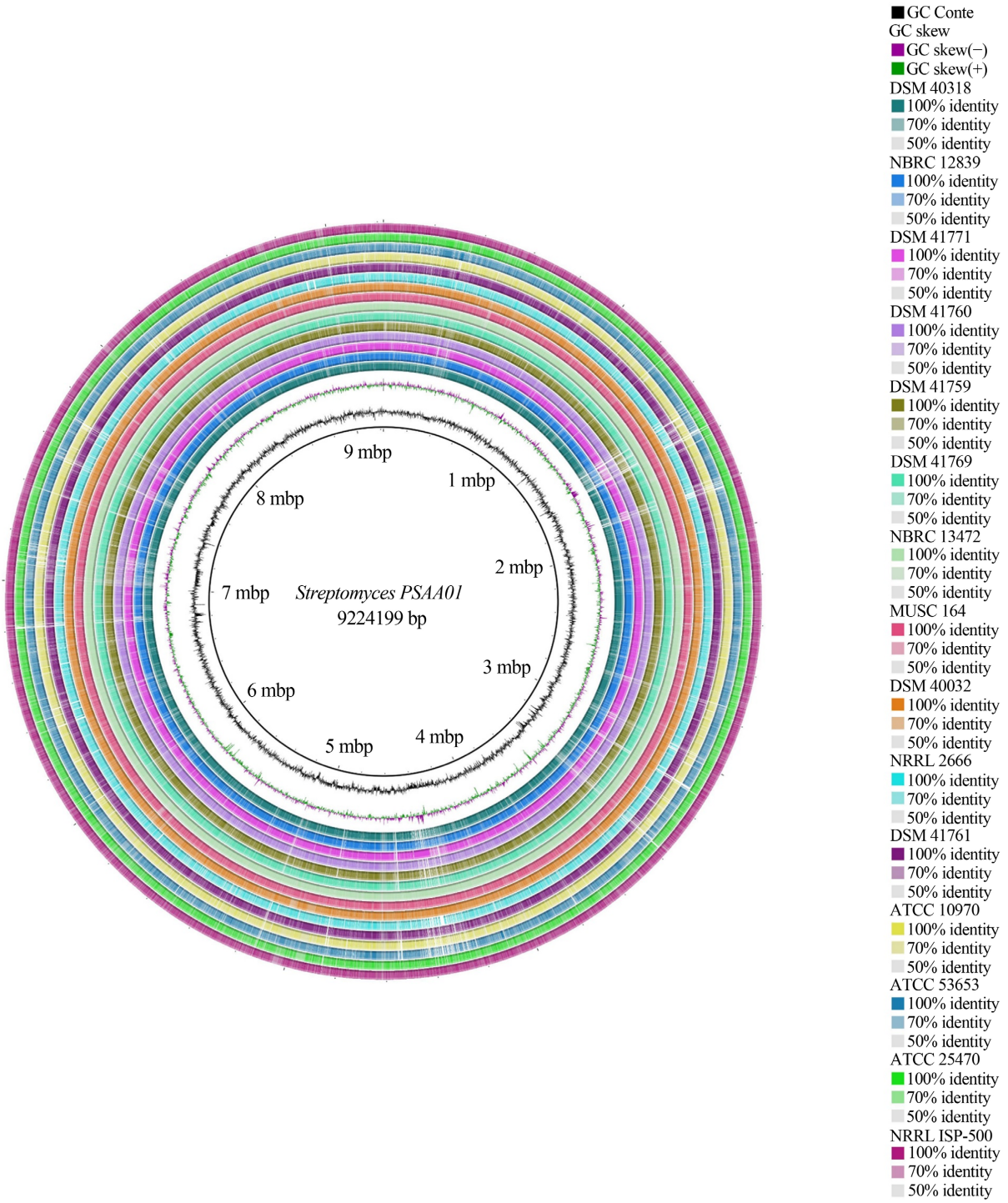

Figure 2. Comparative genomic map of the *Streptomyces* sp. PSA001 genome and the 15 other *Streptomyces* genomes. The BLASTN-based ring image was generated by the BLAST Ring Image Generator (BRIG) version 0.95 [32]. The innermost two rings show GC content (black) and GC skew (purple/green). The remaining 15 rings (from inside to outside) represent a BLASTN comparison with *S. melanosporofaciens* DSM 40318^T; *S. antimycoticus* NBRC 12839^T; *S. yatensis* DSM 41771^T; *S. rhizosphaericus* DSM 41760^T; *S. indonesiensis* DSM 41759^T; *S. cangkringensis* DSM 41769^T; *S. hygroscopicus* subsp. *hygroscopicus* NBRC 13472^T; *S. antioxidans* MUSC 164^T; *S. sioyaensis* DSM 40032^T; *S. decoyicus* NRRL 2666^T; *S. asiaticus* DSM 41761^T; *S. rimosus* subsp. *rimosus* ATCC 10970^T; *S. himastatinicus* ATCC 53653^T; *S. lydicus* ATCC 25470^T; *S. chattanoogensis* NRRL ISP-5002^T, respectively.

In the maximum-likelihood (ML) tree of PSAA01 and its related strains based on the concatenated housekeeping gene sequences (*i.e.*, *atpD-gyrB-recA-rpoB*), PSAA01 is found to share the same clade with *S. yatensis* DSM 41771^T, indicating that they might have descended from the same ancestor (Figure 3). The MLSA pairwise distances between strain PSAA01 and its closely related type strains, except *S. yatensis* DSM 41771^T (0.005), were all found to be above the cut-off point of 0.007, as recommended by [41] for novel species demarcation (Table 2 and Table S2). Furthermore, from the evaluation of the ANI values, it has been found that the values for *S. melanosporofaciens* DSM 40318^T (95.75%), *S. antimycoticus* NBRC 12839^T (95.52%), and *S. yatensis* DSM 41771^T (97.21%) are above the cut-off level (95% - 96%) recommended as the average nucleotide identity (ANI) criterion for interspecies identity [42] (Table 2). Additionally, the AAI values of PSAA01 and related type strains were significantly lower than the recommended cut-off point (95% - 96%) for species delineation [43], except *for S. yatensis* DSM 41771^T, which had an AAI value of 96.53%, which is above the range of the threshold (<95% - 96%) (Table 2). Similarly, except for *S. yatensis* DSM 41771^T (73.80%), the dDDH values for all the related strains were found to be lower than the recommended threshold of 70% for species delineation. The PSAA01 strain was also *found to be closely related to Streptomyces yatensis* by the TYGS-based phylogenomic analysis [37] (Table 2; Figure 4). From all these, we can conclude that the strain PSAA01 is no longer distinguishable from *S. yatensis* DSM 41771^T, but rather appears to be very closely related.

