

UCLA

UCLA Previously Published Works

Title

Genome insights into the pharmaceutical and plant growth promoting features of the novel species *Nocardia alni* sp. nov.

Permalink

<https://escholarship.org/uc/item/1pr696wm>

Journal

BMC Genomics, 23(1)

Authors

Nouioui, Imen

Ha, Sung-Min

Baek, Inwoo

et al.

Publication Date

2022-01-21

DOI

10.1186/s12864-021-08257-y

Peer reviewed

RESEARCH

Open Access



Genome insights into the pharmaceutical and plant growth promoting features of the novel species *Nocardia alni* sp. nov

Imen Nouioui^{1*}, Sung-min Ha², Inwoo Baek³, Jongsik Chun^{3,4} and Michael Goodfellow⁵

Abstract

Background: Recent studies highlighted the biosynthetic potential of nocardiae to produce diverse novel natural products comparable to that of *Streptomyces*, thereby making them an attractive source of new drug leads. Many of the 119 *Nocardia* validly named species were isolated from natural habitats but little is known about the diversity and the potential of the endophytic nocardiae of root nodule of actinorhizal plants.

Results: The taxonomic status of an actinobacterium strain, designated ncl2^T, was established in a genome-based polyphasic study. The strain was Gram-stain-positive, produced substrate and aerial hyphae that fragmented into coccoid and rod-like elements and showed chemotaxonomic properties that were also typical of the genus *Nocardia*. It formed a distinct branch in the *Nocardia* 16S rRNA gene tree and was most closely related to the type strains of *Nocardia nova* (98.6%), *Nocardia jiangxiensis* (98.4%), *Nocardia miyuensis* (97.8%) and *Nocardia vaccinii* (97.7%). A comparison of the draft genome sequence generated for the isolate with the whole genome sequences of its closest phylogenetic neighbours showed that it was most closely related to the *N. jiangxiensis*, *N. miyuensis* and *N. vaccinii* strains, a result underpinned by average nucleotide identity and digital DNA-DNA hybridization data. Corresponding taxogenomic data, including those from a pan-genome sequence analysis showed that strain ncl2^T was most closely related to *N. vaccinii* DSM 43285^T. A combination of genomic, genotypic and phenotypic data distinguished these strains from one another. Consequently, it is proposed that strain ncl2^T (= DSM 110931^T = CECT 30122^T) represents a new species within the genus *Nocardia*, namely *Nocardia alni* sp. nov. The genomes of the *N. alni* and *N. vaccinii* strains contained 36 and 29 natural product-biosynthetic gene clusters, respectively, many of which were predicted to encode for a broad range of novel specialised products, notably antibiotics. Genome mining of the *N. alni* strain and the type strains of its closest phylogenetic neighbours revealed the presence of genes associated with direct and indirect mechanisms that promote plant growth. The core genomes of these strains mainly consisted of genes involved in amino acid transport and metabolism, energy production and conversion and transcription.

Conclusions: Our genome-based taxonomic study showed that isolate ncl2^T formed a new centre of evolutionary variation within the genus *Nocardia*. This novel endophytic strain contained natural product biosynthetic gene clusters predicted to synthesize novel specialised products, notably antibiotics and genes associated with the expression of plant growth promoting compounds.

*Correspondence: imen.nouioui@dsmz.de

¹ Department of Microorganisms, Leibniz Institute DSMZ–German Collection of Microorganisms and Cell Cultures, 38124 Braunschweig, Germany

Full list of author information is available at the end of the article



Keywords: Genome mining, Plant growth promoting properties, Polyphasic taxonomy, Putatively novel antibiotics

Introduction

The actinobacterial genus *Nocardia* [1], the type genus of the family *Nocardiaceae* [2] emend. Zhi et al. [3], has a long and convoluted taxonomic history mainly due to an overreliance placed on morphological properties [4, 5]. The application of polyphasic taxonomic procedures led to marked improvements in the classification of nocardiae and related mycolic acid containing actinobacteria [6]. In general, the genus encompasses aerobic, Gram-stain-positive, acid-alcohol-positive, nonmotile, chemoorganotrophic actinobacteria which form rudimentary to extensively branched substrate hyphae that fragment into coccoid to rod-shaped elements, aerial hyphae may only be visible microscopically; the diamino acid of the peptidoglycan is *meso*-diaminopimelic acid (A₂pm), the characteristic whole-organism sugars are arabinose and galactose; diphosphatidylglycerol, phosphatidylethanolamine, phosphatidylinositol and phosphatidylinositol mannosides are the major polar lipids; the fatty acids consist of straight-chain, saturated, unsaturated and 10-methyl (tuberculosteric) components; mycolic acids have 46–64 carbon atoms and up to four double bonds; the predominant respiratory quinone is a hexahydrogenated menaquinone with eight isoprene units where the two end ones are cyclized (MK-8[H6- ω -cyclo]) and the DNA G+C content ranges from 63–72 mol% [5, 7].

