Streptomyces endocoffea sp. nov. , an endophytic actinomycete isolated from *Coffea arabica* (L.)

Achararak Nammali

King Mongkut's Institute of Technology

Chakapong Intaraudom

National Center for Genetic Engineering and Biotechnology (BIOTEC)

Pattama Pittayakhajonwut

National Center for Genetic Engineering and Biotechnology (BIOTEC)

Chanwit Suriyachadkun

Thailand Bioresource Research Center (TBRC)

Sarin Tadtong

National Science and Technology Development Agency (NSTDA)

Somboon Tanasupawat

Chulalongkorn University

Chitti Thawai (✉ Chitti.th@kmitl.ac.th)

King Monkut's Institute of Tecnology Ladkrabang

Research Article

Abstract

An aerobic, non-motile, Gram-stain positive actinomycete, designated strain CA3R110^T, was isolated from the surface-sterilised root of *Coffea arabica* L. collected from Lampang Province, Thailand. 16S rRNA gene sequence analysis indicated that strain CA3R110^T is a member of the genus *Streptomyces* and showed the closest similarities to *Streptomyces rapamycinicus* NRRL B-5491^T (99.1%), followed by *S. iranensis* HM 35^T (99.0%). Strain CA3R110^T contained *LL*-diaminopimelic acid in cell peptidoglycan, MK-9(H$_6$), and MK-9(H$_8$) as major menaquinone, *iso*-C$_{16:0}$, *iso*-C$_{15:0}$, C$_{16:0}$ as major fatty acids. Diphosphatidylglycerol, phosphatidylethanolamine, hydroxy-phosphatidylethanolamine, phosphatidylglycerol, phosphatidylinositol, phosphatidylinositolmannoside were detected in the cell. The chemotaxonomic characteristics possessed the typical properties of the genus *Streptomyces*. A low digital DNA−DNA hybridization (< 55.6%) and average nucleotide identity-blast (ANIb) (< 92.2%) values revealed that strain CA3R110^T could be distinguished from any known *Streptomyces* species. With the differences in phenotypic and genotypic data, strain CA3R110^T represents a novel species of genus *Streptomyces*, for which the name *Streptomyces endocoffea* sp. nov. is proposed. The type strain is CA3R110^T (= TBRC 11245^T = NBRC 114296^T).

Introduction

An attempt to investigate the novel actinomycete in an unexplored environment is one strategy to discover a new bioactive compound from the microbial resource. It is known that plants are major sources for recovering secondary metabolites. They always accumulate the secondary metabolites in their tissue. Therefore the environment inside the plant tissue may differ significantly from other habitats. Hence the microorganisms residing in the plant tissue are expected to be different from those of soil microorganisms. Many promising endophytic microbes have been reported as candidate biocontrol agents against pathogens. Endophytic actinomycetes, especially *Streptomyces* spp., have been recognized as the promising producer for biocontrol agents to soilborne plant diseases (Solecka et al. 2012). Many reports, for example, 3-acetonylidene-7-prenylindolin-2-one from mycelial cake extracted with EtOH of *Streptomyces* sp. NEAU-D50 showed cytotoxic activities against human adenocarcinoma cell line A549 and antifungal activity against the tested phytopathogenic fungi *Colletotrichum orbiculare, Phytophthora capsici, Corynespora cassiicola* and *Fusarium oxysporum* (Zhang et al. 2014). 6-Prenylindole from the fermented broth of *Streptomyces* sp. TP-A0595 exhibited against plant fungal pathogens, *Alternaria brassicicola* and *Fusarium oxysporum* (Sasaki et al. 2002; Singh and Dubey 2018). Azalomycin B from the fermentation broth of *Streptomyces* sp. HAAG3-15 showed significant antifungal activity against *Fusarium oxysporum* f. sp. *cucumerinum* (Cao et al. 2020). To date, many new species of endophytic *Streptomyces* have continuously been reported. For example, *Streptomyces typhae* sp. nov. (Peng et al. 2021), *Streptomyces albidus* sp. nov. (Kaewkla and Franco 2021), *Streptomyces dioscori* sp. nov. (Wang et al. 2018) were reported as novel endophytic *Streptomyces* species which were isolated from the root of *Typha angustifolia* L., *Callitris preissii*, and the Bulbil of *Dioscorea bulbifera* L., respectively.

In the course of our investigation of endophytic actinomycetes from the coffee plant, strain CA3R110[T] was isolated from the surface-sterilised root of *Coffea arabica* L. collected from Lampang Province, Thailand. The polyphasic taxonomic details of strain CA3R110[T] were reported in this study.