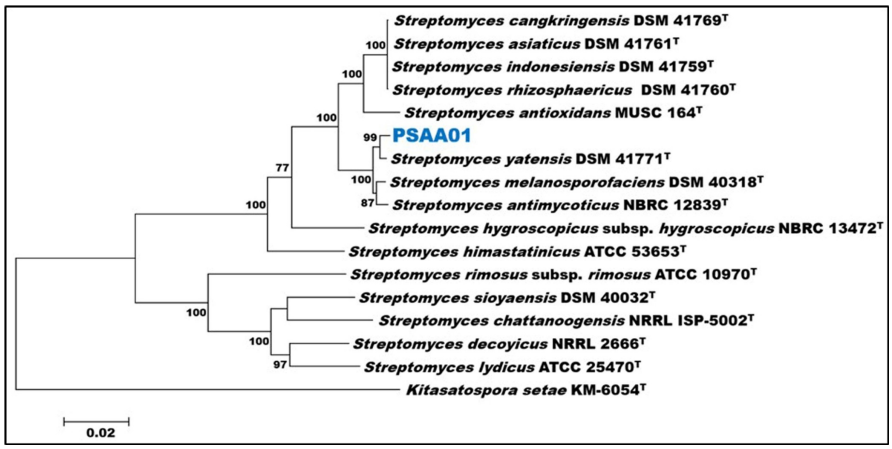

Figure 3. Multi-locus maximum likelihood (ML) phylogenetic tree based on concatenated housekeeping nucleotide sequences (*atpD-gyrB-recA-rpoB*) of strain PSAA01 and its closest type strains. Only bootstrap values >50% (expressed as % of 1000 replications) are shown at the nodes.

According to the results from antiSMASH version 6.0, the genome sequence of PSAA01 harbors 53 BGCs, of which 8 clusters code for NRPS, 25 clusters for PKS, 5 clusters for NRPS-like, 2 clusters for RiPP-like, 3 clusters for siderophore, 6 clusters for terpene, 2 clusters for ladderane, and 1 cluster each for PKS-like, arylpolyene, RRE-containing, hserlactone, ectoine, redox-cofactor, hglE-KS, lanthipep-

tide-class-ii, indole, and butyrolactone, respectively (Table S3). A high number of contigs in a draft genome can significantly affect the integrity, detection, and interpretation of biosynthetic gene clusters (BGCs), especially large modular systems such as polyketide synthases (PKS) and nonribosomal peptide synthetases (NRPS). Below is a structured discussion of the key impacts. A comparative distribution of BGCs present in the genomes of strain PSAA01 and the 10 closest studied strains is shown in Figure 5.

Classic RAST predicted a total of 8418 protein-encoding genes (PEGs), of which 2179 with known functions were classified into 22 groups, each associated with distinct biological roles. Out of these 2179 genes, 54 were related to the categories of virulence, disease, and defense; 8 were related to dormancy and sporulation; 7 were related to secondary metabolism; 413 were related to the synthesis of amino acids and their derivatives; 200 were related to cofactors, vitamins, prosthetic groups, and pigment synthesis (Figure 6).

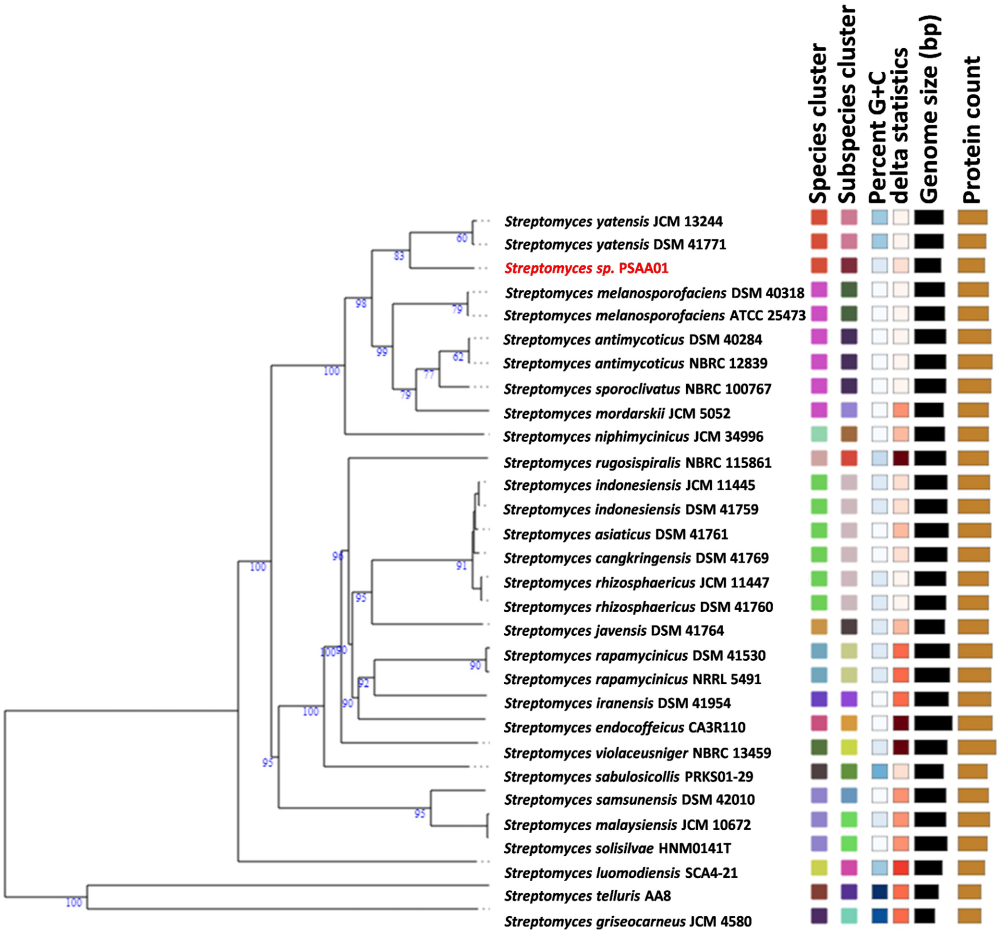

Figure 4. Phylogenetic tree based on genome sequences of the representative *Streptomyces* strains in the TYGS tree inferred with FastME 2.1.6.1 [33] from the Genome BLAST Distance Phylogeny approach (GBDP); distances were calculated from genome sequences. The branch lengths are scaled in terms of the GBDP distance formula d5. The numbers above the branches are GBDP pseudo-bootstrap support values >60% from 100 replications. The tree was rooted at the midpoint [36].

