Genome Sequence of the Streptomyces Strain AgN23 Revealed Expansion and Acquisition of Gene Repertoires Potentially Involved in Biocontrol Activity and Rhizosphere Colonization

Damien D. Gayrard, Clément Nicolle, Marine Veyssière, Kévin Adam, Yves Martinez, Céline Vandecasteele, Marie Vidal, Bernard Dumas, Thomas Rey

PhytoFrontiers™
An Open Access Journal from The American Phytopathological Society

Research

Genome Sequence of the *Streptomyces* Strain AgN23 Revealed Expansion and Acquisition of Gene Repertoires Potentially Involved in Biocontrol Activity and Rhizosphere Colonization

Damien Gayrard[1,2] | Clément Nicolle[1] | Marine Veyssière[1] | Kévin Adam[1] | Yves Martinez[3] | Céline Vandecasteele[4] | Marie Vidal[4] | Bernard Dumas[1,†] | Thomas Rey[1,2,†] |

[1] Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, CNRS, UPS, Toulouse INP, Auzeville-Tolosane, France

[2] De Sangosse, Bonnel, Pont-Du-Casse, France

[3] CNRS, Plateforme Imagerie-Microscopie, Fédération de Recherche FR3450, Castanet-Tolosan, France

[4] INRAE, US 1426, GeT-PlaGe, Genotoul, Castanet-Tolosan, France

† Corresponding authors: T. Rey; reyt@desangosse.com, and B. Dumas; bernard.dumas@univ-tlse3.fr

Accepted for publication 22 January 2023.

Author contributions: D.G. performed the research and wrote the manuscript. M. Veyssière, C.N., K.A., Y.M., C.V., and M. Vidal performed the research. B.D. and T.R. designed the research and wrote the manuscript.

D. Gayrard and C. Nicolle contributed equally.

Funding

This work was funded by the Fond Unique Interministériels (NEOPROTEC project), the Fonds Européen de Développement Économique et Régional (FEDER), the Agence Nationale de la Recherche (LabCom BioPlantProtec ANR-14-LAB7-0001 and STREPTOCONTROL ANR-17-CE20-0030), and the Région Occitanie (projet GRAINE-BioPlantProducts). The Laboratoire de Recherche en Sciences Végétales (LRSV) belongs to the TULIP Laboratoire d'Excellence (ANR-10-LABX-41) and benefits from the "École Universitaire de Recherche (EUR)" TULIP-GS (ANR-18-EURE-0019). Work performed in the GeT core facility, Toulouse, France (https://get.genotoul.fr) was supported by the France Génomique National infrastructure, funded as part of the "Investissement d'Avenir" program managed by the Agence Nationale de la Recherche (contract ANR-10-INBS-09) and by the GET-PACBIO program (FEDER Programme opérationnel FEDER-FSE MIDI-PYRENEES ET GARONNE 2014-2020). D. Gayrard was funded by Agence Nationale de la Recherche Technique, with the Convention Industrielle de Formation par la Recherche and Association Nationale de la Recherche et de la Technologie (grant number 2016/1297). C. Nicolle was funded by the Ministère de l'Enseignement Supérieur et de la Recherche (PhD grant).

e-Xtra: Supplementary material is available online.

The following information may be seen as competing interests. B.D. is one of the inventors of the patent WO2015044585A1 relating to the use of AgN23 in agriculture. T.R. and D.G. are full-time researchers at the AgChem company De Sangosse (Pont-Du-Casse, France), which registers and markets crop-protection products and owns the patent WO2015044585A1. M. Veyssière, M. Vidal, C.N., K.A., Y.M., and C.V. do not declare any competing interests.

Abstract

Streptomycetes are gram-positive actinobacteria largely represented in the plant root microbiota. The genetic determinants involved in the adaption of *Streptomyces* in the rhizosphere environment are mostly unknown but can rely on the ability to release phytohormones, degrade plant cell-wall polysaccharides, and produce specialized metabolites notably involved in microbial competition. Here, we sequenced the genome of the rhizospheric and plant defense-stimulating strain *Streptomyces* sp. AgN23. We found that it belongs to the soil- and plant root-dwelling *S. violaceusniger* clade. The genome annotation of AgN23 revealed the genetic potential of the bacterium to degrade the plant cell wall with a large repertoire of carbohydrate degrading enzymes, to synthesize auxin, a major regulator of plant development, and to produce antimicrobials (rustmicin, mediomycin, niphimycin, nigericin) and plant bioactive compounds (nigericin, echosides, elaiophylin) through a set of biosynthetic gene clusters. We also found that these genomic features are well conserved among members of the *S. violaceusniger* clade. In addition, AgN23 displays original events of biosynthetic gene cluster acquisitions and losses, which may account for its interaction with plants. Taken together, our work supports the hypothesis that evolution of a large set of conserved hydrolytic enzymes directed against plant polymers and specialized metabolite repertoires can mediate the adaptation of *S. violaceusniger* strains to the rhizospheric ecological niches.

Keywords: biosynthetic gene clusters, carbohydrate-active enzymes, long-read sequencing, phylogenomic, specialized metabolites, *Streptomyces*, transcriptome

Streptomycetes are aerobic and gram-positive actinobacteria forming branched vegetative mycelia before developing aerial hyphae-bearing spores (Bush et al. 2015). These bacteria have received considerable attention from a biotechnological point of view, notably regarding their enzymatic repertoire (Book et al. 2016) and in the drug discovery

field, leading to the structure elucidation of more than 6,000 specialized metabolites (Barka et al. 2016; Moumbock et al. 2021). Their tremendous ability to produce antimicrobial compounds relies on the wealth and diversity of biosynthetic gene clusters (BGCs) found in their genomes (Lee et al. 2020). The recent soaring of microbial metabarcoding approaches have highlighted *Streptomyces* species' prominent abundance in plant root microbiota (Fitzpatrick et al. 2018; Lundberg et al. 2012). Plants release 11 to 40% of total fixed carbon (C) through photosynthesis into the surrounding soil root environment. Thus, the rhizosphere is notably enriched in complex carbohydrate compounds, meaning that access to this resource depends on the ability to degrade them. Other factors that can favor the adaptation of microbial species to this environment rely on the competition with other microbial strains. Because several of the metabolite streptomycetes produced display strong antimicrobial activities, *Streptomyces* recruitment in the rhizosphere and inside root tissues presumably help them to fight against competitors and protect the plant from pathogens. Consequently, *Streptomyces*-based products have been developed for agriculture, but these strains only cover a subset of plant-colonizing *Streptomyces* species (Hamedi and Mohammadipanah 2015; Vurukonda et al. 2018). Notwithstanding, little is known regarding the biological function of these bacteria in the plant environment and the gene families involved in their adaptation to this ecological niche (Rey and Dumas 2017; Viaene et al. 2016). More than 100 fully assembled genome sequences of *Streptomyces* strains are currently available, paving the way for genome-based phylogeny and comparisons of BGC content across *Streptomyces* clades (Kautsar et al. 2021a; Medema 2021; Tracanna et al. 2017). This knowledge opens inroads to rationalize and investigate the potential use of *Streptomyces* strains in agriculture.

Here, we report a gapless assembly of *Streptomyces* sp. AgN23 (AgN23), previously isolated from grapevine rhizosphere and identified as a strong inducer of plant defenses (Vergnes et al. 2019). We used this high-quality assembly to position AgN23 in the *S. violaceusniger* genomospecies. We then dissected the original features in the CAZyome and BGC content of AgN23 and other *S. violaceusniger* strains and shed light on specificities in the specialized metabolism and carbohydrate degradation capabilities of this lineage. This comparative genomic approach also led us to identify original features in the genome of AgN23, such as putative BGC losses and acquisition by horizontal gene transfer, which could account for the bacteria adaptation in the rhizospheric niche.

MATERIALS AND METHODS

AgN23 cultivation and high molecular weight DNA extraction

The AgN23 strain was cultivated as described previously (Errakhi et al. 2007; Vergnes et al. 2019). In brief, the strain was grown on solid modified Bennet medium (D-glucose 10 g/liter; soybean peptones 2.5 g/liter; yeast extract 1.5 g/liter; agar 16 g/liter) or International Streptomyces Project media ISP2, ISP3, ISP4, and ISP5 (Shirling and Gottlieb 1966; Kieser et al. 2000). To produce the spore inoculum, we incubated Bennet plates for 2 weeks at 22°C in the darkness before filling them with 10 ml of sterile water. The mycelium was scraped with a spreader, and the resulting solution was filtered in a 50-ml syringe filled with nonabsorbent cotton wool. For DNA extraction, the AgN23 mycelium was grown at 28°C and under 250 rpm in 250-ml Erlenmeyer flasks containing 50 ml of liquid Bennet. Approximately 100 mg of AgN23 pellets was collected by centrifugation at 11,000 × *g* and flash frozen in liquid nitrogen. Genomic DNA was isolated using the Nucleobond RNA/DNA kit (Macherey-Nagel) according to the manufacturer's instructions.

Library preparation for genome sequencing

Library preparation and sequencing were performed at the GeT-PlaGe core facility (Castanet-Tolosan), according to the PacBio instructions, with a 15-kb size cutoff ("20 kb Template Preparation Using BluePippin Size Selection system"). At each step, DNA was quantified using the Qubit dsDNA HS Assay Kit (Life Technologies). DNA purity was tested using the NanoDrop (Thermo Fisher), and size distribution and degradation were assessed using the High-Sensitivity Large Fragment 50-kb Analysis Kit with a fragment analyzer (AATI). Purification steps were performed using 0.45× AMPure PB beads (PacBio). A total of 10 µg of DNA was purified then sheared at 40 kb using the Megaruptor system (Diagenode). Single-molecule real-time (SMRT) sequencing was performed using SMRTBell Template Prep Kit 1.0 (PacBio), and a DNA and END damage repair step was performed on 5 µg of sample. Then, blunt hairpin adapters were ligated to the library. The library was treated with an exonuclease cocktail to digest unligated DNA fragments. A size selection step using a 10-kb cutoff was performed on the BluePippin Size Selection system (Sage Science) with 0.75% agarose cassettes, Marker S1 high Pass 15-20 kb. Conditioned Sequencing Primer V2 was annealed to the size-selected SMRTbell. The annealed library was then bound to the P6-C4 polymerase using a 10:1 ratio of polymerase to SMRTbell. Then, after a magnetic bead-loading step (OCPW), SMRTcell libraries were sequenced on two SMRT cells on an RSII instrument at 0.18 to 0.23 nM with a 360-min movie. The initially generated raw sequencing reads were evaluated in terms of the average quality score at each position, GC content distribution, quality distribution, base composition, and other metrics. Sequencing reads with low quality were also filtered out before the genome assembly and annotation of gene structure. Finally, microbial DNA potential contamination was excluded after comparison by BLAST of the draft assembly of the first SMRT cell against a 16S ribosomal RNA sequences data bank (Bacteria and Archaea).