Materials And Methods

Actinobacterium isolation and cultural conditions

The strain CA3R110[T] was isolated from the surface-sterilized root of *Coffea arabica* L., collected from Ratchaburi province, Thailand. The root sample was gently washed in tap water to remove the dirt and cut into small pieces (1 cm each). After cutting, the sample was washed in sterile 0.1 % tween 20 for 30 sec and then washed twice in sterile distilled water. The sample was surface-sterilized with 95 % ethanol for 10 min and subsequently with 1 % sodium hypochlorite for 6 min, followed by washing in sterile distilled water for three times. Finally, the surface-sterilized sample was ground with sterile mortar and pestle with 0.5 ml sterile distilled water, and the suspension (200 μl) was spread on starch casein agar (SCA) [10 g soluble starch, 2 g KNO_3, 0.3 g casein, 2 g K_2HPO_4, 0.05 g $MgSO_4.7H_2O$, 2 g NaCl, 0.02 g $CaCO_3$, 0.01 g $FeSO_4.7H_2O$, 15 g agar, 1000 ml distilled water, pH: 7.0] supplemented with 25 mg l^{-1} nalidixic acid, and 50 mg l^{-1} nystatin and incubated at 30 °C for 30 days. The final washing water was spread on the same agar medium as the control plates. The pure culture was preserved by freezing at -80 °C and freeze-drying.

Morphology, cultural, biochemical, and physiological tests

Spore chain morphology and spore-surface ornamentation of strain CA3R110[T] were observed after growing on a half strength-ISP2 medium at 30 °C for 30 days by a scanning electron microscope (model JSM-6610 LV; JEOL). The cultural characteristic was performed by cultivation the strain on various International *Streptomyces* Project (ISP) media (ISP2-7) (Shirling and Gottlieb 1966), czapek's sucrose agar (Waksman 1961), glucose-asparagine agar (ISP5 with 1% glucose replacing glycerol), and nutrient agar (Difco) at 30 °C for 14 days. The colours of the substrate and aerial mycelium and diffusible pigments were evaluated by comparison with the ISCC-NBS color charts (Kelly 1964). Growth at different temperatures (10-50 °C), and 0-10%, w/v NaCl (at increment of 1%) were evaluated on ISP2 agar after incubation for 14 days. The effect of pH on growth (4.0-11.0 at an increment of 0.5 pH unit) was determined by cultivation at 30 °C in ISP2 broth for 14 days. Catalase activity was observed by bubble production after applying 3% (v/v) hydrogen peroxide solution. Test for oxidase activity was carried out by oxidation of 1% *N, N, N',N'*-tetramethyl-*p*-phenylenediamine dihydrochloride. Carbon and nitrogen utilisations were examined according to the standard method of Shirling and Gottlieb 1966 and Gordon et al. (1974), respectively. Nitrate reduction, the acid production, the decomposition of insoluble compounds, e.g., xanthine, adenine, hypoxanthine and tyrosine, gelatin liquefaction, hydrolysis of casein (1 %, w/v), and starch (1%, w/v) were evaluated by using the methods of Arai (1975), Williams and Cross (1971) and Gordon et al. (1974). The enzyme activities were determined using the API ZYM system (bioMérieux).

Chemotaxonomy.

Biomass used for chemotaxonomic analyses was collected from cell culture grown in ISP 2 broth on a rotary shaker (200 r.p.m.) at 30˚C for five days.

The diaminopimelic acid isomers and whole-cell sugars were evaluated based on the method of Hasegawa et al. (1983), and Komagata and Suzuki (1987), respectively. The polar lipids in the cell membrane were extracted and analyzed by the methods proposed by Minnikin et al. (1984) and analyzed by the two-dimension TLC technique of Collins and Jones (1980). Menaquinones were extracted as described by Collins et al. (1977) and were analyzed by high-performance liquid chromatography equipped with a Cosmosil 5C18 column (4.6 x 150 mm, Nacalai Tesque) (Tamaoka et al. 1983). Fatty acids were extracted, methylated and analyzed using the Sherlock Microbial Identification (MIDI) system and the ACTIN version 6 database (Sasser 1990), and Kämpfer and Kroppenstedt (1996).

DNA extraction, 16S rRNA gene analysis

Genomic DNA was extracted using the protocol of Tamaoka (1994). The 16S rRNA gene amplification and sequencing were conducted using the method suggested by Thawai (2015). The 16S rRNA gene similarity was calculated using the EzTaxon-e server (www.ezbiocloud.net) (Yoon et al. 2017). To reconstruct the phylogenetic tree, the 16S rRNA gene sequence was multiple-aligned with selected sequences available from the GenBank/EMBL/DDBJ databases by using CLUSTAL W multiple alignment mode within BioEdit program version 7.1.3.0 (Hall 1999). Phylogenetic trees were reconstructed by using neighbor-joining (NJ) (Saitou and Nei 1987), and maximum-likelihood (ML) (Felsenstein 1981) algorithms in MEGA version X program (Kumar et al. 2018). The evolutionary distances were calculated by using Kimura's 2-parameter model (Kimura 1980). The Tamura–Nei model (Tamura and Nei 1993) was applied to the maximum-likelihood analysis using the Subtree-Pruning-Regrafting-Extensive (SPR level 5) program. The confidence values of nodes were evaluated by using the bootstrap resampling method with 1,000 replicates (Felsenstein 1985).