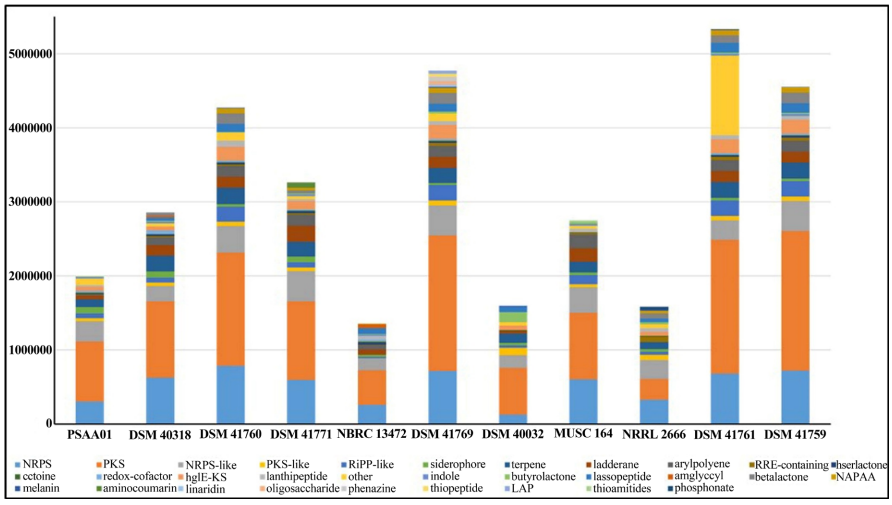

Figure 5. BGCs distribution determined by antiSMASH version 6.0 [37] in the genomes of PSAA01 and 10 related *Streptomyces* sp.

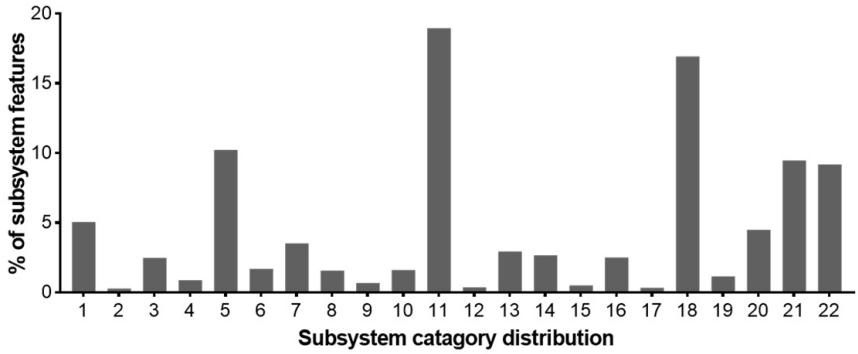

Figure 6. Number of identified protein-encoding genes of known function in *Streptomyces* sp. PSAA01 that are present in different subsystems according to the RAST server subsystem classification. Numbers are represented as: 1: Respiration; 2: Potassium metabolism; 3: Virulence, Disease and Defense; 4: Regulation and Cell signaling; 5: Protein metabolism; 6: Cell Wall and Capsule; 7: DNA Metabolism; 8: Membrane Transport; 9: Sulfur Metabolism; 10: Iron acquisition and metabolism; 11: Amino Acids and Derivatives; 12: Dormancy and Sporulation; 13: Stress Response; 14: RNA Metabolism; 15: Nitrogen Metabolism; 16: Metabolism of Aromatic Compounds; 17: Secondary Metabolism; 18: Carbohydrates; 19: Phosphorus Metabolism; 20: Nucleosides and Nucleotides; 21: Fatty Acids, Lipids, and Isoprenoids; 22: Cofactors, Vitamins, Prosthetic Groups, Pigments.

All of the highly similar compounds were chemically characterized and were present in the Minimum Information about a Biosynthetic Gene cluster (MIBIG) database, which was determined directly through antiSMASH [44]. From the most common BGCs, one showed the antimicrobial compound hygrocin A (NRPS/PKS-like), which exhibited 70% gene similarity with *Streptomyces* sp. LZ35, 51% gene similarity with *S. olivaceus,* and 58% gene similarity with *Streptomyces* sp. HK10576. The gene (hygrocin) cluster contains five core biosynthetic genes, some additional biosynthetic genes, and other regulatory genes. The structure of these core genes (1 - 5) is made up of three modules with the domains KS-

AT-DH-KR-CP, KS-AT-CP, and KS-AT-CP_Docking_C term; one module with the domain KS-AT-DH-ER-KR-CP_Docking_C term; one module with the domain KS-AT-DH-KR-CP_Docking_C term; one module with the domain KS-AT-DH-ER-KR-CP_Docking_C term, and two modules with KS-AT-CP, KS-AT-DH-KR-CP for the 1, 2, 3, 4, and 5 core genes, respectively. Polymer prediction by the aforementioned cluster is—(Me - ccmal - mal - ohmal) + (redemal) + (Me-redmal) + (Me - mal - ccmal), and the putative structure is given in **Figure 7(e)**. The biosynthetic gene cluster contains five core genes: *hgcA*, *hgcB*, *hgcC*, *hgcD,* and *hgcE*. Another antimicrobial compound, curamycin (T2PKS), has been identified in three species: *S. avermitilis* (shows 100% similarity), *S. cyaneus* (shows 100% similarity), and *S. collinus* (shows 85% similarity) (**Figure 7(b)**). Two other important biosynthetic gene clusters produce the antimicrobial compounds nigericin and mediomycin A, which show very high structural similarity to those from *Streptomyces* sp. PSAA01. This compound is a natural by-product and can be used as a potent antimicrobial agent [45] (**Figure 7**).