Genome assembly, annotation, and comparative genomics

The subreads were assembled with PacBio's SMRT analysis software version 2.3.0 using default settings with a minimum subread length of 3 kb to exclude smaller sequenced reads and a read score of better than 0.8 to enrich in reads with a low error rate. The single unitig obtained by long-read sequencing was corrected with Mi-Seq (Illumina) data using Pilon (version 1.21), resulting in 165 substitutions and two deletions of 44 and 5 bases. This final genome assembly was retained for subsequent analysis. The gene annotation was performed on the MicroScope Microbial Genome Annotation & Analysis Platform (Vallenet et al. 2020). Genome completion was checked with CheckM and BUSCO (Parks et al. 2015; Simão et al. 2015). CAZy genes annotation was obtained by running the HMMER tool (e-value < 1e-15, coverage > 0.35) on the dbCAN2 meta server (Yin et al. 2012). For the annotation of auxin biosynthesis genes, we retrieved the protein sequences from UniProt corresponding to the KEGG ontologies associated with the different biosynthetic steps described in Zhang et al. (2019). These sequences were systematically used as queries for BLASTx analysis against AgN23 chromosome sequences. The position on the AgN23 chromosome of any hits showing more than 70% identity and 40% coverage with one query was used to determine the corresponding AgN23 gene model (Zhang et al. 2019). antiSMASH 6.0 was used to

detect BGC-containing regions in the AgN23 chromosome and annotate detected sequences based on the MIBiG 2.0 repository (Blin et al. 2019; Epstein et al. 2018). The comparative genomics were performed with BiG-SCAPE using default parameters with the antiSMASH 6.0-predicted region-containing BGCs as input data to produce similarity networks (Navarro-Muñoz et al. 2020). Complementarily, the BiG-SLiCE software (1.0.0) was used to identify BGCs showing similarity to AgN23 regions annotated by antiSMASH among the 1,225,071 BGCs stored in the BiG-FAM database (Kautsar et al. 2021b). BiG-SLICE was used with standard parameters, including the arbitrary clustering threshold (T = 900.0).

Phylogenetic and phylogenomic analysis

The phylogenetic analysis of AgN23 was performed using genome sequences previously assigned to the *S. violaceusniger* clade (*Streptomyces* sp. M56, *Streptomyces* sp. 11-1-2, *S. rapamycinicus* NRRL 5491, *S. malaysiensis* DSM4137, *S. antimycoticus* NBRC 100767, *S. sabulosicollis* PRKS01-29, *S. albiflaviniger* DSM 41598, *S. rhizosphaericus* NRRL B-24304, *S. javensis* DSM 41764, and *S. milbemycinicus* NRRL) (Díaz-Cruz et al. 2022; Goodfellow et al. 2007; Klassen et al. 2019; Kumar et al. 2007; Kusuma et al. 2021). In addition, we introduced as input the genomes of the model strain *S. coelicolor* A(3)2 and two biocontrol strains, *S. lydicus* WYEC108 and *S. griseoveridis* K61, as well as *Frankia alni* ACN14a as an outgroup (Supplementary Information S1). The *Streptomyces* phylogeny was built using a multi-locus sequence typing (MLST) strategy with the de novo workflow of autoMLST (Alanjary et al. 2019). The concatenated alignment of 85 single-copy conserved genes was built using the Fast alignment mode (MAFFT FFT-NS-2), and the IQ-TREE Ultrafast Bootstrap analysis was performed with 1,000 replicates (Supplementary Information S2) (Alanjary et al. 2019).

Finally, a whole-genome-based phylogenetic tree based on fully assembled chromosomes of isolates from the *S. violaceusniger* clade was inferred with FastME 2.1.6.1 using the distance calculated by genome blast distance phylogeny (GBDP) available on the TYGS platform (Lefort et al. 2015; Meier-Kolthoff and Göker 2019). The branch lengths were scaled in terms of GBDP distance formula d5. The tree had an average branch support of 96.9% after 100 bootstrap replicates and was rooted at midpoint.

Library preparation for transcriptome sequencing and expression analysis

AgN23 was cultivated in liquid Bennett medium at 28°C for 48 h to reach the exponential phase of the growth. Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. Three replicates were prepared for the construction of the libraries. rRNA depletion was performed using the Ribozero rRNA Removal Kit (Illumina). RNA sequencing library preparation was completing using the NEBNext Ultra RNA Library Prep Kit for Illumina following the manufacturer's recommendations (NEB). RNAs were fragmented to generate double-stranded cDNA and subsequently ligated to a universal adapter, followed by index addition and library enrichment with limited-cycle PCR. Sequencing libraries were validated using the Agilent Tapestation 4200 (Agilent Technologies) and quantified by the Qubit 2.0 Fluorometer (Invitrogen), as well as quantitative PCR (Applied Biosystems). RNA-Seq experiments were performed on an Illumina HiSeq4000 using a paired-end read length of 2 × 150 bp with the Illumina HiSeq4000 sequencing kits. The mapping and statistical analysis were performed using the bioinformatics pipeline implemented in the MicroScope Platform (Vallenet et al. 2020) (Supplementary Information S3). Raw reads of each sample (R1 fastq files from the paired-end run) were mapped onto the AgN23 reference genome with BWA-MEM (v.0.7.4) (Li 2013). An alignment score equal to at least half of the read was required for a hit to be retained. SAMtools (v.0.1.8) was then used to extract reliable alignments with a mapping quality (MAPQ) ≥ 1 from SAM formatted files (Li et al. 2009). The number of reads matching each genomic object of the reference sequence was subsequently counted with the toolset BEDTools (v.2.10.1) (Quinlan and Hall 2010). The mean transcripts per million (TPM) for each gene was then calculated from the three independent samples (Supplementary Information S3). The BGC expression levels were determined by calculating the mean expression of biosynthetic genes for each BGC as defined by antiSMASH and listed in Supplementary Information S4. A mean expression for each BGC across the three biological repetitions was then determined.

Scanning electron microscopy

The observation of AgN23 mycelium and spore development by scanning electron microscopy was performed on a Quanta 250 FEG (FEI). Agar plugs of 2-week-old AgN23 cultures were placed on micrometric platen, frozen in liquid nitrogen, and finally metallized with platinum. The samples were observed microscopically at an accelerating voltage of 5.00 kV.

RESULTS AND DISCUSSION

Chromosome-scale assembly of AgN23 and genome-based taxonomic assignation to the *S. violaceusniger* clade

We performed PacBio RSII long-read sequencing and obtained a linear chromosome of 10.86 Mb for AgN23. This genome sequence was polished using Illumina MiSeq sequences to produce the final assembly. The final assembly was uploaded to the MicroScope Platform (Vallenet et al. 2020). The genome annotation retrieved 10,514 protein coding sequences, 10,458 of them being supported by RNA-seq with expression levels comprising 0.09 to 20,661 TPM and mean and median values of 94 and 21 TPM, respectively (Supplementary Information S3) (Table 1). A checkM analysis was performed using 455 genomes and 315 lineage-specific markers and validated the completeness of the assembly and the annotation (Table 1) (Parks et al. 2015). Complementarily, the completeness was also supported by a BUSCO analysis that detected 99.7% of the genes that are expected in the genome of an actinobacterium (Table 1) (Simão et al. 2015).

TABLE 1

Summary of the assembly and the annotation of *Streptomyces* sp. AgN23 complete chromosome obtained by PacBio and Illumina sequencing

Assembly statistics	*Streptomyces* sp. AgN23
Genome size (bp)	10,858,739
Coverage (PacBio)	77×
G + C content (%)	70.9
Number of protein-coding genes	10,514
CheckM completeness	100%
Contamination	1.84%
BUSCO	451/452
CAZymes	278
Number of biosynthetic gene clusters (antiSMASH 6.0)	47

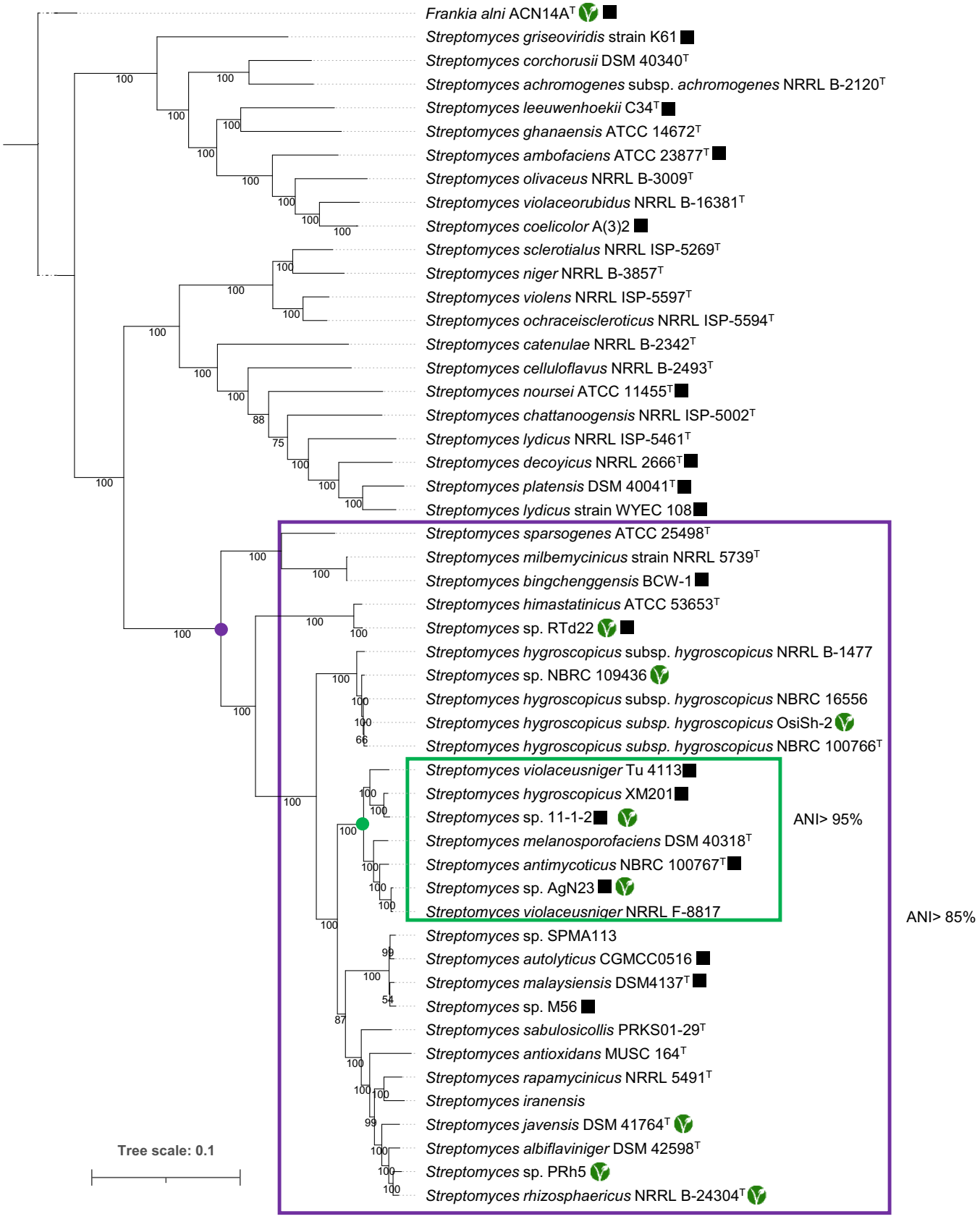

FIGURE 1

Multilocus sequence typing assigned AgN23 to the *Streptomyces violaceusniger* clade. Phylogenetic tree based on the multiple alignment of 85 single-copy homologous genes selected from genomic sequences with AutoMLST. The green node highlights the isolates considered to be from the same species (average nucleotide identity [ANI] > 95%). The black node highlights the clade formed by isolates with ANI > 90% as compared with AgN23. The black squares highlight AgN23 and the 18 strains that were used for the biosynthetic gene cluster (BGC) conservation study. The green logo indicates plant-isolated strains. *Frankia alni* ACN14a was used as the outgroup, bootstrap = 100.