Many of the 119 *Nocardia* species with validly published names (<https://www.bacterio.net/>) are recognized using combinations of genotypic and phenotypic properties [7–9]. Most of these taxa are composed of strains isolated from natural habitats but the best-known species contain causal agents of serious suppurative and granulomatous diseases in humans and animals, especially mycetoma and nocardiosis [10–12]. In contrast, *Nocardia vaccinii* produces galls on blueberry plants [13]. Soil is probably the primary reservoir for *Nocardia* strains as they are found in diverse soil types, including acidic forest [14, 15], arid [16], Cerrado [17], karst cave [18], rhizosphere [19, 20] and saline soils [21, 22]. However, they have also been isolated from marine habitats, especially from sponges [23, 24], as well as from the gut of fungus-growing termites [25] and are increasingly being isolated from plant tissue [26], notably from nodules of actinorhizal plants suggesting that they may have a role in promoting plant growth and inhibiting phytopathogens [27, 28]. Two *Nocardia* strains isolated from *Casuarina glauca* nodules induced root nodule-like structures in the original host plant [29].

Nocardiae are an important source of novel antibiotics [30, 31], as exemplified by the production of amicoumabin B from *Nocardia jinanensis* [32], asterobactin from *Nocardia asteroides* [33], brasiliardin A from *Nocardia brasiliensis* [34], nocardicins from *Nocardia uniformis* subsp. *tsuyamanensis* [35] and tubelactomicin A from *Nocardia vinacea* [36]. A comparative survey of nocardial genomes showed that their biosynthetic potential to produce diverse novel natural products is comparable to that of better studied actinobacterial taxa, such as *Amycolatopsis* and *Streptomyces*, thereby making them an attractive source of new drug leads [37]. These researchers showed that *Nocardia* strains from diverse sources, including clinical material, were equally spread across six phylogenetic clades and found that the genomes of the more pathogenic strains were, on average, slightly smaller than those of most of the other genomes (7.4 Mbp against 7.8 Mbp) and contained fewer BGCs (32.5 against 36.5). Similarly, information from the genome of *Nocardia cyriacigeorgica* shows evidence of adaptation from a saprophytic to a pathogenic lifestyle [38].

The present study was designed to establish the taxonomic status of *Nocardia* strain ncl2^T, isolated from a root nodule of an actinorhizal plant, and to determine its biotechnological and ecological potential. The strain was the subject of a genome-based taxonomic study which showed that it formed a new centre of evolutionary variation within the genus *Nocardia*, the name proposed for this organism is *Nocardia alni* sp. nov. with isolate ncl2^T as the type strain. The genomes of *N. alni* and *N. vaccinii* strains contained natural product biosynthetic gene clusters predicted to synthesize novel specialised products, notably antibiotics and genes associated with the expression of plant growth promoting compounds. Statistical comparison between genomic features of the isolate and its taxonomic neighbours were undertaken to establish any positive correlations between them.

Materials and methods

Isolation, maintenance and cultivation

Strain ncl2^T was isolated from a root nodule of an *Alnus glutinosa* plant growing in Leazes Park, Newcastle upon Tyne, UK, as described in Ghodhbane-Gtari et al. [27]. The permission to collect the root nodules was obtained and this study complies with local and national regulations. Sterile lobes of harvested nodules prepared using the procedure described by these workers were placed onto BAP agar plates [39] which were incubated until single actinobacterial colonies were detected. One such

strain was checked for purity and maintained in 35% (w/v) glycerol at -80°C , as was *Nocardia vaccinii* DSM 43285^T obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). Working cultures of these strains were kept on yeast extract-malt extract agar slopes (International *Streptomyces* Project [ISP] medium 2) [40]. Biomass for the chemotaxonomic analyses carried out on the strains was harvested from ISP2 broths shaken at 200 rpm in baffled flasks for 7 days at 28°C . The harvested cells were washed in distilled water and freeze dried.

Chemotaxonomic, cultural and morphological properties

Smears from ISP2 agar slopes of the isolate were Gram-stained using Hucker's modification [41] and the presence of fragmenting branched hyphae sought by light microscopy. The isolate was examined for its ability to grow over a wide range of temperatures (4°C , 10°C , 15°C , 25°C , 28°C , 37°C , 42°C and 45°C) using ISP2 agar as the basal medium. Growth and cultural properties were recorded on GYM (DSMZ medium 65), nutrient agar (NA), peptone-meat extract-glucose agar (DSMZ medium 250) and tryptic soy agar (TSA) after 7 days at 28°C .