Several biosynthetic gene clusters (BGCs) have been identified that encode putative high-value metabolites, many of which appear to be uniquely associated with specific species. Although many similar products have already been characterized from different sources, there are some compounds that are specific to *Streptomyces* species. Here, we have considered the clusters that have more than 50% similarity. The pentamycin gene cluster has a single T1PKS and five domains. Although the pentamycin gene sequence is between *Streptomyces* sp. S816 and *S. chattanoogensis* has 86% similarity, the domain and structural organization are completely different (**Figure 8(a)**). The domain of *Streptomyces* sp. S816 is composed of five modules (KS-AT-CP, KS-AT-KR-CP, KS-AT-DH-KR-CP, KS-AT-DH-KR-CP, KS-AT-DH-KR-CP) + (KS-AT-KR-CP, KS-AT-KR-CP, KS-AT-KR-CP, KS-AT-KR-CP) + (KS-AT-DH-KR-CP) + (KS-AT-DH-KR-CP, KS-AT-DH-KR-CP) + (KS-AT-KR-CP, KS-AT-KR-CP-TE) and the domain of *S. chattanoogensis* also has five modules but a different composition (KS-AT-KR-CP, KS-AT-DH-KR-CP, KS-AT-DH-KR-CP, KS-AT-DH-KR-CP) + (A-CP-KS-AT-CP) + (KS-AT-DH-KR-CP, KS-AT-KR-CP, KS-AT-KR-CP, KS-AT-KR-CP, KS-AT-CP, KS-AT-KR-CP) + (KS-AT-DH-KR-CP) + (KS-AT-DH-KR-CP-TE). The macrolide group of antibiotics meridamycin isolated from *Streptomyces* sp. NRRL 30748 and from *S. arenicola* CNS-205 shows 80% gene sequence similarity, but they are structurally and module-wise very different (**Figure 8(b)**). Monensin is an ionophoric antibiotic used to treat bacterial, fungal, and parasitic infections. In addition to Monensin, Herboxidiene represents another class of antimicrobial compounds. Although they share more than 50% sequence similarity with PSAA01, their structural features are markedly different (**Figure 8(c)**, **Figure 8(d)**). Secondary metabolites identified in this study were predicted through genome mining based on biosynthetic gene cluster homology to characterized pathways. These compounds were not chemically extracted or structurally validated and therefore remain bioinformatically inferred products.

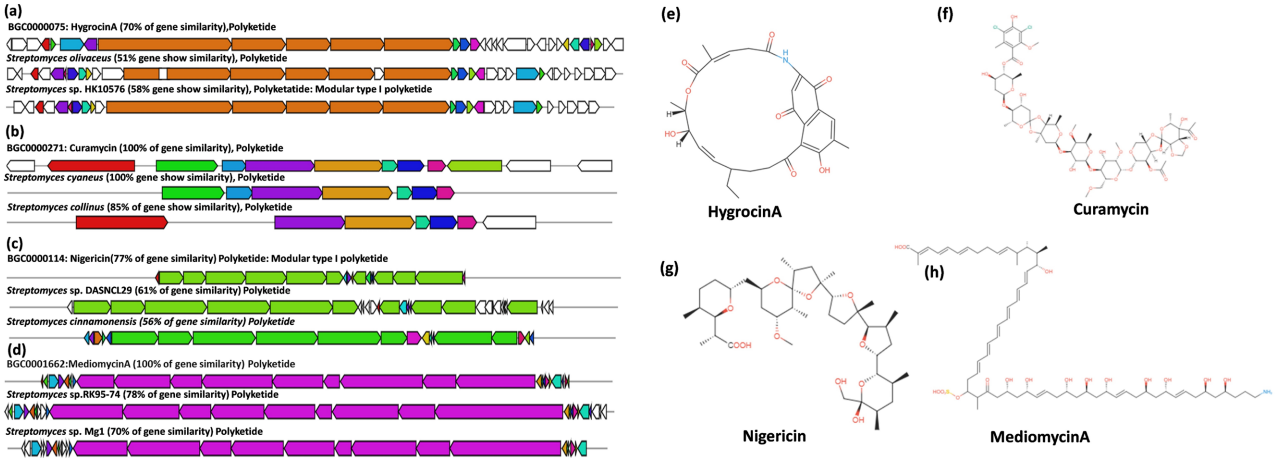

Figure 7. Highly similar antimicrobial gene clusters of *Streptomyces* species compared with known clusters in the antiSMASH database. Gene clusters for Hygrocin A (a), Curamycin (b), Nigericin (c), Mediomycin A (d); and the putative compounds produced by these clusters: Hygrocin A (e), Curamycin (f), Nigericin (g), Mediomycin A (h).