More than 500 *Streptomyces* species have been described based on their 16S rRNA sequences. Previous sequencing of AgN23 16S rRNA showed the strongest conservation with representatives of the *S. violaceusniger* clade, most notably *S. castelarensis* (Vergnes et al. 2019). However, lack of variation in 16S rRNA may confound strains belonging to different species (Labeda et al. 2012). Recent development of long-read technologies and massive sequencing of *Streptomyces* have leveraged genome-based phylogenies (Lee et al. 2020). Thus, we decided to consolidate the taxonomic affiliation of AgN23 with the gapless assembly of the AgN23 chromosome. We performed an autoMLST approach based on publicly available *Streptomyces* sequences. In brief, 85 conserved housekeeping genes showing neutral dN/dS were concatenated and aligned as a basis for tree building (Fig. 1, Supplementary Information S2) (Alanjary et al. 2019).

As a result, using AgN23 as the query, we affiliated the strain to a clade containing six other strains showing an average nucleotide identity (ANI) higher than 95% with AgN23, which is conventionally considered a threshold for species delimitation (Jain et al. 2018). It contains the closely related *S. melanosporofaciens* DSM 40318, *S. antimycoticus* NBRC 100767, *S. violaceusniger* NRRL F-8817, *S. violaceusniger* Tu 4113, *S. hygroscopicus* XM201, and *Streptomyces* sp. 11-1-2 (Caicedo-Montoya et al. 2021; Komaki and Tamura 2020). In total, 28 isolates harbored an ANI > 85% with *Streptomyces* sp. AgN23. Notably, this group contains representative species of the *S. violaceusniger* clade, such as *S. hygroscopicus*, *S. sparsogenes*, *S. malaysiensis* (Goodfellow et al. 2007), *S. himastatinicus*, *S. rapamycinicus*, and other close species *S. autolyticus* (Yin et al. 2017), *S. antioxidans* (Ser et al. 2016), *S. iranensis* (Hamedi et al. 2010), *S. albiflaviniger*, *S. rhizosphaericus*, *S.* sabulosicollis, *S. javensis*, and *S. milbemycinicus* (Kusuma et al. 2021) (Supplementary Information S5). Interestingly, *Streptomyces* strain RT-d22 (Chagas et al. 2016), *Streptomyces* sp. strain PRh5 (Yang et al. 2014), *S. hygroscopicus* Osish-2 (Cao et al. 2021; Gao et al. 2021; Xu et al. 2017, 2019; Zeng et al. 2018), *Streptomyces* sp. NBRC 109436 (Komaki et al. 2016), and *S. rhizosphaericus* NRRL B-24304 (Goodfellow et al. 2007) were isolated from the rhizosphere of diverse plants across the world. Similarly, *Streptomyces* sp. 11-1-2 was isolated from the scab of a potato tuber (Bown and Bignell 2017; Díaz-Cruz et al. 2022). This suggests frequent interactions with plants of the strains from the *S. violaceusniger* clade. Finally, morphological characterization of AgN23 on a range of ISP media confirmed that AgN23 displays typical phenotypes of the *S. violaceusniger* clade regardless of the ISP media tested, with a whitish colony that turns gray during the sporulation process, resulting in the formation of spiraled chains of rugose-ornamented spores (Fig. 2) (Goodfellow et al. 2007; Kumar and Goodfellow 2008; Kumar et al. 2007).

AgN23 exhibits a wide repertoire of CAZymes related to plant cell wall degradation

Streptomycetes display extensive abilities to degrade polysaccharides, notably plant polysaccharides derived from the cell wall, such as cellulose, hemicellulose, and pectins based on their rich repertoires of carbohydrate-active enzymes (CAZymes) (Book et al. 2016; Yin et al. 2012). We annotated the AgN23 CAZymes with dbCAN2 and found 278 CAZymes representing 2.6% of AgN23 coding sequences (Supplementary Information S6). In *Streptomyces*, the amount of CAZymes typically ranges from 1.5 to 3% of the proteome (Book et al. 2016). For example, CAZymes represents 2.9% of coding sequences of the cellulolytic *Streptomyces* sp. SirexAA-E; thus, the AgN23 genome harbors a rather large repertoire of CAZymes among the *Streptomyces* genera (Book et al. 2016). To investigate whether AgN23 displays an expanded repertoire in particular CAZymes families, we compared it with the repertoire of CAZymes from the highly cellulolytic *Streptomyces* sp. SirexAA-E and the 50 genomes used in the phylogenomic analysis, which comprised *S. violaceusniger* representatives, as well as outgroup strains (Fig. 3).

We found that AgN23 and other members of the *S. violaceusniger* clade possess two PL9 polysaccharide lyases involved in pectin degradation that are absent from 20 of the 22 *Streptomyces* outgroup genomes. Similarly, the AgN23 genome harbors seven GH6 (glycoside hydrolases) that are endoglucanases related to cellulose degradation. Such a large repertoire of GH6 is frequent in the *S. violaceusniger* clade (ANI > 85%), whereas zero to four genes were found in the outgroup strains tested. A similar observation could be drawn for GH43 and GH95 involved in hemicellulose catabolism. Strikingly, the AgN23 genome encodes nine GH43, whereas the other *S. violaceusniger* spp. possess up to 15 of these genes, and eight strains among the outgroup (ANI < 85%) had none.

Finally, we analyzed the expression pattern of AgN23 CAZymes, consistently with the whole-genome expression analysis, and only three CAZymes showed no expression in the Bennett liquid culture dataset (Supplementary Information S6). Strikingly, two genes of the PL9 family (AgN23_3063 and AgN23_1031), which is largely restricted to the *S. violaceusniger* group, showed a strong expression at 526 and 279 TPM, respectively. These values rank the two genes among the 10% most expressed CAZymes of AgN23. By contrast, the PL1 representatives of AgN23_3657 and AgN23_1017 have 9 and 8 TPM, respectively. Regarding the GH6 and GH43 families, we no-

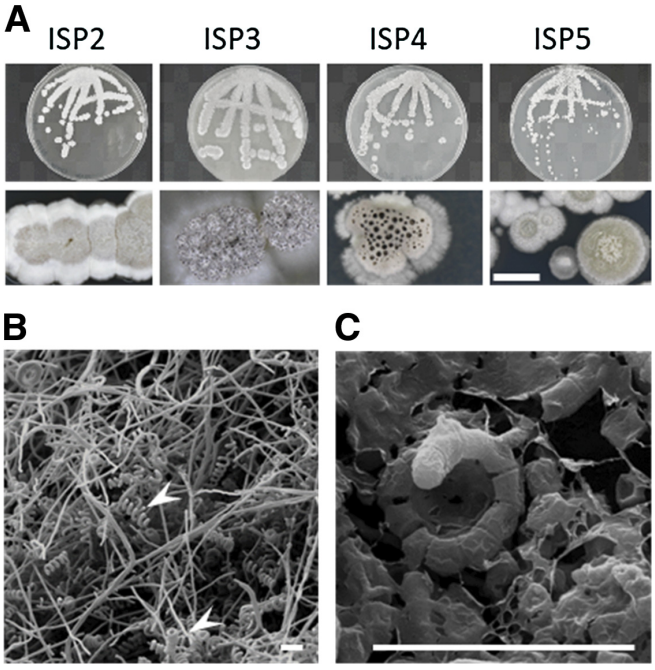

FIGURE 2
AgN23 harbors typical phenotypic features of the *Streptomyces violaceusniger* clade. **A,** AgN23 forms a white mycelium that turns gray at the onset of sporulation as observed on a range of ISP media; the Petri plates are 9 cm in diameter, and the scale bar is 1 mm. **B,** Scanning electron microscopy observation of spiraled chains of spores (white arrow) formed by AgN23 (scale bar = 6 μm). **C,** Scanning electron microscopy observation of AgN23 spores showing a rugose-ornamented surface (scale bar = 6 μm).

ticed that they are dominated by the expression of one single gene, respectively AgN23_4221 (TPM = 199) and AgN23_1548 (TPM = 537), which also belongs to the top 10% in terms of CAZyme expression in AgN23. Taken together, these genomic and transcriptomic data support the hypothesis that AgN23 could degrade plant cell wall cellulose, hemicelluloses, and pectins to use them as carbon sources, a feature which could represent a major advantage in the rhizosphere niche. This hypothesis should be strengthened by investigating CAZyme gene expression patterns in the frame of plant co-cultivation with AgN23.

The AgN23 genome encodes the biosynthesis of antifungal and plant bioactive metabolites

Specialized metabolites are presumed to play key roles in the adaption of streptomycetes to their environment. We used antiSMASH 6.0 to detect and annotate AgN23 BGCs according to their similarity to reference clusters deposited in the MIBiG database (Blin et al. 2019; Epstein et al. 2018). Forty-seven BGCs were detected in the AgN23 genome, all of them being expressed during AgN23 cultivation in the Bennett liquid medium except for region 47 (Table 2, Supplementary Information S4). This result is somehow surprising because it has been shown in many *Streptomyces* strains that the expression of BGCs is tightly regulated, and only a few of them are expressed in standard cultivation media. This suggests particular regulatory mechanisms of the expression of AGN23 BGCs. Only region 20, putatively involved in pigment biosynthesis, stood out in terms of expression with 3,330 TPM, all the others ranging from a few to 200 TPM. Twenty BGCs of AgN23 showed at least 50% similar gene content with MIBiG reference BGCs (Table 2). These candidate BGCs are notably involved in the biosynthesis of volatile terpenes (regions 3 and 10), including geosmin (region 23), indoles (region 21), ribosomally synthesized and post-translationally modified peptide-like (RiPP-like, regions 5 and 28), ochronotic pigment (region 20), and a bicyclomycin-like antibacterial (region 45). Other BGCs are potentially involved in the production of siderophores, including coelichelin and desferrioxamin (regions