Genome sequencing and analysis

Genomic DNA for whole-genome sequencing of strain CA3R110[T] was extracted from 3-day old cultures grown in ISP2 broth at 30 ᵒC and purified following the GeneJET Genomic DNA protocol purification Kit (Thermo Scientific). The genomic DNA of strain CA3R110[T] was sequenced using an Illumina HiSeq 4000 platform (Chulalongkorn University, Thailand). Genome assembly were performed using SPAdes (Bankevich et al. 2012). The online server, Rapid Annotations using Subsystems Technology (RAST) (http://rast.nmpdr.org/), was used for genome annotation (Aziz et al. 2008; Overbeek et al. 2014). The genome was applied to the JSpecies Web Server for calculating the average nucleotide identity (ANI) values (Richter et al. 2016; Richter and Rosselló-Móra 2009). The digital DNA-DNA hybridization (dDDH) values between the genome of strain CA3R110[T], and the most closely related species were calculated using the genome-to-genome distance calculator (GGDC 2.1; blast+method) in which formula 2 (identities/HSP length) was applied to the incomplete draft genome (Meier-Kolthoff et al. 2013). The

phylogenomic tree was reconstructed using the Type (strain) Genome Server (TYGS) (Meier-Kolthoff and Göker 2019). Anti-SMASH program was used to evaluate the secondary metabolite biosynthesis gene clusters in the bacterial genome (Blin et al. 2019).

Results And Discussion

Morphology and cultural characteristic of strain CA3R110^T

Strain CA3R110^T formed well-developed and nonfragmented branched substrate mycelia and grew well on ISP2, ISP4, ISP5, ISP6, and czapek's sucrose agar. Moderate growth was observed on ISP3, ISP7, nutrient agar, and glucose-asparagine agar. The strain could produce pale greenish-yellow to dark brownish-gray substrate mycelia on the above media tested. Greyish-white aerial mycelium was easily formed on all media tested after 21 days of cultivation at 30 ^{o}C. The brilliant greenish-yellow diffusible pigment was observed on ISP5 and ISP7 (Table S1). The spiral chain of spore with the rugose surface was detected (Fig. S1). Morphological detail revealed that it formed the same morphological characteristics as other members of the genus *Streptomyces* (Waksman and Henrici 1943; Pridham et al. 1958).

Chemotaxonomic characteristics of strain CA3R110^T

The strain showed *LL*-diaminopimelic acid (*LL*-DAP) as the diagnostic diamino acid in the cell wall peptidoglycan. This chemotaxonomic property was generally found in all members of the genus *Streptomyces* (Wellington et al. 1992; Xu et al. 2012). MK-9(H$_6$) (73.2%) and MK-9(H$_8$) (23.7%) were detected as the major menaquinone (MK) in the cell, while MK-9(H$_4$) (2.9%) and MK-9(H$_2$) (0.2%) were also observed. The reducing sugars in cell-hydrolysates were galactose, glucose, ribose. The polar lipid profile consisted of diphosphatidylglycerol (DPG), phosphatidylethanolamine (PE), hydroxy-phosphatidylethanolamine (OH-PE), phosphatidylglycerol (PG), phosphatidylinositol (PI), phosphatidylinositolmannoside (PIM) and six unidentified phospholipids (PLs) (Fig. S2). The predominant fatty acids (>10 %), *iso*-C$_{16:0}$, *iso*-C$_{15:0}$, C$_{16:0}$, were detected, while *iso*-C$_{14:0}$, C$_{14:0}$, *anteiso*-C$_{15:0}$, *iso*-C$_{17:0}$, *anteiso*-C$_{17:0}$, cyclo-C$_{17:0}$ were found to be minor fatty acid components in the cell (Table S2). The pattern of major fatty acids in the cell of strain CA3R110^T was similar to those predominant fatty acids reported in *S. rapamycinicus* DSM 41530^T and *S. iranensis* DSM 41954^T with different proportions. All these chemotaxonomic data confirmed that strain CA3R110^T should be assigned to the genus *Streptomyces*.

16S rRNA gene and genome analyses

According to the result from the EzBioCloud (http://www.ezbiocloud.net/) server (Yoon et al. 2017), strain CA3R110^T related to the genus *Streptomyces*. The strain is closely related to *S. rapamycinicus* NRRL B-5491^T (99.1%), followed by *S. iranensis* HM 35^T (99.0%). Strain CA3R110^T forms a distinct monophyletic line within the genus *Streptomyces*, supported by the two treeing algorithms, maximum-likelihood (ML),

and neighbor-joining (NJ), but low bootstrap values (<50%) were observed on both phylogenetic trees (Fig. 1 and Fig. S3). To confirm the taxonomic position of strain CA3R110[T] at the species level, genome-based taxonomic characterization of strain CA3R110[T] was performed. Genome sequencing of strain CA3R110[T] resulted in 269 contigs, an N50 of 328 kb, genome coverage of 138x, and comprised 13128250 bp. The genomic DNA G+C content of this strain was 70.9 mol%, which was in the range of the genus *Streptomyces* (Pridham et al. 1958). The draft genome of the strain showed 10797 protein-coding genes, 94 tRNA genes, 1 tmRNA gene and 46 copies of misc_RNA genes.