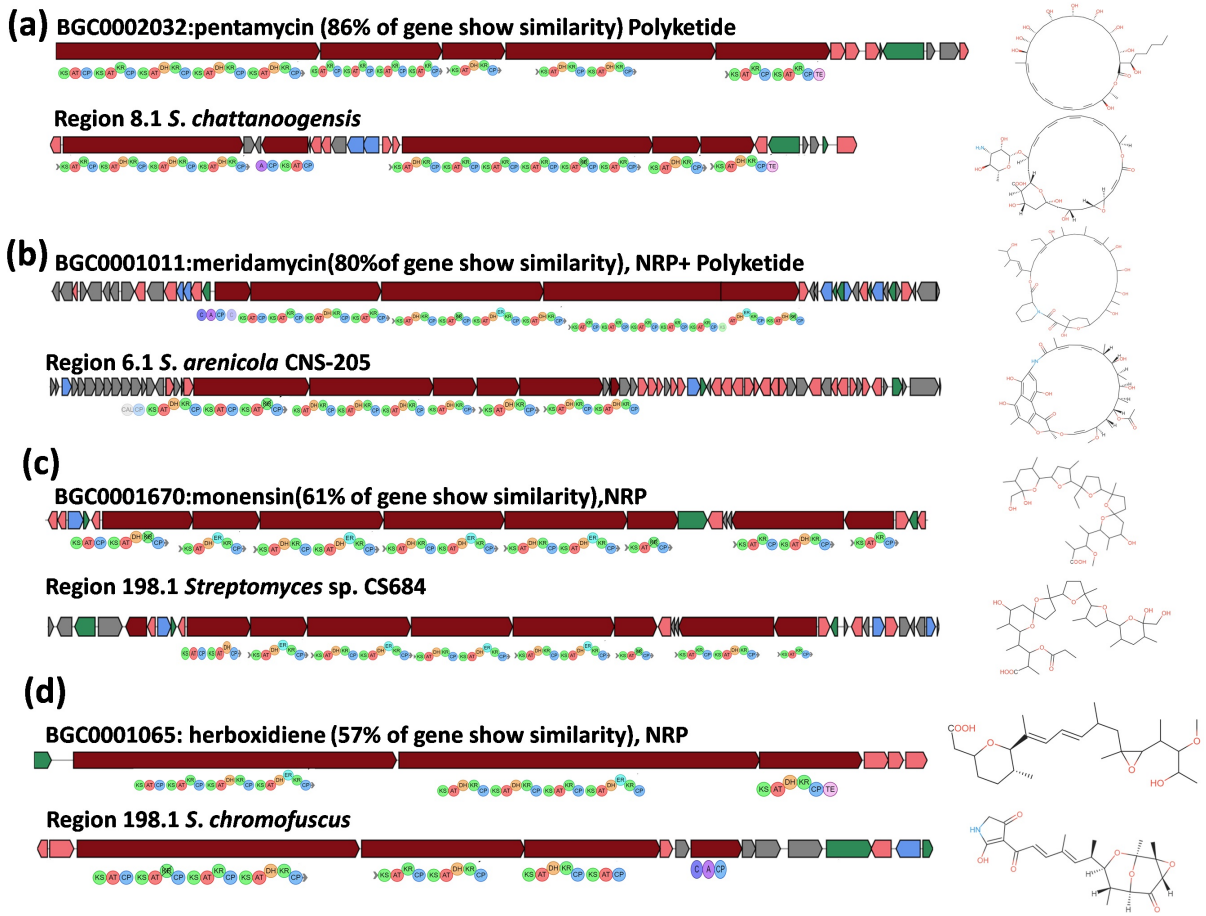

Figure 8. Species-specific clusters and their putative products from *Streptomyces* compared to the known clusters and their products from the antiSMASH database. (a) The biosynthetic cluster of pentamycin, a macrolide group of antibiotics, shows 86% gene similarity with *S. chattanoogensis*. (b) Another macrolide group of antibiotic, the meridamycin cluster, shows 80% similarity with *S. arenicola* CNS-205. (c) The monensin, an ionophoric group of antibiotic cluster, shows 61% gene similarity with *Streptomyces* sp. CS684. (d) The herboxidiene cluster shows 57% gene similarity with *S. chromofuscus*.

4. Conclusion

The strain *Streptomyces* sp. PSAA01 is a soil isolate from the Manas National Park, Assam, India, and has been found to be very closely related to *S. yatensis* DSM 41771^T. The strain exhibits amylase, catalase, lipase, cellulase, and urease activity. It can hydrolyze xylan and hypoxanthine. Strain PSAA01 can utilize L-rhamnose, D-sucrose, inositol, melezitose, maltose, D-galactose, raffinose, adonitol, pyruvate, and melibiose as carbon sources; it can also utilize amino acids like L-alanine, L-arginine, and L-asparagine. The strain has also been found to harbor diverse BGCs encoding various secondary metabolites like NRPS, PKS, NRPS-like, RiPP-like, siderophore, terpene, ladderane, PKS-like, arylpolyene, RRE-containing, hserlactone, ectoine, redox-cofactor, hglE-KS, lanthipeptide-class-ii, indole, and butyrolactone. According to the RAST server, 8418 protein-encoding genes (PEGs) were predicted in total, and among these, 2179 genes with known functions were divided into 22 categories, each with a unique biological role. The antibacterial chemical hygrocin A (similar to NRPS/PKS) exhibits gene similarity with *Streptomyces* sp. LZ35 of 70%, *S. olivaceus* of 51%, and *Streptomyces* sp. HK10576 of 58%, as determined by antiSMASH. Two further types of antibacterial compounds with sequences longer than 50% but extremely different structures are monensin and herboxidiene, which have also been found by the server.

Data Availability

The *16S rRNA* gene sequence and the genome of *Streptomyces* sp. PSAA01 were deposited in GenBank (NCBI) under the accession numbers MT829328 and JAKKUU000000000, respectively. Other data are available in this manuscript.

Author Contributions

Prasenjit Das: Conceptualization, performed bioinformatic and formal analysis, writing review; Biraj Sarkar: Performed bioinformatic analysis and writing-original draft; Dipanwita Patra: Performed partial bioinformatic data analysis; Sukhendu Mandal: Conceptualization, editing and supervision. All authors have read and approved the manuscript.

Consent for Publication

All the authors agree to submit the manuscript for publication.

Acknowledgements

P.D. is appreciative of financial support in the form of a scholarship from the University Grants Commission (UGC), Government of India. B.S. is appreciative of the University of Calcutta's financial support in the form of a fellowship.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] Jose, P.A. and Jha, B. (2016) New Dimensions of Research on Actinomycetes: Quest for Next Generation Antibiotics. *Frontiers in Microbiology*, **7**, Article No. 1295. https://doi.org/10.3389/fmicb.2016.01295

[2] Jiang, Y., Li, Q., Chen, X. and Jiang, C. (2016) Isolation and Cultivation Methods of Actinobacteria. In: *Actinobacteria—Basics and Biotechnological Applications*, InTech, 39-57. https://doi.org/10.5772/61457

[3] van der Meij, A., Worsley, S.F., Hutchings, M.I. and van Wezel, G.P. (2017) Chemical Ecology of Antibiotic Production by Actinomycetes. *FEMS Microbiology Reviews*, **41**, 392-416. https://doi.org/10.1093/femsre/fux005

[4] Prudence, S.M.M., Addington, E., Castaño-Espriu, L., Mark, D.R., Pintor-Escobar, L., Russell, A.H., *et al.* (2020) Advances in Actinomycete Research: An Actinobase Review of 2019. *Microbiology*, **166**, 683-694. https://doi.org/10.1099/mic.0.000944

[5] Parte, A.C., Sardà Carbasse, J., Meier-Kolthoff, J.P., Reimer, L.C. and Göker, M. (2020) List of Prokaryotic Names with Standing in Nomenclature (LPSN) Moves to the DSMZ. *International Journal of Systematic and Evolutionary Microbiology*, **70**, 5607-5612. https://doi.org/10.1099/ijsem.0.004332

[6] Suga, T., Kimura, T., Inahashi, Y., Iwatsuki, M., Nonaka, K., Také, A., *et al.* (2018) Hamuramicins a and B, 22-Membered Macrolides, Produced by an Endophytic Actinomycete Allostreptomyces Sp. K12-0794. *The Journal of Antibiotics*, **71**, 619-625. https://doi.org/10.1038/s41429-018-0055-x

[7] Bérdy, J. (2005) Bioactive Microbial Metabolites. *The Journal of Antibiotics*, **58**, 1-26. https://doi.org/10.1038/ja.2005.1

[8] Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H., *et al.* (2016) Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews*, **80**, 1-43. https://doi.org/10.1128/mmbr.00019-15

[9] Bentley, S.D., Chater, K.F., Cerdeño-Tárraga, A., Challis, G.L., Thomson, N.R., James, K.D., *et al.* (2002) Complete Genome Sequence of the Model Actinomycete Streptomyces Coelicolor A3(2). *Nature*, **417**, 141-147. https://doi.org/10.1038/417141a

[10] Bao, J., He, F., Li, Y., Fang, L., Wang, K., Song, J., *et al.* (2018) Cytotoxic Antibiotic Angucyclines and Actinomycins from the *Streptomyces* sp. Xzhg99t. *The Journal of Antibiotics*, **71**, 1018-1024. https://doi.org/10.1038/s41429-018-0096-1

[11] Iliĉ, S.B., Konstantinoviĉ, S.S., Todoroviĉ, Z.B., Laziĉ, M.L., Veljkoviĉ, V.B., Joković, N. and Radovanoviĉ, B.C. (2007) Characterization and Antimicrobial Activity of the Bioactive Metabolites in Streptomycete Isolates. *Mikrobiologiia*, **4**, 480-487.