The isolate and *N. vaccinii* DSM 43285^T were examined for chemotaxonomic properties known to be of value in nocardial systematics [5, 25]. Standard chromatographic procedures were used to establish the diamino acid of the wall peptidoglycan [42], whole organism sugars [43] and polar lipid profiles [44]. Cellular fatty acids were extracted and methylated after Miller [45], as modified by Kyukendall et al. [46], and analysed by gas chromatography (Agilent instrument, model 6890N). The resultant peaks were identified using the Standard Microbial Identification (MIDI) system, version 4.5 and the ACTINO6 database [47]. Mycolic acids were extracted using the procedure described by Minnikin and Goodfellow [48], purified and their profiles determined by gas chromatography (Agilent instrument, model 6890N).

Genome sequencing

Genomic DNA was extracted from wet biomass of a single colony of isolate ncl2^T grown on ISP2 agar for 10 days at 28°C . The extracted DNA was purified and quantified following the protocol of MicrobesNG, Birmingham (UK). Genomic DNA libraries and sequencing were achieved using an Illumina HiSeq instrument and the 250bp paired end protocol, as used in the service provided by MicrobesNG. The draft genome sequence was annotated using the RAST-SEED webserver with default options [49].

Phylogeny

An almost full length 16S rRNA gene sequence (1523 bp) extracted directly from the draft genome sequence of isolate ncl2^T was deposited in the GenBank databases under accession number MZ014381. The resultant gene sequence was compared with corresponding gene sequences of closely related *Nocardia* strains retrieved from the EzBioCloud server [50]. Phylogenetic trees based on single 16S rRNA genes and corresponding genome sequences were inferred using the Type Strain Genome Server (TYG), the high throughput Genome to Genome Distance Calculator (GGDC) webserver of Meier-Kolthoff et al. [51]. Average nucleotide identity (ortho ANI, [52]) and digital DNA-DNA hybridization (dDDH) [53] similarities were determined between the isolate and its phylogenomic neighbours using the ANI calculator from the EZBioCloud (<http://www.ezbiocloud.net/tools/ani>) and the GGDC webserver, respectively. The type strain of *Nocardia casuarinae*, isolated from root nodules of *Casuarina glauca* [27] was included for comparative purposes, as was that of *Nocardia pseudobrasiliensis* which was isolated from a leg abscess of a patient suffering from ulcerative colitis [54].

Comparative genomic analyses

The genome sequence of isolate ncl2^T was compared with that of *N. vaccinii* NBRC 15922^T (GenBank accession number BDCC00000000.1), its nearest taxonomic neighbour, and with those of other phylogenomic relatives, namely *Nocardia jiangxiensis* NBRC 101359^T (GenBank accession number BAGB00000000.1), *Nocardia miyunensis* NBRC 108239^T (GenBank accession number BDBQ00000000.1), as well as with *N. casuarinae* BMG 51109^T (GenBank accession number JAFQ00000000) and *N. pseudobrasiliensis* DSM 44290^T (GenBank accession number QQBC00000000.1). These strains were from diverse sources, namely from bud-proliferating galls on blueberry [13], the rhizosphere of goose-grass (*Eleusine indica*), a pine forest soil [14], the root nodule of *Casuarina glauca* and a leg abscess of the patient with ulcerative colitis, respectively. The draft genome assemblies of the strains were annotated using the RAST-SEED webserver [49, 55] with default options.

Genome-based species identification was achieved using the TrueBac ID System v1.92, DB:20190603 [<https://www.truebacid.com/>] [56] and the algorithm proposed by Chun et al. [57]. For the comparative genomic analyses, homologous regions in the target genomes were determined to query ORFs using the USE-ARCH program version v8.1.1861, and aligned using a pairwise global alignment [58]. The matched region in the subject contig was extracted and saved as a homolog.

A pairwise gene-to-gene comparison of each genome was conducted using USEARCH and the gene contents among the isolates and related strains compared using the reciprocal homology search tool as described in Chun et al. [59] and Ha et. al [60]. Reciprocal relations are determined if two different genes give bidirectionally top hits to one another. A pairwise orthologous group is defined if the pair of genes has reciprocal relations. The term pairwise orthologous groups (POGs) was coined for these collections of reciprocally linked orthologs. After the initial grouping, partial genes that are grouped due to their short sequence length are targeted for clustering analysis against the POGs using UCLUST (≥ 95 % identity). The coding sequences (CDSs) were classified into groups based on their roles, with reference to orthologous groups (EggNOG 4.5; <http://eggnogdb.embl.de>) [61].

Statistical comparisons of the genomic properties of the isolate and those of closely related strains (genome size, digital G+C content, number and median length of codon sequences) were carried out to determine possible correlations between them.