The genome annotation result obtained by rapid annotations using subsystems technology (RAST) (http://rast.nmpdr.org/) (Aziz et al. 2008; Overbeek et al. 2014) revealed that strain CA3R110[T] possessed 367 subsystems belonging to 24 categories (Fig. S5). Among these subsystems, "amino acids and derivatives" was the largest subsystem (647 feature counts), followed by "carbohydrates" (570 feature counts), "fatty acids, lipids, and isoprenoids" (336 feature counts), "cofactors, vitamins, prosthetic groups, pigments" (285 feature counts), "protein metabolism" (269 feature counts), and "nucleosides and nucleotides" (138 feature counts). Moreover, the "secondary metabolism" subsystem was also detected. The antiSMASH server predicted that strain CA3R110[T] contained several secondary metabolite biosynthesis gene clusters (smBGCs) e.g. type I polyketide synthase (T1PKS), type II polyketide synthase (T2PKS), type III polyketide synthase (T3PKS), terpene, non-ribosomal peptide synthetase (NRPS), NRPS-like, betalactone, butyrolactone, indole, redox-cofactor, lanthipeptide-class-i and ii, PKS-like, heterocyst glycolipid synthase-like PKS (hglE-KS), siderophore, ectoine, non-alpha poly-amino acids like e-Polylysin (NAPAA), homoserine lactone (hserlactone), RRE-element containing cluster, Other unspecified ribosomally synthesised and post-translationally modified peptide product (RiPP-like), lassopeptide, ladderane, arylpolyene, aminocoumarin gene clusters. Among these gene clusters, twenty one clusters showed more than 50 % similarities to known biosynthetic gene clusters. There were coelichelin biosynthetic gene cluster (100 %), azalomycin biosynthetic gene cluster (100 %), nigericin biosynthetic gene cluster (100 %), 2-methylisoborneol biosynthetic gene cluster (100 %), echoside A biosynthetic gene cluster (100 %), pristinol biosynthetic gene cluster (100 %), desferrioxamine B biosynthetic gene cluster (100 %), geosmin biosynthetic gene cluster (100 %), ectoine biosynthetic gene cluster (100 %), coelibactin biosynthetic gene cluster (100 %), elaiophylin biosynthetic gene cluster (87 %), feglymycin biosynthetic gene cluster (78 %), geldanamycin biosynthetic gene cluster (69 %), mediomycin biosynthetic gene cluster (68 %), sceliphrolactam biosynthetic gene cluster (60 %), atratumycin biosynthetic gene cluster (57 %), glycinocin biosynthetic gene cluster (53 %),hopene biosynthetic gene cluster (53 %), griseochelin biosynthetic gene cluster (53 %), 5-isoprenylindole-3-carboxylate-β-D-glycosyl ester biosynthetic gene cluster (52 %), and bafilomycin B1 biosynthetic gene cluster (50 %).

The genomic comparison between strain CA3R110[T] and the type strain of the genus *Streptomyces* indicated that strain CA3R110[T] shared the highest ANIb and ANIm values (92.2 and 93.8%, respectively) with *S. rapamycinicus* NRRL B-5491[T] followed by *S. iranensis* HM 35[T] (92.1% and 93.9%, respectively) which were below an ANI *threshold* range (95−96 %) for *species demarcation* (Richter and Rosselló-Móra 2009; Chun et al. 2018) (Table S3). The phylogenomic tree indicated that strain CA3R110[T] formed a

stable clade with *S. rapamycinicus* NRRL B-5491^T and *S. iranensis* HM 35^T that was supported by 95 % bootstrap value (Fig. S4). As evident from genomic analysis, *S. rapamycinicus* NRRL B-5491^T and *S. iranensis* HM 35^T should be selected for phenotypically comparative study.