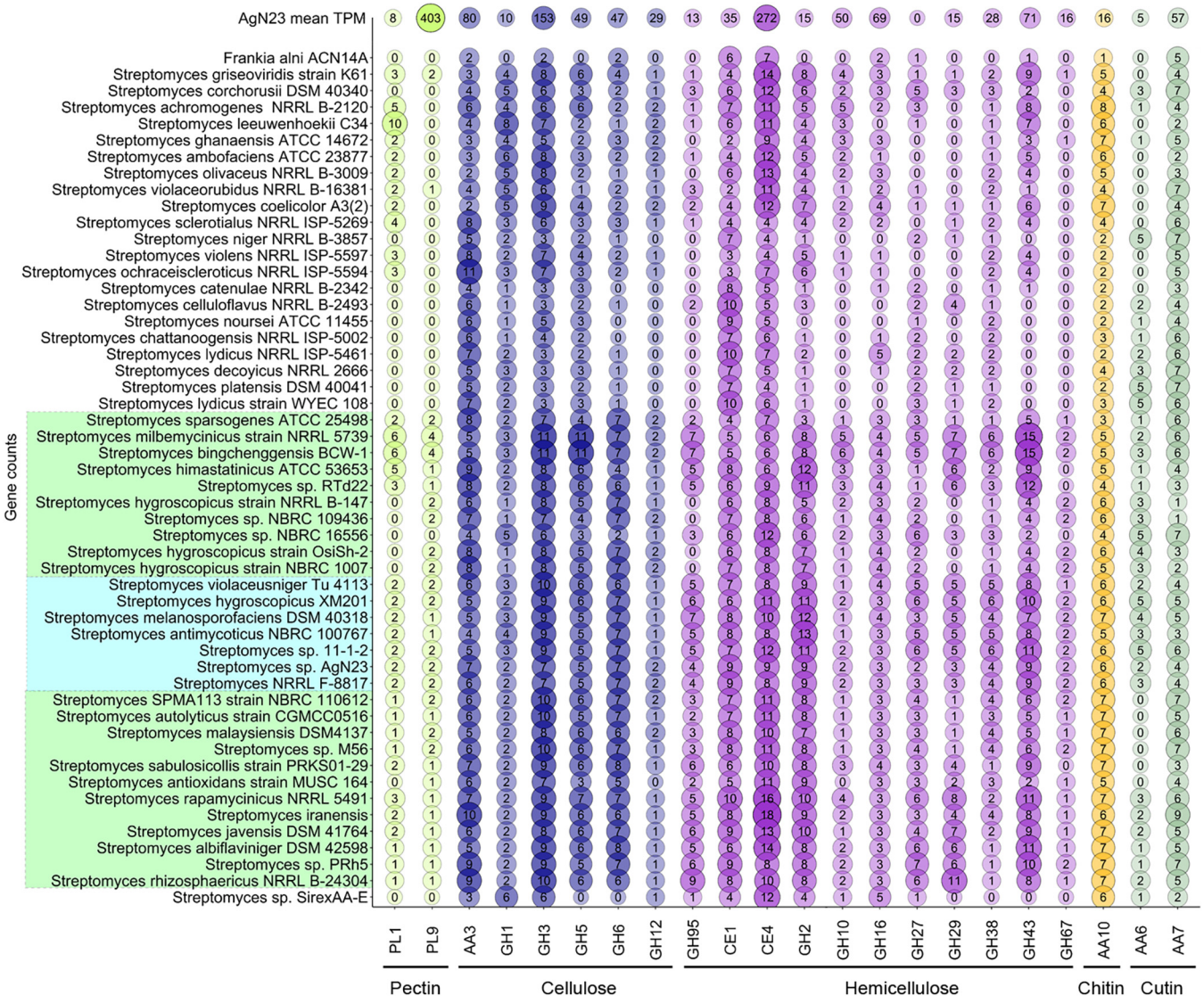

FIGURE 3

Plant cell wall-associated carbohydrate-active enzymes (CAZymes) repertoire is expanded in *Streptomyces violaceusniger* sp. as compared with *S. coelicolor* and *Streptomyces* sp. SirexAA-E. CAZymes families with at least two genes in AgN23 are ordered according to their target polymers and the type of enzymatic activity: glycoside hydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), and auxiliary activities (AAs). The transcriptions of AgN23 CAZymes are displayed as the mean transcript per million (TPM) reads for each family ($n = 3$).

24, 27, and 44), or osmotic and cold-stress protectants (ectoin, region 15). Additional BGCs are likely involved in the regulation of the bacterium life cycle, such as spore pigment (region 31), hopene (region 32), and butyrolactone (region 37). BGCs encoded in regions 2, 17, 41, and 42 are similar to the ones belonging to the biosynthesis pathways of the antifungal compounds rustmicin, also known as galbonolide A, mediomycin A, nigericin, and niphimycins C-E, respectively (Cai et al. 2007; Harvey et al. 2007; Hu et al. 2018; Kim et al. 2014). In addition, echoside (region 26), elaiophylin (region 40), and nigericin (region 41) are structural analogues of terfestatin, pteridic acid, and monensin, respectively, three compounds affecting plant immunity and development (Hayashi et al. 2008; Igarashi et al. 2002; Maintz et al. 2018; Wang et al. 2015; Yamazoe et al. 2005).

In addition to the BGC-mediated production of specialized metabolites, some *Streptomyces* spp. produce auxin, a phytohormone with a strong impact on root growth (Rashad et al. 2015). We undertook the annotation of auxin biosynthesis genes of AgN23 by retrieving protein sequences associated with KEGG ontologies, leading to the biosynthesis of indole-3 acetic acid (IAA), the most common plant auxin, from its precursor tryptophane through the indole-3-acetamide (IAM) pathway, the indole-3-acetaldoxime/indole-3-acetonitrile (IAN/IAOx) pathway, and the tryptamine (TAM) or indole-3-pyruvate (IPyA) pathway, both leading to indole-3-acetaldehyde (IAAld) before its conversion to IAA (Zhang et al. 2019). As a result, we identified a complete biosynthetic route for the IAM pathway (AgN23_8393 and AgN23_8392), the TAM-IAAld pathway (AgN23_3181, AgN23_0775, AgN23_0524, and AgN23_0525), and the IPyA

TABLE 2

antiSMASH annotation of AgN23 chromosomal regions coding for biosynthetic gene clusters (BGCs)[a]

Region	BGC type	Putative compound (MIBiG similarity score)	MIBiG accession	Species	Transcripts per million
1	NRPS	Atratumycin (7%)	BGC0001975	*S. atratus*	13
2	PKS-like, terpene	Rustmicin (20%)	BGC0000065	*S. galbus* KCCM 41354	15
3	Terpene	Tiancilactone (17%)	BGC0002019	*Streptomyces* sp. CB03234	7
4	NRPS	Dechlorocuracomycin (8%)	BGC0001569	*S. noursei* ATCC 11455	74
5	RiPP-like				13
6	T1PKS, hglE-KS	Leinamycin (2%)	BGC0001101	*S. atroolivaceus*	8
7	NRPS-like				6
8	T1PKS, NRPS	Meridamycin (52%)	BGC0001011	*Streptomyces* sp. NRRL 30748	14
9	Terpene	2-methylisoborneol (100%)	BGC0000658	*S. griseus* NBRC 13350	31
10	Terpene	Pristinol (100%)	BGC0001746	*S. pristinaespiralis* ATCC 25486	10
11	NRPS-like, T1PKS	Amipurimycin (90%)	BGC0001957	*S. novoguineensis*	26
12	T1PKS, T3PKS, NRPS-like, NRPS, betalactone	Totopotensamide A / totopotensamide B (43%)	BGC0001807	*S. pactum* SCSIO 02999	22
13	Lanthipeptide-class-iii	Lipopolysaccharide (43%)	BGC0000774	*Xanthomonas campestris*	18
14	T1PKS, siderophore	Apoptolidin (23%)	BGC0000021	*Nocardiopsis* sp. FU 40	24
15	Ectoine	Ectoine (100%)	BGC0000853	*S. anulatus*	43
16	Terpene				260
17	T1PKS	Mediomycin A (68%)	BGC0001932	*Kitasatospora mediocidica*	10
18	RRE-containing	Granaticin (10%)	BGC0000227	*S. violaceoruber*	25
19	Ladderane	Atratumycin (31%)	BGC0001975	*Streptomyces atratus*	14
20	NRPS	Ochronotic pigment (75%)	BGC0000918	*S. avermitilis*	3,330
21	Indole		BGC0001483	*Streptomyces* sp. RM-5-8	4
22	NRPS	RP-1776 (46%)	BGC0000429	*Streptomyces* sp. Acta 2897	87
23	Terpene	Geosmin (100%)	BGC0001181	*S. coelicolor* A3(2)	182
24	Siderophore	Desferrioxamin B (100%)	BGC0000941	*S. griseus* NBRC 13350	8
25	Terpene	Carotenoid (63%)	BGC0000633	*S. avermitilis*	155
26	NRPS-like	Echosides (100%)	BGC0000340	*Streptomyces* sp. LZ35	33
27	Siderophore				17
28	RiPP-like				13
29	Redox-cofactor				37
30	NRPS	Formicamycins A-M (18%)	BGC0001590	*Streptomyces* sp. KY5	32
31	T2PKS	Spore pigment (83%)	BGC0000271	*S. avermitilis*	141
32	Terpene	Hopene (76%)	BGC0000663	*S. coelicolor* A3(2)	47
33	Lanthipeptide-class-i	Steffimycin D (16%)	BGC0000273	*S. steffisburgensis*	29
34	Other	Mitomycin (16%)	BGC0000915	*S. lavendulae*	71
35	NRPS	Cadaside A / cadaside B (14%)	BGC0001968	Uncultured bacterium	67
36	T1PKS				80
37	Butyrolactone				6
38	Hserlactone	Daptomycin (3%)	BGC0000336	*S. filamentosus* NRRL 11379	76
39	Redox-cofactor				42
40	T1PKS	Elaiophylin (87%)	BGC0000053	Unknown	49
41	T1PKS	Nigericin (100%)	BGC0000114	*S. violaceusniger*	55
42	T1PKS	Niphimycins C-E (87%)	BGC0001700	*Streptomyces* sp. IMB7-145	13
43	NRPS-like	BD-12 (67%)	BGC0001379	*S. luteocolor*	5
44	NRPS	Coelichelin (90%)	BGC0000325	*S. coelicolor* A3(2)	12
45	CDPS	Bicyclomycin (100%)	BGC0001468	*S. cinnamoneus*	5
46	Betalactone				13
47	NRPS-like	Naphthyridinomycin (7%)	BGC0000394	*Streptomyces lusitanus*	0

[a] The functional category of each BGC was determined by antiSMASH. The BGC type and best hit in the MIBiG database as its percentage of similarity to the query are indicated, along with the bacterial strain from which the cluster was described. The expression level of each BGC was determined by determining the mean of the expression level in transcripts per million (TPM) of the core biosynthetic genes of each BGC from the RNA-seq data.

to IAAld pathway (AgN23_1600), as well as enzymes converting IAN to IAM (AgN23_1182 and AgN23_1183) (Fig. 4). All those enzymes were found to be expressed in our transcriptome analysis of the bacteria pure culture (Supplementary Information S7). Taken together, these data reveal that AgN23 likely produces specialized metabolites and auxins with a potential to regulate host plant biology along with its root microbiota.

A core set of BGCs conserved across *S. violaceusniger* species

The *S. violaceusniger* clade hosts several strains possessing BGCs involved in the synthesis of antifungal polyene macrolides such as nigericin, elaiophylin, and geldanamycin (Chen et al. 2013; Goodfellow et al. 2007; Hayakawa et al. 2004; Hayashi et al. 1992; Kang et al. 2010; Kumar and Goodfellow 2008; Nagpure et al. 2014; Riclea et al. 2014). However, strains sharing phylogenetic vicinity may differ in their specialized metabolism due to variations in their BGCs (Choudoir et al. 2018; Martinet et al. 2020). To assess the diversity of BGCs within the *S. violaceusniger* clade, we set up a comparative genomic approach based on the chromosome-scale assemblies used for our phylogenetic analysis (Supplementary Information S1). antiSMASH retrieved 873 BGC-containing regions in the genomes of AgN23 and the 18 other selected genomes. These sequences were introduced into BiG-SCAPE, which grouped them into 415 gene cluster families (GCFs) and built a sequence similarity network. To support our BiG-SCAPE analysis, the sequences of the 47 AgN23 BGCs were then compared with the 1,225,071 BGCs of the BiG-FAM database using the BiG-SLICE algorithm (Kautsar et al. 2021a, b).

We first observed that BiG-SCAPE distributed the AgN23 BGCs among four categories comprising *S. violaceusniger*-specific BGCs (ANI > 85%) (Fig. 5A), BGCs found in the outgroup (ANI < 85%) (Fig. 5B), and AgN23 BGCs not involved in any GCFs (Fig. 5C). In addition, BiG-SCAPE identified GCFs involving BGCs of other *S. violaceusniger* strains but not from AgN23 (Fig. 5D). To back up the identification of BGCs specific to *S. violaceusniger* (Fig. 5A) or unique to AgN23 (Fig. 5C), we inspected any potential hits of the corresponding AgN23 BGCs in the BiG-FAM database (Supplementary Information S8).