Phenotypic properties

Strain ncl2^T and *N. vaccinii* DSM 43285^T were screened for a broad range of phenotypic properties, including their ability to metabolize diverse sole carbon and nitrogen sources, to grow in the presence of several concentrations of sodium chloride, at a range of pH values and in the presence of antibiotics using GENIII microplates and an OMNILOG device (Biolog Inc., Hayward, CA, USA), as described previously [8]. The resultant data were analysed using version 1.3.36 of the OPM package [62, 63]. They were also tested for their ability to produce niacin, arylsulfatase after 3 days [64], and to reduce tellurite [65]. All of these tests were carried out in duplicate using a standard inoculum. Enzymatic and additional metabolic properties of the strains were determined using API-ZYM kits and the protocol provided by the manufacturer (Biomérieux, France).

Results and discussion

Chemotaxonomic, cultural and micromorphological properties

Isolate ncl2^T showed a range of phenotypic properties consistent with its classification in the genus *Nocardia* [5]. It is aerobic, Gram-stain-positive and forms extensively branched substrate hyphae and aerial hyphae that fragment into coccoid to rod-like elements. Beige-pinkish aerial mycelia were formed on DSMZ 65; ISP2 and TSA agar after 10 days at 28°C and 37°C (Fig. S1). The strain grew well on DSMZ 65 and ISP2 agar at 28°C and 37°C, but not at 4°C, 10°C, 15°C, 25°C, 42°C

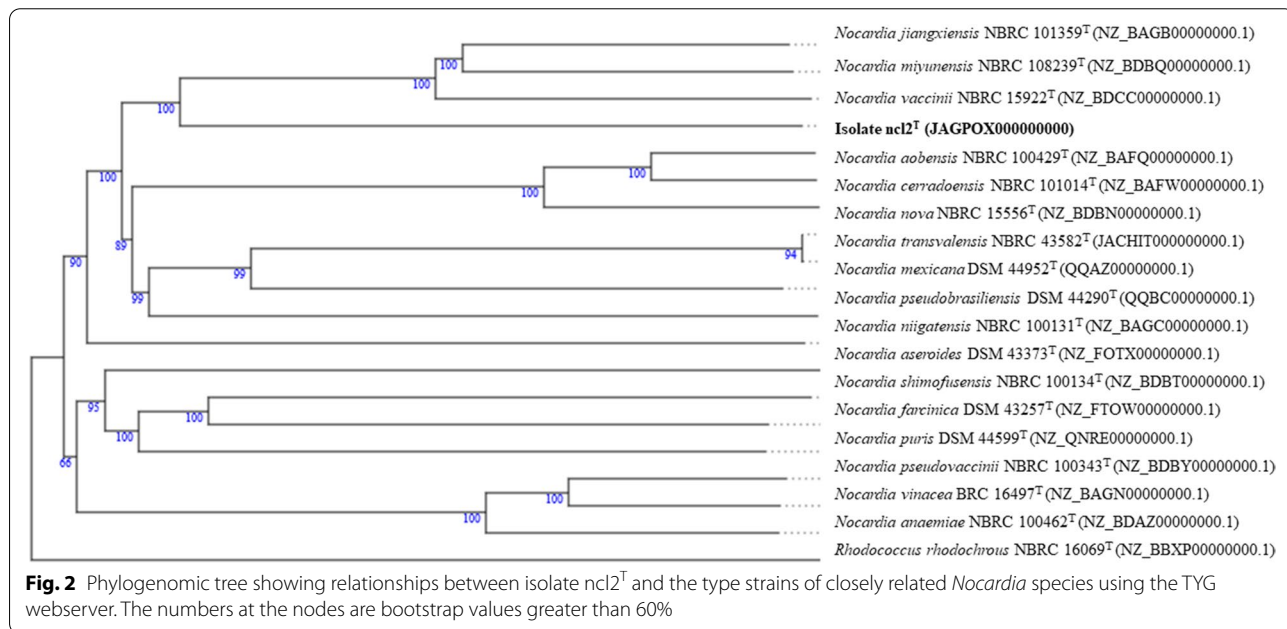
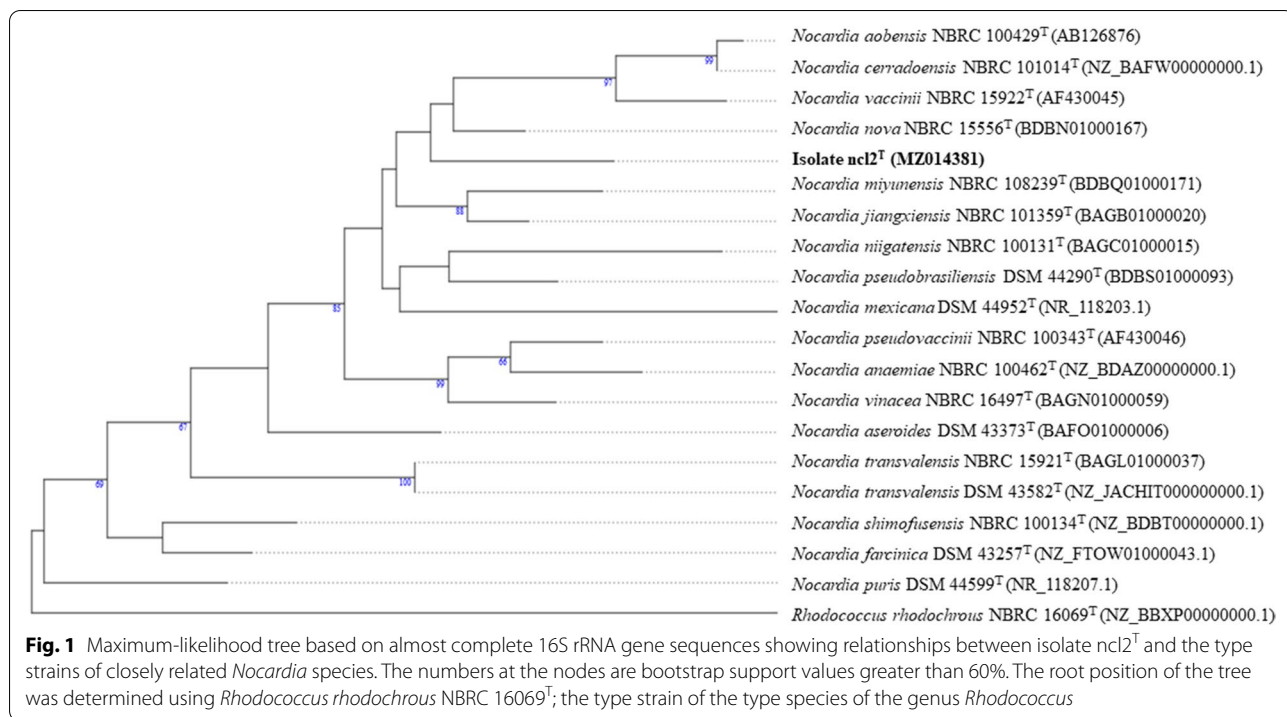
or 45°C, nor did it grow on nutrient agar. It contained *meso*-A₂pm as the diamino acid of the peptidoglycan, arabinose, galactose and glucose as cellular sugars; mycolic acids with 42-62 carbon atoms; the major fatty acids (>10%) were C_{16.0} and C_{18.1} ω9c (Table S1) and the polar lipid pattern consisted of diphosphatidylglycerol (DPG), phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphoglycolipid (PGL), and unidentified phospholipids (PL1-2), aminolipid (AL) and lipids (L1-2) as shown in Fig. S5.