Differential characteristics between strain CA3R110^T and the closest neighbors

To clarify the novelty at species level of strain CA3R110^T, the phenotypic properties of strain CA3R110^T and its closest relatives, *S. rapamycinicus* DSM 41530^T and *S. iranensis* DSM 41954^T were compared in the same culture conditions. Unlike *S. rapamycinicus* DSM 41530^T and *S. iranensis* DSM 41954^T, strain CA3R110^T grew in the presence of up to 3 % NaCl. The pH and temperature for cell growth were in the range of 6-9 and 20-40 ^{o}C, respectively. Physiological and biochemical properties showed that strain CA3R110^T significantly differed from *S. rapamycinicus* DSM 41530^T and *S. iranensis* DSM 41954^T. For example, the utilization of L-arabinose, D-mannitol, D-xylose, D-melibiose, D-mannose, L-raffinose, D-ribose, dextran, sucrose, and xylitol as sole carbon sources and the utilization of L-histidine, 4-hydroxyproline, and L-proline were the different points for discriminating between strain CA3R110^T and both type strains. It can also be distinguishable based on the ability to degrade compounds such as adenine, hypoxanthine tyrosine, and gelatin and the ability to produce acid from inulin, lactose, and D-melezitose. Moreover, the composition of sugar of the strain CA3R110^T made it different from diagnostic sugars found in *S. rapamycinicus* DSM 41530^T and *S. iranensis* DSM 41954^T. Unlike *S. rapamycinicus* NRRL B-54917^T and *S. iranensis* DSM 41954^T, strain CA3R110^T was found to have galactose, glucose and ribose, while the closest relatives contained only glucose and ribose in cell-hydrolysates. In addition, the levels of digital DNA-DNA relatedness between strain CA3R110^T and *S. rapamycinicus* NRRL B-5491^T, *S. iranensis* HM 35^T were 55.6% and 57.0%, respectively. These values are below the threshold value of 70 %, as suggested by Wayne et al. (1987) for determining novel species for bacterial strains (Table S3). These phenotypic, chemotaxonomic, and genotypic data indicated that strain CA3R110^T merit classification as a novel species of the genus *Streptomyces*, for which we propose the name *Streptomyces endocoffea* sp. nov.

Description of *Streptomyces endocoffea* sp. nov.

Streptomyces endocoffea (en.do. cof.fe'ae. Gr. *endo* within; N.L. gen. n. *coffeae* of *Coffea arabica*, the coffee plant, *endocoffeae*, pertaining to the original isolation from coffee plant tissue).

Cells are Gram-stain-positive and aerobic. Pale greenish-yellow to dark brownish-gray substrate mycelia is observed. Greyish-white aerial mycelium is produced on ISP2, ISP3, ISP4, ISP5, ISP6, ISP7, Czapek's, nutrient and glucose-asparagine agar media after 21 days of cultivation. Brilliant greenish-yellow diffusible pigments are detected on ISP5 and ISP7. The spiral chain of spore was observed on any media tested. Produce spiral chains of spores with rugose surfaces and are non-motile. The coagulation of milk, urease activity, hydrolysis of starch, gelatin liquefaction, and catalase and oxidase activities are positive, but hydrolysis of casein is weakly positive. Negative for the reduction of nitrate. Utilizes D-cellobiose, D-

galactose, D-glucose, D-fructose, D-mannitol, D-mannose, D-melibiose, L-rhamnose, *myo*-inositol, , dextran, trehalose, and xylitol as a sole carbon source, weakly utilizes D-melezitose, lactose, and inulin but not L-arabinose, D-raffinose, D-ribose, D-xylose, sucrose, and glycerol. Utilizes L-arginine, L-asparagine, L-histidine, 4-hydroxyproline, L-methionine, L-phenylalanine, L-proline and L-threonine as a sole nitrogen source, but not L-cysteine, L-serine, and L-valine. Decomposes adenine, hypoxanthine, tyrosine, but not cellulose, and xanthine. The growth temperature is between 15-40 $^\circ$C while the optimum range is 25-30 $^\circ$C. Maximum NaCl for growth is 3% (w/v). The pH range for growth is 6-9. According to the API ZYM system, it shows alkaline phosphatase, leucine arylamidase, α-chymotrypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-glucosidase, and N-acetyl-β-glucosaminidase activities. Esterase (C4), valine arylamidase, cystine arylamidase, trypsin, β-galactosidase, β-glucosidase, α-mannosidase activities are weak and shows no activities of α-galactosidase, β-glucuronidase, esterase lipase (C8), lipase (C14), or α-fucosidase. Cell wall peptidoglycan contains *LL*-diaminopimelic acid. The major menaquinone is MK-9(H_6) and MK-9(H_8), while MK-9(H_4) and MK-9(H_2) are minor components. Glucose, ribose, and galactose are detected as reducing sugar of the cell. The polar lipid profile contains diphosphatidylglycerol, phosphatidylethanolamine, hydroxy-phosphatidylethanolamine, phosphatidylglycerol, phosphatidylinositol, phosphatidylinositolmannoside (PIM), and six unidentified phospholipids (PLs). The main fatty acids (>10 %) are *iso*-$C_{15:0}$, *iso*-$C_{16:0}$, and $C_{16:0}$. The DNA G+C content of the type strain is 70.9 mol%.