[12] Law, J.W., Ser, H., Duangjai, A., Saokaew, S., Bukhari, S.I., Khan, T.M., *et al.* (2017) *Streptomyces colonosanans* Sp. Nov., a Novel Actinobacterium Isolated from Malaysia Mangrove Soil Exhibiting Antioxidative Activity and Cytotoxic Potential against Human Colon Cancer Cell Lines. *Frontiers in Microbiology*, **8**, Article No. 877. https://doi.org/10.3389/fmicb.2017.00877

[13] Schneider, O., Simic, N., Aachmann, F.L., Rückert, C., Kristiansen, K.A., Kalinowski, J., *et al.* (2018) Genome Mining of *Streptomyces* sp. YIM 130001 Isolated from Lichen Affords New Thiopeptide Antibiotic. *Frontiers in Microbiology*, **9**, Article No. 3139. https://doi.org/10.3389/fmicb.2018.03139

[14] Law, J.W., Chan, K., He, Y., Khan, T.M., Ab Mutalib, N., Goh, B., *et al.* (2019) Diversity of *Streptomyces* spp. from Mangrove Forest of Sarawak (Malaysia) and Screening of Their Antioxidant and Cytotoxic Activities. *Scientific Reports*, **9**, Article No. 15262. https://doi.org/10.1038/s41598-019-51622-x

[15] Ventola, C.L. (2015) The Antibiotic Resistance Crisis: Part 1: Causes and Threats. *P & T*: *A Peer-Reviewed Journal for Formulary Management*, **4**, 277-283.

[16] Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., *et al.* (2003) Complete Genome Sequence and Comparative Analysis of the Industrial Microorganism *Streptomyces avermitilis*. *Nature Biotechnology*, **21**, 526-531. https://doi.org/10.1038/nbt820

[17] Ohnishi, Y., Ishikawa, J., Hara, H., Suzuki, H., Ikenoya, M., Ikeda, H., *et al.* (2008) Genome Sequence of the Streptomycin-Producing Microorganism *Streptomyces griseus* IFO 13350. *Journal of Bacteriology*, **190**, 4050-4060. https://doi.org/10.1128/jb.00204-08

[18] Cruz-Morales, P., Vijgenboom, E., Iruegas-Bocardo, F., Girard, G., Yáñez-Guerra, L.A., Ramos-Aboites, H.E., *et al.* (2013) The Genome Sequence of *Streptomyces lividans* 66 Reveals a Novel tRNA-Dependent Peptide Biosynthetic System within a Metal-Related Genomic Island. *Genome Biology and Evolution*, **5**, 1165-1175. https://doi.org/10.1093/gbe/evt082

[19] El-Nakeeb, M.A. and Lechevalier, H.A. (1963) Selective Isolation of Aerobic Actinomycetes. *Applied Microbiology*, **11**, 75-77. https://doi.org/10.1128/am.11.2.75-77.1963

[20] Küster, E. and Williams, S.T. (1964) Selection of Media for Isolation of Streptomycetes. *Nature*, **202**, 928-929. https://doi.org/10.1038/202928a0

[21] Shirling, E.B. and Gottlieb, D. (1966) Methods for Characterization of *Streptomyces* Species. *International Journal of Systematic Bacteriology*, **16**, 313-340. https://doi.org/10.1099/00207713-16-3-313

[22] Maiti, P.K. and Mandal, S. (2019) Majority of Actinobacterial Strains Isolated from Kashmir Himalaya Soil Are Rich Source of Antimicrobials and Industrially Important Biomolecules. *Advances in Microbiology*, **9**, 220-238. https://doi.org/10.4236/aim.2019.93016

[23] Maiti, P.K., Das, S., Sahoo, P. and Mandal, S. (2020) *Streptomyces* sp SM01 Isolated from Indian Soil Produces a Novel Antibiotic Picolinamycin Effective against Multi Drug Resistant Bacterial Strains. *Scientific Reports*, **10**, Article No. 10092. https://doi.org/10.1038/s41598-020-66984-w

[24] Marmur, J. (1961) A Procedure for the Isolation of Deoxyribonucleic Acid from Micro-Organisms. *Journal of Molecular Biology*, **3**, 208-IN1. https://doi.org/10.1016/s0022-2836(61)80047-8

[25] Yoon, S., Ha, S., Kwon, S., Lim, J., Kim, Y., Seo, H., *et al.* (2017) Introducing Ezbiocloud: A Taxonomically United Database of 16S rRNA Gene Sequences and Whole-Genome Assemblies. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 1613-1617. https://doi.org/10.1099/ijsem.0.001755

[26] Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, **30**, 2725-2729. https://doi.org/10.1093/molbev/mst197

[27] Felsenstein, J. (1981) Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach. *Journal of Molecular Evolution*, **17**, 368-376. https://doi.org/10.1007/bf01734359

[28] Jukes, T.H. and Cantor, C.R. (1969) Evolution of Protein Molecules. In: *Mammalian Protein Metabolism*, Elsevier, 21-132. https://doi.org/10.1016/b978-1-4832-3211-9.50009-7

[29] Tamura K. (1992) Estimation of the Number of Nucleotide Substitutions When There Are Strong Transition-Transversion and G+C-Content Biases. *Molecular Biology and Evolution*, **4**, 678-687.

[30] Kimura, M. (1980) A Simple Method for Estimating Evolutionary Rates of Base Substitutions through Comparative Studies of Nucleotide Sequences. *Journal of Molecular Evolution*, **16**, 111-120. https://doi.org/10.1007/bf01731581

[31] Nei, M. and Kumar S. (2000) Molecular Evolution and Phylogenetics. Oxford University Press.