Phylogeny

The phylogenetic tree (Fig. 1) shows that strain ncl2^T forms a distinct branch in the *Nocardia* 16S rRNA gene tree which lies between lineages composed of the type strains of *Nocardia nova* [66], *Nocardia aobensis* [67], *Nocardia cerradoensis* [17], and *N. vaccinii*, and *N. jiangxiensis* and *N. miyunensis*, respectively. The close relationship between the *N. jiangxiensis* and *N. miyunensis* strains has been recorded repeatedly [14, 24, 25, 68, 69]. Isolate ncl2^T was most closely related to *N. nova* NBRC 15556^T, these strains shared a 16S rRNA gene sequence similarity of 98.6%, which corresponds to 20 nucleotides differences. It also showed relatively low sequence similarities with the *N. aobensis* (97.2%), *N. cerradoensis* (97.0%), *N. jiangxiensis* (98.4%), *N. miyunensis* (97.8%) and *N. vaccinii* (97.7%), these relationships were not supported by high bootstrap values (Fig. 1).

Greater confidence can be placed in the topology of phylogenomic tree when compared against corresponding 16S rRNA gene tree as the former is based on millions, not hundreds, of unit characters [70]. The phylogenomic tree (Fig. 2) shows that isolate ncl2^T forms a distinct branch that is most closely related to an evolutionary group composed of the type strains of *N. jiangxiensis*, *N. miyunensis* and *N. vaccinii*, a relationship supported by a 100% bootstrap value. The members of this lineage form a subclade next to a well-supported taxon composed of the type strains of *N. aobensis*, *N. cerradoensis* and *N. nova*.