As a result, six BGCs found to be *S. violaceusniger* specific in BiG-SCAPE (Fig. 5A) actually had outgroup hits in BiG-FAM. Additionally, one BGC with no hits in BiG-SCAPE had a similar BGC in Big-FAM. The results of BiG-SCAPE and Big-FAM analyses are consolidated in Figure 6.

These analyses revealed that BGCs involved in the biosynthesis of geldanamycin, a largely conserved phytotoxic, antifungal, and antibacterial compound of *S. violaceusniger* strains, were detected in six strains but neither in AgN23 nor its closest neighbor *S. antimycoticus* (Fig. 5D) (Heisey and Putnam 1986). Additionally, BiG-SCAPE analysis clustered regions containing rustmicin-like BGCs from four *S. violaceusniger* strains, but nei-

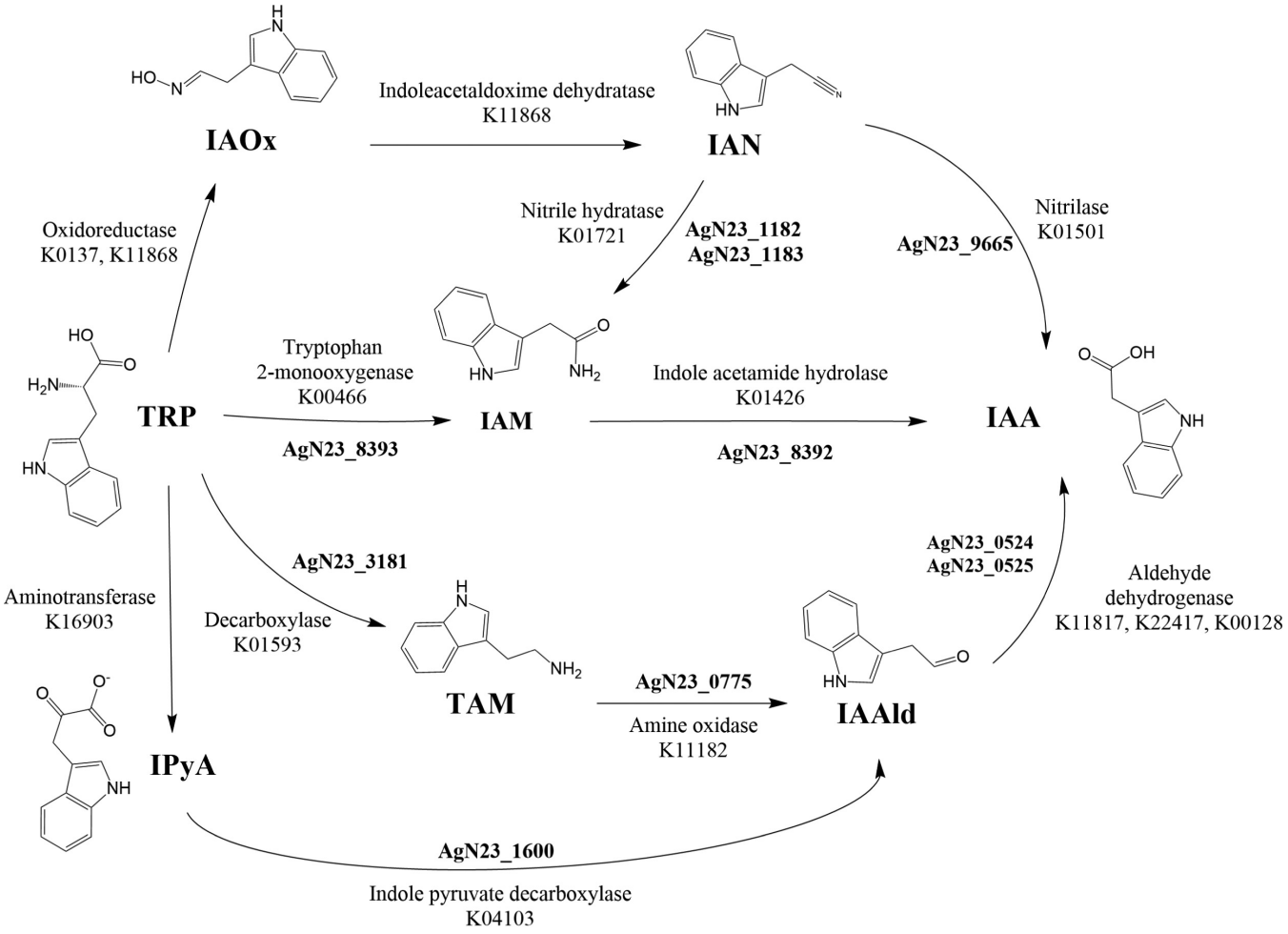

FIGURE 4

Overview of tryptophan-dependent indole-3 acetic acid (IAA) biosynthesis pathways in bacteria. Putative AgN23 genes encoding for enzymes of the KEGG pathways are indicated. Compound abbreviations: tryptophan (Trp), indole-3-acetaldoxime (IAOx), indole-3-acetonitrile (IAN), indole-3-acetamide (IAM), indole-3-pyruvate (IPyA), indole-3-acetaldehyde (IAAld), and tryptamine (TAM).

ther connected region 2 of AgN23 to it although that region contains a BGC annotated as rustmicin by antiSMASH (Table 2). Because the organization of the five BGCs is very similar to the subcluster GalA-E of *S. galbus* responsible for the biosynthesis of the macrolactone ring of the compound, we propose that the biosynthesis of a rustmicin-like compound is commonly found within the *S. violaceusniger* clade (Kim et al. 2014) (Fig. 7).

Importantly, both BiG-SCAPE and BiG-FAM confirmed the conservation of BGCs involved in antifungal activity (regions 2, 17, 41, and 42) and plant bioactive metabolites (regions 26, 40, and 41) that we identified in AgN23 and the *S. violaceusniger* clade. This suggests that this specialization could have arisen during the *Streptomyces* genus radiation and may result, or have participated, in the *S. violaceusniger* clade adaptation to its different ecological niche (Chater 2016; Doroghazi and Buckley 2010, 2014).

BGCs synteny among *S. violaceusniger* spp. and putative BGC horizontal transfers

Streptomyces chromosome organization consists of a central conserved genome, whereas terminal sequences contain more variable gene content described as an accessory genome (Kim et al. 2015; Lioy et al. 2021; Szafran et al. 2021). This accessory genome undergoes frequent rearrangement, amplification, and deletion events, as well as interspecies homologous recombination (Andam et al. 2016; Cheng 2016; Choudoir and Buckley 2018; Doroghazi and Buckley 2010; McDonald and Currie 2017;

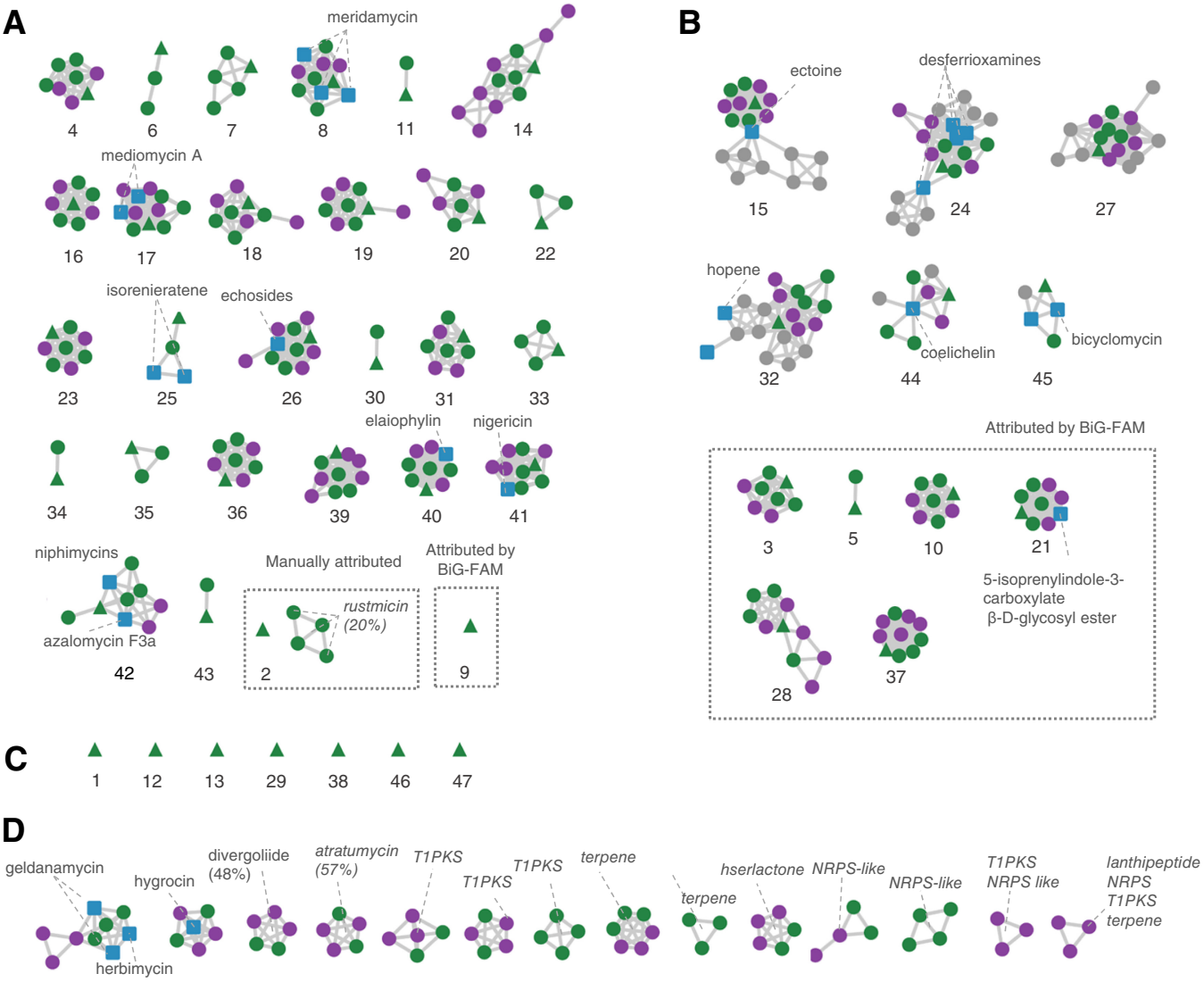

FIGURE 5

Twenty-eight biosynthetic gene cluster (BGC)-containing regions in the *Streptomyces* sp. AgN23 genome are specifically found within *S. violaceusniger* strains. **A,** The BiG-SCAPE similarity network analysis of the genomes of AgN23 and 18 other isolates corrected by BiGFAM reveals that 28 of AgN23 BGC-containing regions are similar to regions from the genomes of at least one other member of the *S. violaceusniger* clade. **B,** Twelve are broadly spread and **C,** seven do not match with any other tested sequence. **D,** AgN23 does not show similarity with any of the 14 BGC-containing regions that are found in at least three other members of the *S. violaceusniger* clade. Among them, the largely conserved geldanamycin, herbimycin, and hygrocin biosynthesis pathways are absent. Each node represents a BGC-containing region from one of the compared genomes, and they are clustered within regions displaying similar organization. They are colored as follows: BGC-containing regions from the genome of the same species as AgN23 are in green (average nucleotide identity [ANI] > 95%), other members of the *S. violaceusniger* clade are in purple (95%>ANI > 85%), and phylogenetically distant isolates are in gray. Green triangles represent AgN23 BGC sequences, and blue squares represent curated sequences from the MiBIG database. Dot frames highlight the results that differ from the BiG-SCAPE output based on manual annotation or BiG-FAM analysis.