[32] Alikhan, N., Petty, N.K., Ben Zakour, N.L. and Beatson, S.A. (2011) BLAST Ring Image Generator (BRIG): Simple Prokaryote Genome Comparisons. *BMC Genomics*, **12**, Article No. 402. https://doi.org/10.1186/1471-2164-12-402

[33] Meier-Kolthoff, J.P. and Göker, M. (2019) TYGS Is an Automated High-Throughput Platform for State-of-the-Art Genome-Based Taxonomy. *Nature Communications*, **10**, Article No. 2182. https://doi.org/10.1038/s41467-019-10210-3

[34] Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., *et al.* (2008) The RAST Server: Rapid Annotations Using Subsystems Technology. *BMC Genomics*, **9**, Article No. 75. https://doi.org/10.1186/1471-2164-9-75

[35] Rodriguez-R, L.M. and Konstantinidis, K.T. (2016) The Enveomics Collection: A Toolbox for Specialized Analyses of Microbial Genomes and Metagenomes (No. e1900v1).

[36] Meier-Kolthoff, J.P., Auch, A.F., Klenk, H. and Göker, M. (2013) Genome Sequence-Based Species Delimitation with Confidence Intervals and Improved Distance Functions. *BMC Bioinformatics*, **14**, Article No. 60. https://doi.org/10.1186/1471-2105-14-60

[37] Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., Medema, M.H., *et al.* (2021) antiSMASH 6.0: Improving Cluster Detection and Comparison Capabilities. *Nucleic Acids Research*, **49**, W29-W35. https://doi.org/10.1093/nar/gkab335

[38] Skinnider, M.A., Johnston, C.W., Gunabalasingam, M., Merwin, N.J., Kieliszek, A.M., MacLellan, R.J., *et al.* (2020) Comprehensive Prediction of Secondary Metabolite Structure and Biological Activity from Microbial Genome Sequences. *Nature Communications*, **11**, Article No. 6058. https://doi.org/10.1038/s41467-020-19986-1

[39] Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S.Y., *et al.* (2019) Antismash 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Research*, **47**, W81-W87. https://doi.org/10.1093/nar/gkz310

[40] Saintpierre, D., Amir, H., Pineau, R., Sembiring, L. and Goodfellow, M. (2003) *Streptomyces yatensis* sp. Nov., a Novel Bioactive Streptomycete Isolated from a New-Caledonian Ultramafic Soil. *Antonie van Leeuwenhoek*, **83**, 21-26. https://doi.org/10.1023/a:1022906325397

[41] Rong, X. and Huang, Y. (2010) Taxonomic Evaluation of the *Streptomyces griseus* Clade Using Multilocus Sequence Analysis and DNA-DNA Hybridization, with Proposal to Combine 29 Species and Three Subspecies as 11 Genomic Species. *International Journal of Systematic and Evolutionary Microbiology*, **60**, 696-703. https://doi.org/10.1099/ijs.0.012419-0

[42] Richter, M. and Rosselló-Móra, R. (2009) Shifting the Genomic Gold Standard for the Prokaryotic Species Definition. *Proceedings of the National Academy of Sciences*, **106**, 19126-19131. https://doi.org/10.1073/pnas.0906412106

[43] Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D.R., da Costa, M.S., *et al.* (2018) Proposed Minimal Standards for the Use of Genome Data for the Taxonomy of Prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, **68**, 461-466. https://doi.org/10.1099/ijsem.0.002516

[44] Liu, B., Ge, B., Ma, J., Wei, Q., Khan, A.A., Shi, L., *et al.* (2018) Identification of wysPII as an Activator of Morphological Development in *Streptomyces albulus* CK-15. *Frontiers in Microbiology*, **9**, Article No. 2550. https://doi.org/10.3389/fmicb.2018.02550

[45] Yu, D., Xu, F., Valiente, J., Wang, S. and Zhan, J. (2013) An Indigoidine Biosynthetic Gene Cluster from *Streptomyces chromofuscus* ATCC 49982 Contains an Unusual IndB Homologue. *Journal of Industrial Microbiology and Biotechnology*, **40**, 159-168. https://doi.org/10.1007/s10295-012-1207-9

Supplementary

Table S1. Showing the biochemical test results of Strain PSAA01.

Biochemical analyses of strain PSAA01	
Carbon Utilization	
L-Rhamnose	+
D-Ribose	–
D-Sucrose	+
Inositol	+
Melezitose	+
Glucose	–
Maltose	+
Arabinose	–
Xylose	–
D-Galactose	+
Dulcitol	–
Raffinose	+
Adonitol	+
Pyruvic acid	+
Melibiose	+
Degradation/hydrolysis	
Gelatin	–
Xanthine	–
Hypoxanthine	+
Starch	+
Cellulose	+
Urea	+
Xylan	+
Tyrosine	–
Casein	–
Amino acid utilization	
L-Alanine	+
L-Arginine	+
L-Asperagine	+
Other biochemical tests	
H_2S production	–
Nitrate reduction	+
Catalase	+
Methyl red test	–
Voges-Proskauer test	–
Indole test	–
Lipase test	+

Table S2. Showing the pairwise distance of PSAA01 and other related strains based on concatenated sequences of the house-keeping gene sequences (*atpD-gyrB-recA-rpoB*). The MLSA pairwise distance was calculated on MEGA 6 by Kimura parameter 2 model. 1, PSAA01; 2, *S. melanosporofaciens* DSM 40318^T; 3, *S. antimycoticus* NBRC 12839^T; 4, *S. yatensis* DSM 41771^T; 5, *S. rhizosphaericus* DSM 41760^T; 6, *S. indonesiensis* DSM 41759^T; 7, *S. cangkringensis* DSM 41769^T; 8, *S. hygroscopicus* subsp. *hygroscopicus* NBRC 13472^T; 9, *S. antioxidans* MUSC 164^T; 10, *S. sioyaensis* DSM 40032^T; 11, *S. decoyicus* NRRL 2666^T; 12, *S. asiaticus* DSM 41761^T; 13, *S. rimosus* subsp. *rimosus* ATCC 10970^T; 14, *S. himastatinicus* ATCC 53653^T; 15, *S. lydicus* ATCC 25470^T; 16, *S. chattanoogensis* NRRL ISP-5002^T; 17, *Kitasatospora setae* KM-6054^T.