The recommended threshold used to distinguish between closely related strains based on ANI and DDH values are 95% to 96% and 70%, respectively [57, 71]. Table 1 shows that on this basis isolate ncl2^T can be distinguished from its closest phylogenomic neighbours and from the type strains of *N. casuarinae* and *N. pseudobrasiliensis*. It is also clear that it is most closely related to the *N. jiangxiensis*, *N. miyunensis* and *N. vaccinii* strains although the shared ANI and dDDH similarities are low falling within the range 80.2% to 80.7% and 24.4% to 24.9%, respectively.



Genome features

Key genomic features of the isolate and the associated reference strains are shown in Table 2 The *in silico* G+C content of the strains fall within a narrow range, namely 66.7 to 68.9%. In contrast, the corresponding genomes show more variation ranging from 8.4 up to 10.5 Mbp in size. The isolate is most closely related to *N. vaccinii* DSM

43285^T, as these strains have genomes of a similar size (9.9 Mbp and 9.2 Mbp, respectively), very similar dG+C values (67.0% and 66.7%, respectively), and a similar number of CDCs (8969 and 8502) and tRNA genes (49). In contrast, the *N. jiangxiensis* and *N. miyunensis* strains have larger genomes, namely 10.45 and 10.52 Mbp, respectively, and a similar, albeit larger, number of CDS;

Table 1 Average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) values between strain ncl2^T and its closest phylogenomic neighbours and with the type strains of *N. casuarinae* and *Nocardia pseudobrasiliensis*

Species	Strain designation	dDDH values	ANI values
<i>N. aobensis</i>	NBRC 100429 ^T	22.4	77.9
<i>N. casuarinae</i>	BMG 51109 ^T	23.0	78.9
<i>N. cerradoensis</i>	NBRC 101014 ^T	22.4	77.8
<i>N. jiangxiensis</i>	NBRC 101359 ^T	24.6	80.6
<i>N. miyunensis</i>	NBRC 108239 ^T	24.9	80.7
<i>N. nova</i>	NBRC 15556 ^T	22.4	77.9
<i>N. pseudobrasiliensis</i>	DSM 44290 ^T	22.4	78.7
<i>N. vaccinii</i>	NBRC 15922 ^T	24.4	80.2

these results underpin the close relationship between these strains as found in 16S rRNA gene sequence analyses. The *N. pseudobrasiliensis* strain, a representative of a species associated with invasive human diseases [54], has the smallest genome, thereby providing further evidence that genome sizes of clinically significant nocardiae are lower than corresponding results from non-pathogenic strains [37]. It is also interesting that the type strain of *N. casuarinae*, which induces root nodule formation in *C. glauca*, has a genome size similar to that of *N. pseudobrasiliensis* DSM 44285^T.

A comparison of the taxogenomic features (genome size, digital GC content, number of CDSs, median length of CDS) of strain ncl2^T and the five associated reference strains shows that the number of CDSs is positively correlated with genome size and with the coefficient of determination (R²=0.94), this means that 94% of the data points support the predicted regression lines (y=c[176]+c[896]x), shown in Fig. S2. The frequency plot of the pan genome orthologous groups (POGs) of the strains highlight POGs involved in amino acid and carbohydrate metabolism, information storage and processing (eg. recombination and replication) and cellular processes and signaling (Fig. S3). However, no clear correlation was found between genome size and the number

of orthologous gene groups though the frequency plots of the POGs for the genomes of the *N. casuarinae*, *N. jiangxiensis*, *N. miyunensis* strains were similar with little evidence of quantitative variation.

Comparison of the functional categories between genes in the core and pan genomes using COG/EggNog software gives results in good agreement with those found in the SEED analysis. Apart from genes with unknown function and unassigned categories, the core genomes are composed mainly of genes involved in amino acid transport and metabolism, energy production, and conversion and transcription, as shown in Fig. S4a. However; the pan genome of the strains also contain genes associated with carbohydrate metabolism, clustering based systems and the metabolism of amino acids and derivatives, as presented in Fig. S4b. Around 75% of the core genomes are composed of genes assigned to defined categories whereas less than 40% of those in the pan genomes are associated with functional categories based on COG and SEED pathways. When the strains were examined for strain specific CDSs the highest number were for strain ncl2^T with 2162 and the *N. pseudobrasiliensis* DSM 44290^T with 2187. The highest similarity was between the pan genome of strain ncl2^T and its closest taxogenomic neighbour, *N. vaccinii*, then by the *N. jiangxiensis*, *N. miyunensis* strains which were isolated from acidic soil; no correlation was found between the genome gene contents of the strains and the habitats from which they were isolated.