The type strain, CA3R110[T] (= TBRC 11245[T] = NBRC 114296[T]), is an endophytic actinobacterium isolated from the root of *Coffea arabica* L., which grows in Lampang province, Thailand. The GenBank/EMBL/DDBJ accession number for the 16S rRNA gene sequence of strain CA3R110[T] is MN116545. The whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAERRG000000000. The version described in this paper is version JAERRG010000000.

Abbreviations

ANIb, Average Nucleotide Identity-BLAST; ANIm, Average Nucleotide Identity-MUMmer algorithm; DAP, Diaminopimelic acid; dDDH, Digital DNA-DNA Hybridisation; DPG, diphosphatidylglycerol; PE, phosphatidylethanolamine; OH-PE, hydroxy-phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; PL, unidentified phospholipid; NPG, ninhydrin-positive lipid; L, unidentified lipid; GGDC, Genome to Genome Distance Calculator; GGDC, Genome to Genome Distance Calculator; HPLC, high performance liquid chromatography; ISP, International Streptomyces Project; ISCC-NBS, Inter-Society Color Council-the National Bureau of Standards; MK, menaquinone; ML, maximum-likelihood; MP, maximum-parsimony; NJ, neighbour-joining; r.p.m., round per minute; PE, phosphatidylethanolamine; PI, phosphatidylinositol; tetra, tetra nucleotide signature correlation index; TLC, Thin Layer Chromatography; TYGS, Type strain Genomic Server; TYGS, Type strain Genomic Server.

Declarations

Conflicts of interest

The authors declare that there are no conflicts of interest.

Data availability

The DDBJ accession number for the 16S rRNA gene sequence of strain CA3R110^T is MN116545. The whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAERRG000000000. The version described in this paper is version JAERRG010000000.

Compilance with ethical standards

Conflict of interest

The authors declare no conflict of interest.

Ethics approval

This article does not contain any studies with human participants and/or animals performed by any of the authors. Formal consent is not required in this study.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

References

Arai T (1975) Culture Media for Actinomycetes. The Society for Actinomycetes Japan, Tokyo

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al (2008) The RAST Server: Rapid annotations using subsystems technology. BMC Genom 8:9-75.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Com Biol 19:455-477.

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T (2019) antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids ResOxford University Press 47:W81–W87.

Cao P, Li C, Wang H, Yu Z, Xu X, Wang X, Zhao J, Xiang W (2020) Community structures and antifungal activity of root-associated endophytic actinobacteria in healthy and diseased cucumber plants and *Streptomyces* sp. HAAG3-15 as a promising biocontrol agent. Microor 8:236-253.

Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, Costa MS, Rooney AP, Yi H, Xu XW, Meyer SD, Trujillo ME (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol 68:461–466.

Collins MD, Jones D (1980) Lipids in the classification and identification of coryneform bacteria containing peptidoglycans based on 2,4-diaminobutyric acid. J Appl Bacteriol 48:459-470.

Collins MD, Pirouz T, Goodfellow M, Minnikin DE (1977) Distribution of menaquinones in actinomycetes and corynebacteria. J Gen Microbiol 100:221-230.

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368-376.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783-791.

Gordon RE, Barnett DA, Handerhan JE, Pang CHN (1974) *Nocardia coeliaca*, *Nocardia autotrophica*, and the nocardin strain. Int J Syst Bacteriol 24:54-63.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98.

Hasegawa T, Takizawa M, Tanida S (1983) A rapid analysis for chemical grouping of aerobic actinomycetes. J Gen Appl Microbiol 29:319-322.

Kaewkla O and Franco CMM (2021) Genome mining and description of *Streptomyces albidus* sp. nov., an endophytic actinobacterium with antibacterial potential. Antonie Van Leeuwenhoek. Doi: 10.1007/s10482-021-01539-1

Kämpfer P, Kroppenstedt RM (1996) Numerical analysis of fatty acid patterns of coryneform bacteria and related taxa. Can J Microbiol 42:989-1005.

Kelly KL (1964) Inter-Society Color Council – National Bureau of Standard Color Name Charts Illustrated with Centroid Colors. US Government Printing Office, Washington DC

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol l6:111-120.

Komagata K, Suzuki KI (1987) Lipid and cell-wall analysis in bacterial systematics. Methods Microbiol 19:161-207.

Kumar S, Stecher G, Li M, Knyaz C, and Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol 35:1547-1549.

Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60-73.

Meier-Kolthoff JP, Göker M (2019) TYGS is an automated highthroughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10:2182.

Minnikin DE, O'Donnell AG, Goodfellow M, Alderson G, Athalye M, Schaala A, Parletta JH (1984) An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. J Microbiol Methods 2:233-241.

Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al (2014) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST).   Nucleic Acids Res 42:D20-D214.