Tidjani et al. 2019; Zhang et al. 2022). We decided to investigate commonalities and differences in the BGC organization across the *S. violaceusniger* clade and to pinpoint original features of our newly sequenced AgN23 strain.

We found that the core set of 26 BGCs found in the *S. violaceusniger* clade is syntenically positioned in the "left" and "right" arms of the chromosome or in the central region in AgN23, *Streptomyces* sp. 11-1-2, *S. antimycoticus* NBRC10767, *S. autolyticus* CGMCC0516, *S. malaysiensis* DSM4137, and *Streptomyces* sp. M56 (Fig. 6). By contrast, discrepancies in BGC positions exist between this latter group of genomes and *S. violaceusniger* Tü 4113, as well as *S. hygroscopicus* XM201, suggesting that these strains underwent drastic chromosome rearrangements but still retained BGCs connected to the synthesis of plant and fungal bioactive metabolites.

The regions 1, 46, and 47, which harbor original BGCs according to antiSMASH (<10% similarity with MIBiG) and the BiG-SCAPE analysis, are located on the extremities (left and

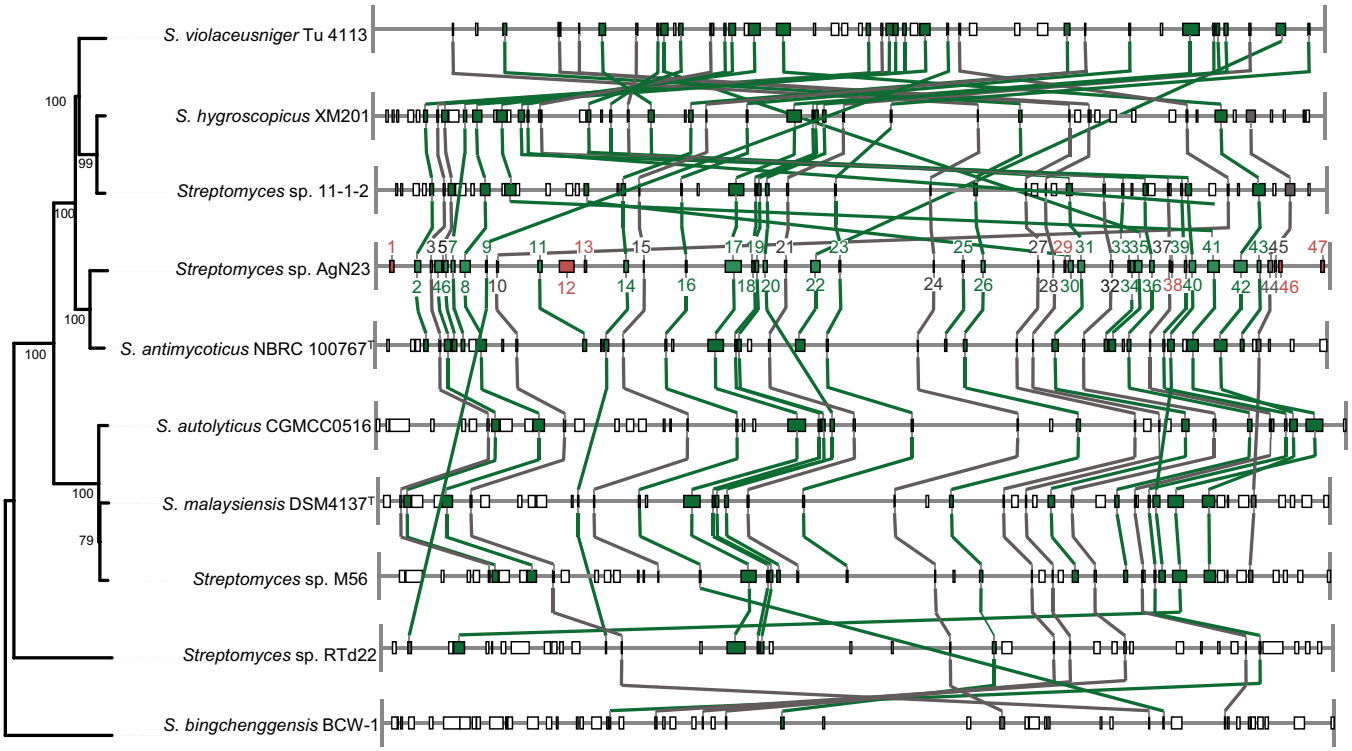

FIGURE 6
Synteny analysis of the biosynthetic gene cluster (BGC)-containing regions in genomes of the *Streptomyces violaceusniger* clade. The squares each represent BGC-containing regions that are linked together between different genomes by lines. They are colored in green if they have been shown to be specific to the *S. violaceusniger* clade, gray if they are widespread among *Streptomyces* spp., and orange if no similar regions have been detected in other genomes. The tree is inferred from GBDP distances calculated from fully assembled chromosomes of isolates from the *S. violaceusniger* clade. The numbers above branches are GBDP pseudo-bootstrap support values >60% from 100 replications.

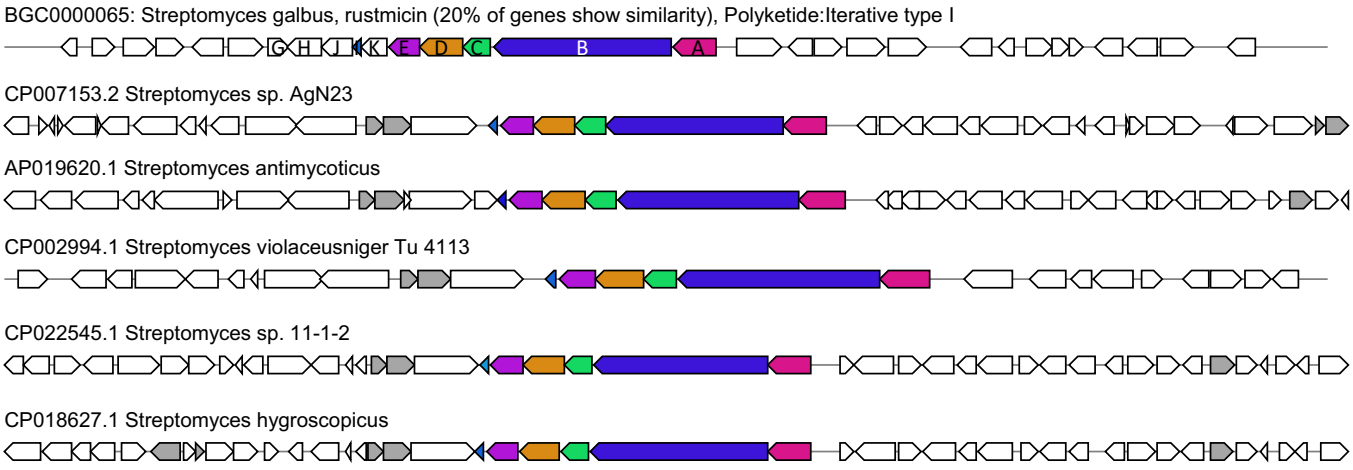

FIGURE 7
The rustmicin ABCDE biosynthetic gene cluster is conserved across the *Streptomyces violaceusniger* clade. Comparison of the rustmicin BGC sequence from MIBiG (BGC0000065) with genomic regions that are similar in AgN23 and four other *S. violaceusniger* genomes. Colored genes are putative homologs according to antiSMASH, with ClusterBLAST e-value < 1E-05, identity > 30%, and query coverage > 25%.

right arms) of the AgN23 chromosome and syntenic strains. Region 45 contains a BGC with 100% similarity to the biosynthetic pathway of the antimicrobial bicyclomycin, which is absent from other *S. violaceusniger* strains but is found outside of the clade (Table 2, Supplementary Information S8). Thus, region 45 might have been acquired by AgN23 through horizontal gene transfer. In conclusion, our synteny analysis is consistent with a preferential accumulation of unique or horizontally transferred BGCs on the chromosome extremities of AgN23. Conversely, the biosynthesis of specialized metabolites such as niphimycin or galbonolide was reported outside the *S. violaceusniger* clade, suggesting that some of these specific BGCs might have been horizontally transferred outside of the clade (Hu et al. 2018; Kim et al. 2014). Taken together, these data support the view that AgN23 accessory BGCs are concentrated at chromosome extremities.

CONCLUSIONS

Plant roots recruit abundant and diverse consortia of microorganisms while exploring the soil. The relevance of this root microbiota in plant nutrition and resistance to stresses is being unveiled through metabarcoding studies of plant microbiota. The *Streptomyces* genus constitutes one of the most prominent bacterial genera colonizing plant roots (Lundberg et al. 2012). Streptomycetes are filamentous gram-positive bacteria universally found around and within host plant tissues. These actinobacteria have been extensively investigated for their tremendous ability to produce diverse specialized metabolites (e.g., antibiotics). By contrast, their impact on host plant physiology is widely neglected. Additionally, few studies have addressed the species diversity of plant-associated *Streptomyces* and the genomic features enabling their colonization of the plant root environment. Therefore, genome sequencing of root-associated *Streptomyces* is an important step toward the description of their molecular interaction with the host plant. Here, we produced a complete chromosome sequence of *Streptomyces* sp. AgN23, a strain isolated from grapevine rhizosphere and shown to elicit plant defense responses (Vergnes et al. 2019). This reference sequence enabled us to perform a multilocus sequence typing approach and to position AgN23 in the *S. violaceusniger* clade from which several representatives have been isolated worldwide from the rhizosphere of unrelated plants. Comparative genomic studies suggest that *S. violaceusniger* spp. produce a large repertoire of CAZymes with expansion of plant cell wall-degrading enzyme families and a conserved specialized metabolism acting on plants and microorganisms. These genomic features may underly adaptation to the rhizospheric niche of some of the *S. violaceusniger* spp., such as AgN23. In addition, this phylogenetic lineage is considered to have a high potential in terms of specialized metabolism as its members have a large genome between 10.7 and 12.7 Mb and possess from 45 to 55 BGCs (Chung et al. 2021). The wealth of genomic data available in this clade allowed us to unveil common trends in the BGCs specifically found in *S. violaceusniger* isolates. The ability of these strains to produce antifungal compounds (nigericin, niphimycin, mediomycin, and rustmicin), as well as nigericin, which has phytotoxic or plant defense-elicitor activity, suggests that these metabolites play an important role in protection from pathogens. In addition, echosides and elaiophylin, from which structural analogues have been shown to interfere with auxin-like responses, could interfere with plant development (Igarashi et al. 2002; Wang et al. 2015). Thorough comparisons of the BGC content and chromosome organization of AgN23 with other *S. violaceusniger* spp. also enabled us to pinpoint BGC gains (i.e., bicyclomycin, an antibacterial product) and losses (i.e., geldanamycin, a phytotoxin), that could account for specific activities of the strain in the rhizosphere microbiota and colonization of plant tissues, respectively.

Availability of data and materials

The PacBio raw read sequences of the AgN23 genome are available at NCBI on the Sequence Read Archive "SRA" repository, SRR13990229. The MiSeq raw read sequences of the AgN23 genome are available at NCBI on the Sequence Read Archive "SRA" repository, SRR14028548. The genome assembly of AgN23 is available on the NCBI "Nucleotide" repository, NZ_CP007153.1. The RNA-seq raw reads of AgN23 are available at NCBI on the "BioProject" repository, PRJNA745930.