Phenotypic properties

Identical results were obtained for nearly all of the duplicated cultures included in the phenotypic tests. Table 3 shows that the isolate can be distinguished from *N. vaccinii* DSM 43285^T as it produced niacin, utilized L-alanine and L-serine and an extensive range of organic compounds, including β-gentiobiose, dextrin, D-maltose, D-mannose, pectin, D-raffinose, L-rhamnose, sucrose, D-trehalose and D-turanose, was resistant to fusidic acid, lincomycin, minocycline, troleandomycin and

Table 2 Genomic features of isolate ncl2^T (1) and the type strains of *N. vaccinii* (2), *N. miyunensis* (3), *N. jiangxiensis* (4), its closest phylogenetic relatives, and with those of *Nocardia casuarinae* (5) and *N. pseudobrasiliensis* (6)

Genomic features	1	2	3	4	5	6
Genome size (Mbp)	9.9	9.2	10.5	10.4	8.8	8.4
Digital G+C content	67.0	66.7	66.9	66.7	68.9	67.3
No. of rRNA genes	12	4	4	3	9	8
No. of tRNA genes	49	49	49	49	56	51
No. of CDSs	8969	8502	9567	9695	7746	7949
Median length of CDS (bp)	831	816	831	822	819	822

Table 3 Phenotypic, enzymatic and biochemical properties that distinguish strain ncl2^T from *Nocardia vaccinii* DSM 43285^T

Characteristics	Isolate ncl2 ^T	<i>N. vaccinii</i> DSM 43285 ^T
GEN III Biolog microplate tests		
Sugars:		
D-Cellobiose, D-galactose, β-gentiobiose, D-maltose, D-mannose, D-raffinose, L-rhamnose, sucrose, D-trehalose, and D-turanose	+	-
Polymers:		
Dextrin, pectin	+	-
Tolerance to:		
Lithium chloride, guanidine hydrochloride, sodium bromate, sodium chloride (4% and 8%), tetrazolium blue, and tetrazolium violet	+	-
Sugar alcohols:		
D-Arabitol	+	-
D-Mannitol	-	+
Oxidation of amino acids:		
L-Alanine, L-serine	+	-
L-Aspartic acid	-	+
Resistant to:		
Fusidic acid, lincomycin, minocycline, troleandomycin, and vancomycin	+	-
Sugar acids:		
L-Galactonic acid-γ-lactone, and D-galacturonic acid	+	-
Organic acids		
γ-amino-n-Butyric acid, α-hydroxy-butyric acid, L-lactic acid, D-lactic acid methyl ester, p-hydroxy-phenylacetic acid, methyl pyruvate, and quinic acid,	+	-
α-keto-Glutaric acid, L-malic acid, mucic acid, propionic acid, and bromo-succinic acid,	-	+
API ZYM tests:		
Alkaline phosphatase, α-chymotrypsin, cystine aminopeptidase, esterase (C4), esterase lipase (C8), lipase (C14), trypsin, and valine aminopeptidase	-	+
Other tests:		
Niacin	+	-

+, positive result; -, negative result

Both strains reduced potassium tellurite and gave positive results for acid phosphatase, β-galactosidase α-glucosidase; oxidized, D-glucose, D-gluconic acid, glycerol, and myo-inositol (sugars); acetic acid, acetoacetic acid, butyric acid, β-hydroxy-butyric acid, α-keto-butyric acid, citric acid, and D-saccharic acid, (organic acids); L-arginine, L-glutamic acid and D-serine #2 (amino acids); were resistant to nalidixic acid, and rifamycin sv (antibiotics); and grew in presence of potassium tellurite, and Tween 40 (inhibitory compounds); and in the presence 1% sodium lactate, sodium chloride (1%) (salts) and at pH 5- 6. The strains gave negative results for arylsulfatase after 3 days, D-fucose, L-fucose, D-fructose, D-fructose-6-phosphate, N-acetyl-D-galactosamine, N-acetyl-D-glucosamine, D-glucose-6-phosphate, glucuronamide, 3-O-methyl-D-glucose, β-methyl-D-glucoside, α-D-lactose, D-melibiose, D-salicin, D-sorbitol, and D-stachyose (sugars); D-gluconic acid, D-malic acid, N-acetyl-neuraminic acid, and L-pyroglyutamic acid (organic acids), D-aspartic acid, glycine-proline, L-histidine, and D-serine #1 (amino acids); inosine (nucleoside); aztreonam (antibiotic); gelatin (polymer), niaproof, and sodium formate (inhibitory compounds); α-fucosidase; α-galactosidase, N-acetyl-β-glucosaminidase; β-glucuronidase, and α-mannosidase (API-ZYM tests)

vancomycin, and grew in the presence of lithium chloride, guanidine hydrochloride, sodium bromate, sodium chloride (4% and 8%), tetrazolium blue and tetrazolium violet. In contrast, only the *N. vaccinii* strain utilized α-keto-glutaric acid, L-malic acid, mucic acid, propionic acid and bromo-succinic acid, it also showed greater activity in the API-ZYM tests.