Peng C, Zhuang X, Gao C, Wang Z, Zhao J, Huang SX, Liu C, Xiang W (2021) *Streptomyces typhae* sp. nov., a novel endophytic actinomycete with antifungal activity isolated the root of cattail (*Typha angustifolia* L.). Anto van Leeu. Doi: 10.1007/s10482-021-01561-3

Pridham TG, Hesseltine CW, Benedict RG (1958) A guide for the classification of   streptomycetes according to selected groups; placement of strains in morphological sections. Appl Microbiol 6:52–79.

Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci USA 106:19126–19131.

Richter M, Rosselló-Móra R, Glöckner FO, Peplies J (2016) JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929-931.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425.

Sasaki T, Igarashi Y, Ogawa M, Furumai T (2002) Identification of 6-prenylindole as an antifungal metabolite of *Streptomyces* sp. TP- A0595 and synthesis and bioactivity of 6-substituted indoles. J. Antibiot 55:1009-1012.

Sasser M (1990) Identification of bacteria by gas chromatography of cellular fatty acids. MIDI Technical Note 101. Microbial ID, Inc, Newark

Shirling EB, Gottlieb D (1966) Methods for characterization of *Streptomyces* species. Int J Syst Bacteriol 16:313-340.

Singh R and Dubey AK (2018) Diversity and applications of endophytic actinobacteria of plants in special and other ecological niches. Front Microbiol 9:1767.

Solecka J, Zajko J, Postek M, Rajnisz A (2012) Biologically active secondary metabolites from Actinomycetes. Cent Eur J Biol 7:373-390.

Tamaoka J (1994) Determination of DNA Base Composition. In: Goodfellow M, O'Donnell AG (eds). Chemical Methods in Prokaryotic Systematics, John Wiley and Sons, Chichester, pp 463-470.

Tamaoka J, Katayama-Fujimura Y, Kuraishi H (1983) Analysis of bacterial menaquinone mixtures by high performance liquid chromatography. J Appl Bacteriol 54:31-36.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526.

Thawai C (2015) *Micromonospora costi* sp. nov., isolated from a leaf of *Costus speciosus*. Int J Syst Evol Microbiol 55:1456-1461.

Waksman SA (1961) The Actinomycetes, vol. 2, Classification, Identification and Description of Genera and Species. Williams and Wilkins, Baltimore

Waksman SA, Henrici AT (1943) The nomenclature and classification of the actinomycetes. J Bacteriol 46:337–341.

Wang Z, Tian J, Li X, Gan L, He L, Chu Y, Tian Y (2018) *Streptomyces dioscori* sp. nov., a novel endophytic actinobacterium isolated from bulbil of *Dioscorea bulbifera* L. Curr Microbiol 10:1384-1390.

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP, Truper HG (1987) Report of the ad hoc committee on reconciliation of

approaches to bacterial systematics. Int J Syst Bacteriol 37:463-464.

Wellington EMH, Stackebrandt E, Sanders D, Wolstrup J, Jorgensen NOG (1992) Taxonomic Status of *Kitasatosporia*, and Proposed Unification with *Streptomyces* on the Basis of Phenotypic and 16S rRNA Analysis and Emendation of *Streptomyces* Waksman and Henrici 1943, 339[AL]. Int J Syst Bact 42:156-160.

Williams ST, Cross T (1971) Actinomycetes. In: Booth C (editor). Methods in Microbiology, vol 4. Academic Press, London, pp 295-334.

Xu Y, He J, Tian XP, Li J, Yang LL, Xie Q, Tang SK, Chen YG, Zhang S, Li WJ (2012) *Streptomyces glycovorans* sp. nov., *Streptomyces xishensis* sp. nov. and *Streptomyces abyssalis* sp. nov., isolated from marine sediments. Int J Syst Evol Microbiol 62:2371-2377.

Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J (2017) Introducing EzBioCloud: A taxonomically united database of 16S rRNA and whole genome assemblies. Int J Syst Evol Microbiol 67:1613-1617.

Zhang J, Wang JD, Liu CX, Yuan JH, Wang XJ, Xiang WS (2014) A new prenylated indole derivative from endophytic actinobacteria *Streptomyces* sp. neau-D50. Nat Prod Res 7:431-437.

Table

Table 1. Differential characteristics of strain CA3R110[T] and its closest phylogenetic relatives.

Characteristics	1	2	3
Urease activity	+	-	-
Coagulation of milk	+	-	+
Maximum NaCl tolerance (%w/v)	3	4	3
Maximum temperature for growth (°C)	40	40	37
The pH range for growth	6-9	6-9	6-11
Carbon utilization:			
D-Raffinose	-	-	+
L-Arabinose	-	-	+
D-Mannitol	+	-	+
D-Ribose	-	-	+
D-Xylose	-	-	+
D-Melibiose	+	-	+
Dextran	+	w	-
D-Mannose	+	-	+
Sucrose	-	-	+
Xylitol	+	w	-
Nitrogen utilization:			
L-Histidine	w	+	-
4-Hydroxyproline	+	-	+
L-Valine	-	+	+
Decomposition of:			
Adenine	+	+	-
Hypoxanthine	+	+	-
Acid production from:			
Inulin	+	-	-
D-Melezitose	+	-	-
Lactose	-	+	+

Strains: 1, CA3R110[T]; 2, *S. rapamycinicus* DSM 41530[T], and 3, *S. iranensis* DSM 41954[T].