ACKNOWLEDGMENTS

We thank Olivier Bouchez for helpful discussions on the next-generation sequencing strategy throughout the project. We are grateful to Sylvie Lautru and Jean-Luc Pernodet (I2BC, Paris) for helpful discussions. We thank Thierry Grollier and Valérie Arnal for helpful comments on earlier versions of the manuscript.

LITERATURE CITED

Alanjary, M., Steinke, K., and Ziemert, N. 2019. AutoMLST: An automated web server for generating multi-locus species trees highlighting natural product potential. Nucleic Acids Res. 47:W276-W282.

Andam, C. P., Choudoir, M. J., Vinh Nguyen, A., Sol Park, H., and Buckley, D. H. 2016. Contributions of ancestral inter-species recombination to the genetic diversity of extant *Streptomyces* lineages. ISME J. 10:1731-1741.

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Meier-Kolthoff, J. P., Klenk, H. P., Clément, C., Ouhdouch, Y., and van Wezel, G. P. 2016. Taxonomy, physiology, and natural products of *Actinobacteria*. Microbiol. Mol. Biol. Rev. 80:1-43.

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H., and Weber, T. 2019. antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 47:W81-W87.

Book, A. J., Lewin, G. R., Mcdonald, B. R., Takasuka, T. E., Wendt-Pienkowski, E., Doering, D. T., Suh, S., Raffa, K. F., Fox, B. G., and Currie, C. R. 2016. Evolution of high cellulolytic activity in symbiotic streptomyces through selection of expanded gene content and coordinated gene expression. PLoS Biol. 14:e1002475.

Bown, L., and Bignell, D. R. D. 2017. Draft genome sequence of the plant pathogen *Streptomyces* sp. strain 11-1-2. Genome Announc. 5:e00968-17.

Bush, M. J., Tschowri, N., Schlimpert, S., Flärdh, K., and Buttner, M. J. 2015. c-di-GMP signalling and the regulation of developmental transitions in streptomycetes. Nat. Rev. Microbiol. 13:749-760.

Cai, P., Kong, F., Fink, P., Ruppen, M. E., Williamson, R. T., and Keiko, T. 2007. Polyene antibiotics from *Streptomyces mediocidicus*. J. Nat. Prod. 70:215-219.

Caicedo-Montoya, C., Manzo-Ruiz, M., and Ríos-Estepa, R. 2021. Pan-genome of the genus *Streptomyces* and prioritization of biosynthetic gene clusters with potential to produce antibiotic compounds. Front. Microbiol. 12:677558.

Cao, L., Gao, Y., Yu, J., Niu, S., Zeng, J., Yao, Q., Wang, X., Bu, Z., Xu, T., Liu, X., and Zhu, Y. 2021. *Streptomyces hygroscopicus* OsiSh-2-induced mitigation of Fe deficiency in rice plants. Plant Physiol. Biochem. 158:275-283.

Chagas, F. O., Ruzzini, A. C., Bacha, L. V., Samborskyy, M., Conti, R., Pessotti, R. C., de Oliveira, L. G., Clardy, J., and Pupo, M. T. 2016. Genome sequence of *Streptomyces* sp. strain RTd22, an endophyte of the Mexican sunflower. Genome Announc. 4:e00693-16.

Chater, K. F. 2016. Recent advances in understanding *Streptomyces*. F1000Research 5:2795.

Chen, X., Zhang, B., Zhang, W., Wu, X., Zhang, M., Chen, T., Liu, G., and Dyson, P. 2013. Genome sequence of *Streptomyces violaceusniger* strain SPC6, a halotolerant streptomycete that exhibits rapid growth and development. Genome Announc. 1:e00494-13.

Cheng, K., Rong, X., and Huang, Y. 2016. Widespread interspecies homologous recombination reveals reticulate evolution within the genus *Streptomyces*. Mol. Phylogenet. Evol. 102:246-254.

Choudoir, M. J., and Buckley, D. H. 2018. Phylogenetic conservatism of thermal traits explains dispersal limitation and genomic differentiation of *Streptomyces* sister-taxa. ISME J. 12:2176-2186.

Choudoir, M. J., Pepe-Ranney, C., and Buckley, D. H. 2018. Diversification of secondary metabolite biosynthetic gene clusters coincides with lineage divergence in *Streptomyces*. Antibiotics (Basel) 7:12.

Chung, Y.-H., Kim, H., Ji, C.-H., Je, H.-W., Lee, D., Shim, S. H., Joo, H.-S., and Kang, H.-S. 2021. Comparative genomics reveals a remarkable biosynthetic potential of the *Streptomyces* phylogenetic lineage associated with rugose-ornamented spores. mSystems 6:e00489-00421.

Díaz-Cruz, G. A., Liu, J., Tahlan, K., and Bignell, D. R. D. 2022. Nigericin and geldanamycin are phytotoxic specialized metabolites produced by the plant pathogen. Microbiol. Spectr. 10:e0231421.

Doroghazi, J. R., and Buckley, D. H. 2010. Widespread homologous recombination within and between *Streptomyces* species. ISME J. 4: 1136-1143.

Doroghazi, J. R., and Buckley, D. H. 2014. Intraspecies comparison of *Streptomyces pratensis* genomes reveals high levels of recombination and gene conservation between strains of disparate geographic origin. BMC Genomics 15:970.

Epstein, S. C., Charkoudian, L. K., and Medema, M. H. 2018. A standardized workflow for submitting data to the Minimum Information about a Biosynthetic Gene cluster (MIBiG) repository: Prospects for research-based educational experiences. Stand. Genomic Sci. 13:16.

Errakhi, R., Bouteau, F., Lebrihi, A., and Barakate, M. 2007. Evidences of biological control capacities of *Streptomyces* spp. against *Sclerotium rolfsii* responsible for damping-off disease in sugar beet (*Beta vulgaris* L.). World J. Microbiol. Biotechnol. 23:1503-1509.

Fitzpatrick, C. R., Copeland, J., Wang, P. W., Guttman, D. S., Kotanen, P. M., and Johnson, M. T. J. 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. Proc. Natl. Acad. Sci. U. S. A. 115:E1157-E1165.

Gao, Y., Ning, Q., Yang, Y., Liu, Y., Niu, S., Hu, X., Pan, H., Bu, Z., Chen, N., Guo, J., Yu, J., Cao, L., Qin, P., Xing, J., Liu, B., Liu, X., and Zhu, Y. 2021. Endophytic *Streptomyces hygroscopicus* OsiSh-2-mediated balancing between growth and disease resistance in host rice . mBio 12:e0156621.

Goodfellow, M., Kumar, Y., Labeda, D. P., and Sembiring, L. 2007. The *Streptomyces violaceusniger* clade: A home for streptomycetes with rugose ornamented spores. Antonie Van Leeuwenhoek 92:173-199.

Hamedi, J., and Mohammadipanah, F. 2015. Biotechnological application and taxonomical distribution of plant growth promoting actinobacteria. J. Ind. Microbiol. Biotechnol. 42:157-171.

Hamedi, J., Mohammadipanah, F., Klenk, H. P., Potter, G., Schumann, P., Sproer, C., von, J. M., and Kroppenstedt, R. M. 2010. *Streptomyces iranensis* sp. nov., isolated from soil. Int. J. Syst. Evol. Microbiol. 60: 1504-1509.

Harvey, B. M., Mironenko, T., Sun, Y., Hong, H., Deng, Z., Leadlay, P. F., Weissman, K. J., and Haydock, S. F. 2007. Insights into polyether biosynthesis from analysis of the nigericin biosynthetic gene cluster in *Streptomyces* sp. DSM4137. Chem. Biol. 14:703-714.

Hayakawa, M., Yoshida, Y., and Iimura, Y. 2004. Selective isolation of bioactive soil actinomycetes belonging to the *Streptomyces violaceusniger* phenotypic cluster. J. Appl. Microbiol. 96:973-981.

Hayashi, K., Hashimoto, M., Shigematsu, N., Nishikawa, M., Ezaki, M., Yamashita, M., Kiyoto, S., Okuhara, M., Kohsaka, M., and Imanaka, H. 1992. WS9326A, a novel tachykinin antagonist isolated from *Streptomyces violaceusniger* no. I. Taxonomy, fermentation, isolation, physicochemical properties and biological activities. J. Antibiot. (Tokyo) 45: 1055-1063.

Hayashi, K., Yamazoe, A., Ishibashi, Y., Kusaka, N., Oono, Y., and Nozaki, H. 2008. Active core structure of terfestatin A, a new specific inhibitor of auxin signaling. Bioorg. Med. Chem. 16:5331-5344.

Heisey, R. M., and Putnam, A. R. 1986. Herbicidal effects of geldanamycin and nigericin, antibiotics from *Streptomyces hygroscopicus*. J. Nat. Prod. 49:859-865.

Hu, Y., Wang, M., Wu, C., Tan, Y. i, Li, J., Hao, X., Duan, Y., Guan, Y., Shang, X., Wang, Y., Xiao, C., and Gan, M. 2018. Identification and proposed relative and absolute configurations of niphimycins C-E from the marine-derived *Streptomyces* sp. IMB7-145 by genomic analysis. J. Nat. Prod. 81:178-187.

Igarashi, Y., Iida, T., Yoshida, R., and Furumai, T. 2002. Pteridic acids A and B, novel plant growth promoters with auxin-like activity from *Streptomyces hygroscopicus* TP-A0451. J. Antibiot. (Tokyo) 55:764-767.

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9:5114.

Kang, M. J., Strap, J. L., and Crawford, D. L. 2010. Isolation and characterization of potent antifungal strains of the *Streptomyces violaceusniger* clade active against *Candida albicans*. J. Ind. Microbiol. Biotechnol. 37:35-41.

Kautsar, S. A., Blin, K., Shaw, S., Weber, T., and Medema, M. H. 2021a. BiG-FAM: The biosynthetic gene cluster families database. Nucleic Acids Res. 49:D490-D497.

Kautsar, S. A., van der Hooft, J. J. J., de Ridder, D., and Medema, M. H. 2021b. BiG-SLiCE: A highly scalable tool maps the diversity of 1.2 million biosynthetic gene clusters. Gigascience 10:giaa154.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. 2000. Practical *Streptomyces* Genetics. John Innes Foundation, Norwich, U.K.

Kim, H. J., Karki, S., Kwon, S. Y., Park, S. H., Nahm, B. H., Kim, Y. K., and Kwon, H. J. 2014. A single module type I polyketide synthase directs de novo macrolactone biogenesis during galbonolide biosynthesis in *Streptomyces galbus*. J. Biol. Chem. 289:34557-34568.

Kim, J. N., Kim, Y., Jeong, Y., Roe, J. H., Kim, B. G., and Cho, B. K. 2015. Comparative genomics reveals the core and accessory genomes of *Streptomyces* species. J. Microbiol. Biotechnol. 25:1599-1605.

Klassen, J. L., Lee, S. R., Poulsen, M., Beemelmanns, C., and Kim, K. H. 2019. Efomycins K and L from a termite-associated *Streptomyces* sp. M56 and their putative biosynthetic origin. Front. Microbiol. 10:1739.

Komaki, H., Ichikawa, N., Oguchi, A., Hamada, M., Harunari, E., Kodani, S., Fujita, N., and Igarashi, Y. 2016. Draft genome sequence of *Streptomyces* sp. TP-A0867, an alchivemycin producer. Stand. Genomic Sci. 11:85.

Komaki, H., and Tamura, T. 2020. Reclassification of *Streptomyces castelarensis* and *Streptomyces sporoclivatus* as later heterotypic synonyms of *Streptomyces antimycoticus*. Int. J. Syst. Evol. Microbiol. 70:1099-1105.