Strains	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1																	
2	**0.009**																
3	**0.009**	0.006															
4	**0.005**	0.008	0.009														
5	**0.029**	0.028	0.027	0.027													
6	**0.028**	0.028	0.027	0.027	0.000												
7	**0.029**	0.028	0.027	0.027	0.000	0.000											
8	**0.045**	0.044	0.044	0.044	0.044	0.044	0.044										
9	**0.031**	0.030	0.030	0.030	0.018	0.018	0.018	0.046									
10	**0.094**	0.093	0.092	0.093	0.096	0.096	0.097	0.095	0.099								
11	**0.096**	0.095	0.093	0.095	0.098	0.098	0.098	0.095	0.103	0.040							
12	**0.028**	0.028	0.027	0.027	0.000	0.000	0.000	0.044	0.018	0.096	0.098						
13	**0.099**	0.097	0.096	0.098	0.098	0.098	0.098	0.098	0.100	0.068	0.068	0.098					
14	**0.052**	0.050	0.049	0.051	0.046	0.046	0.046	0.046	0.051	0.094	0.093	0.046	0.093				
15	**0.097**	0.096	0.096	0.097	0.098	0.099	0.099	0.097	0.103	0.044	0.036	0.099	0.068	0.095			
16	**0.098**	0.096	0.095	0.097	0.100	0.101	0.101	0.095	0.104	0.043	0.044	0.101	0.071	0.096	0.047		
17	**0.139**	0.139	0.138	0.140	0.136	0.137	0.137	0.138	0.141	0.134	0.135	0.137	0.135	0.140	0.135	0.132	

Table S3. Showing the diverse BGCs present in the genome of PSAA01 as predicted by antiSMASH 6.0.

Cluster	Type	From	To	Most similar known cluster	Similarity (%)	MIBiG accession
Cluster 1	NRPS	350,607	376,969	polyoxypeptin	10	BGC0000457
Cluster 2	T1PKS	56,760	148,040	salinomycin	28	BGC0001244
Cluster 3	T1PKS	1	26,477	nigericin	77	BGC0000041
Cluster 4	Terpene	2118	20,604	hopene	61	BGC0000663
Cluster 5	NRPS, T1PKS	1	34,169	meridamycin	28	BGC0000457
Cluster 6	T1PKS	1	17,440	laidlomycin	34	BGC0000041
Cluster 7	T2PKS	1	17,197	Spore pigment	83	BGC0000215
Cluster 8	T1PKS	1	13,922	nigericin	61	BGC0001068
Cluster 9	Siderophore, T1PKS	1	56,211	apoptolidin	23	BGC0000946
Cluster 10	terpene	71,496	92,569	BE-43547A1, BE-43547A2, BE-43547B1, BE-43547B2B3, BE-43547C1, BE-43547C2	20	BGC0000649

Continued

Cluster 11	Ladderane, arylpolyene	1	11,894	atratumycin	39	BGC0001444
Cluster 12	T1PKS	1	10,771	lydicamycin	32	BGC0000056
Cluster 13	T1PKS	1	10,646	Divergolide A/B/C/D	17	BGC0000041
Cluster 14	Siderophore	20,354	32,267	ficellomycin	3	BGC0000946
Cluster 15	T1PKS	1	8384	primycin	21	BGC0000041
Cluster 16	T1PKS	1	6974	-	-	BGC0001273
Cluster 17	T1PKS	1	6377	-	-	BGC0000056
Cluster 18	terpene	1	6024	carotenoid	36	BGC0000633
Cluster 19	NRPS	1	5825	telomycin	8	BGC0000396
Cluster 20	RRE-containing	1	5631	Chromomycin A3	5	BGC0001295
Cluster 21	T1PKS	1	4037	-	-	BGC0000056
Cluster 22	Hserlactone	96,510	109,829	Heronamide A/B/C/D/E/F	8	BGC0001682
Cluster 23	T1PKS	22,648	70,204	Argimycin PI/argimycin PII, nigrifactin/argimycin PIV/ argimycin PV/argimycin PVI/argimycin PIX	43	BGC0000675
Cluster 24	terpene	6349	28,676	geosmin	100	BGC0001181
Cluster 25	NRPS-like, NRPS, T1PKS	48,196	107,409	Griseoviridin, fijimycin A	8	BGC0000417
Cluster 26	NRPS	1	1310	Rhizomide A/B/C	100	BGC0001833
Cluster 27	T1PKS	80,398	106,590	Hygrocin A/B	70	BGC0001858
Cluster 28	T1PKS	1	1004	-	-	BGC0000260
Cluster 29	NRPS	1	24,103	Ochronotic pigment	75	BGC0002075
Cluster 30	ladderane	68,783	104,375	atratumycin	31	BGC0000056
Cluster 31	siderophore	22,749	34,539	Desferrioxamin B	100	BGC0001478
Cluster 32	ectoine	7021	17,425	ectoine	100	BGC0002052
Cluster 33	terpene	33,884	54,897	2-methylisoborneol	100	BGC0000658
Cluster 34	Redox-cofactor	7722	29,861	Lankacidin C	13	BGC0001484
Cluster 35	NRPS	56,402	90,716	RP-1776	18	BGC0000389
Cluster 36	hglE-KS, T1PKS, RiPP-like	13,146	67,557	-	-	BGC0002031
Cluster 37	T3PKS	25,793	66,968	7-deoxypactamycin	16	BGC0000280
Cluster 38	NRPS	26,563	86,830	phthoxazolin	4	BGC0000389
Cluster 39	T1PKS	22,650	85,901	ECO-02301	57	BGC0000041
Cluster 40	T1PKS, lanthipeptide-class-ii	61,498	83,906	actagardine	9	BGC0001909
Cluster 41	RiPP-like	13,217	24,524	-	-	BGC0001803
Cluster 42	PKS-like	124,492	165,511	rustmicin	20	BGC0001911
Cluster 43	T1PKS, NRPS-like	2	77,306	meridamycin	36	BGC0000041
Cluster 44	other	1	26,299	mitomycin	20	BGC0000719

Continued

Cluster 45	NRPS-like	91,444	134,371	Echoside A/B/C/D/E	100	BGC0000340
Cluster 46	NRPS, T1PKS, other, NRPS-like	1	64,442	polyoxypeptin	32	BGC0000427
Cluster 47	indole	12,079	33,224	5-isoprenylindole-3-carboxylate β-D-glycosyl ester	61	BGC0001483
Cluster 48	T1PKS	1	45,723	Mediomycin A	46	BGC0001164
Cluster 49	terpene	5686	26,642	pristinol	100	BGC0001746
Cluster 50	T1PKS	1	31,609	fostriecin	23	BGC0000041
Cluster 51	T1PKS	1	30,407	apoptolidin	12	BGC000067
Cluster 52	butyrolactone	18,184	29,116	-	-	BGC0000848
Cluster 53	NRPS-like	1	29,431	Echoside A/B/C/D/E	11	BGC0000888