[a]All data was done in this study. +, Positive; -, Negative; w, Weakly positive.
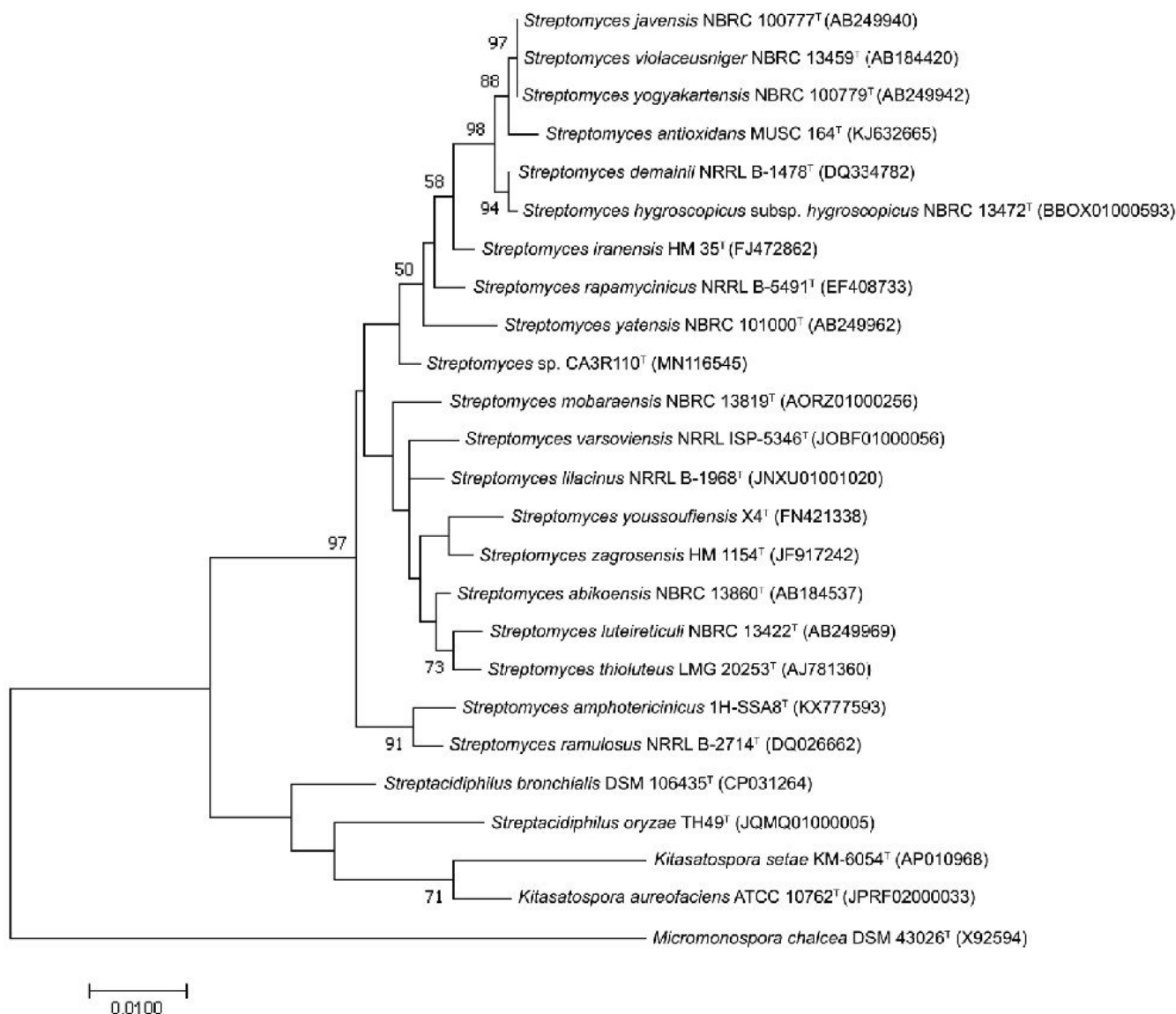
Figures

Figure 1

Maximum-Likelihood tree based on almost-complete 16S rRNA gene sequences showing the relationships between strain CA3R110T, the Streptomyces species with validly published names, members of the genera Streptacidiphilus and Kitasatospora. Micromonospora chalcea DSM 43026T was used as an outgroup. Asterisks (*) indicating the branches of the tree that were also found using the neighbor-joining method. The numbers on the branches indicate the percentage bootstrap values of 1,000 replicates; only values ≥50% are shown. Bar, 0.010 substitutions per nucleotide position.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SubmitSupplementarydataofCA3R11016April2021.pdf