Kumar, Y., Aiemsum-Ang, P., Ward, A. C., and Goodfellow, M. 2007. Diversity and geographical distribution of members of the *Streptomyces violaceusniger* 16S rRNA gene clade detected by clade-specific PCR primers. FEMS Microbiol. Ecol. 62:54-63.

Kumar, Y., and Goodfellow, M. 2008. Five new members of the *Streptomyces violaceusniger* 16S rRNA gene clade: *Streptomyces castelarensis* sp. nov., comb. nov., *Streptomyces himastatinicus* sp. nov., *Streptomyces mordarskii* sp. nov., *Streptomyces rapamycinicus* sp. nov. and *Streptomyces ruanii* sp. nov. Int. J. Syst. Evol. Microbiol. 58:1369-1378.

Kusuma, A. B., Nouioui, I., and Goodfellow, M. 2021. Genome-based classification of the *Streptomyces violaceusniger* clade and description of *Streptomyces sabulosicollis* sp. nov. from an Indonesian sand dune. Antonie Van Leeuwenhoek 114:859-873.

Labeda, D. P., Goodfellow, M., Brown, R., Ward, A. C., Lanoot, B., Vanncanneyt, M., Swings, J., Kim, S.-B., Liu, Z., Chun, J., Tamura, T., Oguchi, A., Kikuchi, T., Kikuchi, H., Nishii, T., Tsuji, K., Yamaguchi, Y., Tase, A., Takahashi, M., Sakane, T., Suzuki, K. I., and Hatano, K. 2012. Phylogenetic study of the species within the family *Streptomycetaceae*. Antonie Van Leeuwenhoek 101:73-104.

Lee, N., Kim, W., Hwang, S., Lee, Y., Cho, S., Palsson, B., and Cho, B. K. 2020. Thirty complete *Streptomyces* genome sequences for mining novel secondary metabolite biosynthetic gene clusters. Sci. Data 7:55.

Lefort, V., Desper, R., and Gascuel, O. 2015. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. Mol. Biol. Evol. 32:2798-2800.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.

Lioy, V. S., Lorenzi, J.-N., Najah, S., Poinsignon, T., Leh, H., Saulnier, C., Aigle, B., Lautru, S., Thibessard, A., Lespinet, O., Leblond, P., Jaszczyszyn, Y., Gorrichon, K., Varoquaux, N., Junier, I., Boccard, F., Pernodet, J.-L., and Bury-Moné, S. 2021. Dynamics of the compartmentalized *Streptomyces* chromosome during metabolic differentiation. Nat. Commun. 12:5221.

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., Rio, T. G. D., Edgar, R. C., Eickhorst, T., Ley, R. E., Hugenholtz, P., Tringe, S. G., and Dangl, J. L. 2012. Defining the core *Arabidopsis thaliana* root microbiome. Nature 488:86-90.

Maintz, J., Suliman, M., Joglekar, S., Halder, V., Kombrink, E., Bautor, J., Bartsch, M., Parker, J. E., and Zeier, J. 2018. Chemical activation of EDS1/PAD4 signaling leading to pathogen resistance in *Arabidopsis*. Plant Cell Physiol. 59:1592-1607.

Martinet, L., Naômé, A., Baiwir, D., De Pauw, E., Mazzucchelli, G., and Rigali, S. 2020. On the risks of phylogeny-based strain prioritization for drug discovery. Biomolecules 10:1027.

McDonald, B. R., and Currie, C. R. 2017. Lateral gene transfer dynamics in the ancient bacterial genus *Streptomyces*. MBio 8:1027.

Medema, M. H. 2021. The year 2020 in natural product bioinformatics: An overview of the latest tools and databases. Nat. Prod. Rep. 38:301-306.

Meier-Kolthoff, J. P., and Göker, M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat. Commun. 10:2182.

Moumbock, A. F. A., Gao, M., Qaseem, A., Li, J., Kirchner, P. A., Ndingkokhar, B., Bekono, B. D., Simoben, C. V., Babiaka, S. B., Malange,

Y. I., Sauter, F., Zierep, P., Ntie-Kang, F., and Günther, S. 2021. StreptomeDB 3.0: An updated compendium of streptomycetes natural products. Nucleic Acids Res. 49:D600-D604.

Nagpure, A., Choudhary, B., and Gupta, R. K. 2014. Mycolytic enzymes produced by *Streptomyces violaceusniger* and their role in antagonism towards wood-rotting fungi. J. Basic Microbiol. 54:397-407.

Navarro-Muñoz, J. C., Selem-Mojica, N., Mullowney, M. W., Kautsar, S. A., Tryon, J. H., Parkinson, E. I., De Los Santos, E. L. C., Yeong, M., Cruz-Morales, P., Abubucker, S., Roeters, A., Lokhorst, W., Fernandez-Guerra, A., Cappelini, L. T. D., Goering, A. W., Thomson, R. J., Metcalf, W. W., Kelleher, N. L., Barona-Gomez, F., and Medema, M. H. 2020. A computational framework to explore large-scale biosynthetic diversity. Nat. Chem. Biol. 16:60-68.

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25:1043-1055.

Quinlan, A. R., and Hall, I. M. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics 26:841-842.

Rashad, F. M., Fathy, H. M., El-Zayat, A. S., and Elghonaimy, A. M. 2015. Isolation and characterization of multifunctional *Streptomyces* species with antimicrobial, nematicidal and phytohormone activities from marine environments in Egypt. Microbiol. Res. 175:34-47.

Rey, T., and Dumas, B. 2017. Plenty is no plague: *Streptomyces* symbiosis with crops. Trends Plant Sci. 22:30-37.

Riclea, R., Citron, C. A., Rinkel, J., and Dickschat, J. S. 2014. Identification of isoafricanol and its terpene cyclase in *Streptomyces violaceusniger* using CLSA-NMR. Chem. Commun. (Camb.) 50:4228-4230.

Ser, H.-L., Tan, L. T.-H., Palanisamy, U. D., Abd Malek, S. N., Yin, W.-F., Chan, K.-G., Goh, B.-H., and Lee, L.-H. 2016. *Streptomyces antioxidans* sp. nov., a novel mangrove soil actinobacterium with antioxidative and neuroprotective potentials. Front. Microbiol. 7:889.

Shirling, E. B., and Gottlieb, D. 1966. Methods for characterization of *Streptomyces* species. Int. J. Syst. Bacteriol. 16:313-340.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31: 3210-3212.

Szafran, M. J., Małecki, T., Strzałka, A., Pawlikiewicz, K., Duława, J., Zarek, A., Kois-Ostrowska, A., Findlay, K. C., Le, T. B. K., and Jakimowicz, D. 2021. Spatial rearrangement of the *Streptomyces venezuelae* linear chromosome during sporogenic development. Nat. Commun. 12:5222.

Tidjani, A. R., Lorenzi, J. N., Toussaint, M., van Dijk, E., Naquin, D., Lespinet, O., Bontemps, C., and Leblond, P. 2019. Massive gene flux drives genome diversity between sympatric *Streptomyces* conspecifics. MBio 10:e01533-19.

Tracanna, V., de Jong, A., Medema, M. H., and Kuipers, O. P. 2017. Mining prokaryotes for antimicrobial compounds: From diversity to function. FEMS Microbiol. Rev. 41:417-429.

Vallenet, D., Calteau, A., Dubois, M., Amours, P., Bazin, A., Beuvin, M., Burlot, L., Bussell, X., Fouteau, S., Gautreau, G., *et al.* 2020. MicroScope: An integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. Nucleic Acids Res. 48:D579-D589.

Vergnes, S., Gayrard, D., Veyssiere, M., Toulotte, J., Martinez, Y., Dumont, V., Bouchez, O., Rey, T., and Dumas, B. 2019. Phyllosphere colonisation by a soil *Streptomyces* sp. promotes plant defense responses against fungal infection. Mol. Plant-Microbe Interact. 33: 223-234.

Viaene, T., Langendries, S., Beirinckx, S., Maes, M., and Goormachtig, S. 2016. *Streptomyces* as a plant's best friend?. FEMS Microbiol. Ecol. 92: fiw119.

Vurukonda, S., Giovanardi, D., and Stefani, E. 2018. Plant growth promoting and biocontrol activity of *Streptomyces* spp. as endophytes. Int. J. Mol. Sci. 19:952.

Wang, X., Reynolds, A. R., Elshahawi, S. I., Shaaban, K. A., Ponomareva, L. V., Saunders, M. A., Elgumati, I. S., Zhang, Y., Copley, G. C., Hower, J. C., Sunkara, M., Morris, A. J., Kharel, M. K., Van Lanen, S. G., Prendergast, M. A., and Thorson, J. S. 2015. Terfestatins B and C, new p-terphenyl glycosides produced by *Streptomyces* sp. RM-5-8. Org. Lett. 17:952fiw1192796-2799.

Xu, T., Cao, L., Zeng, J., Franco, C. M. M., Yang, Y., Hu, X., Liu, Y., Wang, X., Gao, Y., Bu, Z., Shi, L., Zhou, G., Zhou, Q., Liu, X., and Zhu, Y. 2019. The antifungal action mode of the rice endophyte *Streptomyces hygroscopicus* OsiSh-2 as a potential biocontrol agent against the rice blast pathogen. Pestic. Biochem. Physiol. 160:58-69.

Xu, T., Li, Y., Zeng, X., Yang, X., Yang, Y., Yuan, S., Hu, X., Zeng, J., Wang, Z., Liu, Q., Liu, Y., Liao, H., Tong, C., Liu, X., and Zhu, Y. 2017. Isolation and evaluation of endophytic *Streptomyces endus* OsiSh-2 with potential application for biocontrol of rice blast disease. J. Sci. Food Agric. 97:1149-1157.

Yamazoe, A., Hayashi, K., Kepinski, S., Leyser, O., and Nozaki, H. 2005. Characterization of terfestatin A, a new specific inhibitor for auxin signaling. Plant Physiol. 139:779-789.

Yang, H., Zhang, Z., Yan, R., Wang, Y., and Zhu, D. 2014. Draft genome sequence of *Streptomyces* sp. strain PRh5, a novel endophytic actinomycete isolated from dongxiang wild rice root. Genome Announc. 2: e00012-14.

Yin, M., Jiang, M., Ren, Z., Dong, Y., and Lu, T. 2017. The complete genome sequence of *Streptomyces autolyticus* CGMCC 0516, the producer of geldanamycin, autolytimycin, reblastatin and elaiophylin. J. Biotechnol. 252:27-31.

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. 2012. dbCAN: A web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 40:W445-W451.

Zeng, J., Xu, T., Cao, L., Tong, C., Zhang, X., Luo, D., Han, S., Pang, P., Fu, W., Yan, J., Liu, X., and Zhu, Y. 2018. The role of iron competition in the antagonistic action of the rice endophyte *Streptomyces sporocinereus* OsiSh-2 against the pathogen *Magnaporthe oryzae*. Microb. Ecol. 76: 1021-1029.

Zhang, P., Jin, T., Kumar Sahu, S., Xu, J., Shi, Q., Liu, H., and Wang, Y. 2019. The distribution of tryptophan-dependent indole-3-acetic acid synthesis pathways in bacteria unraveled by large-scale genomic analysis. Molecules 24:1411.

Zhang, Z., Shitut, S., Claushuis, B., Claessen, D., and Rozen, D. E. 2022. Mutational meltdown of putative microbial altruists in *Streptomyces coelicolor* colonies. Nat. Commun. 13:14112266.