Strains ncl2^T and *N. vaccinii* NBRC 15922^T can also be distinguished using chemotaxonomic and genomic traits. They showed qualitative and quantitative differences in fatty acid profiles, as exemplified by the presence of a trace of C_{17:1}ω9c and a high proportion of

C_{18:0}10-methyl in the *N. vaccinii* strain and a higher proportion of C_{18:1}ω9c in the isolate (24.4% against 9.6%) (Table S1). These properties distinguish them from the type strains of *N. jiangxiensis* and *N. miyunensis* as shown in Table S1. Further, only the isolate contains glucose in whole cell hydrolysates and unidentified aminolipid and glycolipid in its polar lipid profile, unlike *N. vaccinii* DSM 43285^T (Fig. S5). The isolate contains mycolic acids ranging from 42 to 62 carbons with predominant amounts of C42 and C44 whereas the *N. vaccinii* strain has mycolic acids ranging from 52 to 62 carbons with C54, C56, and C58 as major components. Both strains have meso-A₂pm

as the diamino acid of the wall peptidoglycan, arabinose and galactose as diagnostic whole cell sugars and $C_{16:0}$ as the major fatty acids. Furthermore, the genome size of the isolate is larger than that of the *N. vaccinii* strain (9.9 against 9.2 Mbp).

Specialized metabolite biosynthetic gene clusters

Antismash 5.0 [72] predicts NP-BGCs based on the percentage of genes from the closest known bioclusters which show BLAST hits to the genomes of the strains under consideration. The genomes of strain ncl2^T and *N. vaccinii* NBRC 15922^T contained 36 and 29 well-defined bioclusters that are predicted to encode for a broad range of specialized metabolites albeit with low levels of gene identity, as shown in Table S2. The genomes of the strains are well equipped to synthesize non-ribosomal peptide synthetases, type I polyketides, ribosomally synthesized and post-translationally modified peptides, as well as betalactone (3% gene similarity) and carotenoid-like terpene (18% gene similarity) compounds. They have the genetic capacity to produce products most closely related to himastatin (3% gene similarity), an antitumor antibiotic produced by *Streptomyces hygrosopicus* [73], steffimycin D (8% gene similarity), which was initially produced by a *Streptomyces* strain and inhibits ras-oncogen expressed cells [74], and teicoplanin, a product of a *Streptomyces* strain that inhibits growth of Gram-positive bacteria, including *Enterococcus faecalis* and methicillin-resistant *Staphylococcus aureus* (MRSA) strains [75]. The strains also contain bioclusters predicted to synthesise arylpolyene-like compounds that are structurally and functionally similar to caretonoids [76] and which show antimicrobial and antioxidant activity [77]. They also have bioclusters predicted to encode for ectoine (100% gene similarity), a protective molecule which enables bacteria to survive extreme conditions [78].

The genome of strain ncl2^T, unlike that of *N. vaccinii* NBRC 15922^T, contains presumptive NP-BGCs associated with the synthesis of amycolamycin A and B (2% gene similarity), type I polyketides that are cytotoxic for breast cancer cell lines [79], pepticinnamin E (6% gene similarity) which has the potential to treat cancer and malaria [80] and echosides A, B, C, D and E (11% gene similarity) and stambomycin A to D (36% gene similarity), which are antitumor antibiotics [81]. Strain ncl2^T also contains bioclusters predicted to encode for the antitubercular polyketides atratumycin (7% gene similarity) and capreomycin A and B (6% gene similarity), which are produced by '*Streptomyces aratus*' and '*Streptomyces capreolus*,' respectively [32, 82]. Additional presumptive bioclusters are linked to the synthesis of lasolocid (3% gene similarity), a betalactone produced by '*Streptomyces lasaliensis*,' which has antibacterial and

coccidiostatic properties [83] and tiacumicin B (3% gene similarity), a type I polyketide active against nosocomial diarrhea caused by *Clostridium difficile* [84]. Other bioclusters are predicted to encode for the antifungal agents fengycin (3% gene similarity) and nystatin (31% gene similarity) that are produced by *Bacillus subtilis* and *Streptomyces noursei* strains, respectively [85–87].

The genome of the *N. vaccinii* strain harbours several bioclusters absent from that of strain ncl2^T, including ones predicted to encode for polyketides (Table S2). These bioclusters include cyphomycin (2% gene similarity), which is produced by a *Streptomyces* strain and is used to control multidrug resistant fungal pathogens [88] and caniferolides A–D, that are synthesized by *Streptomyces caniferus* and inhibit the growth of *Candida albicans* and *Aspergillus* spp. [89]; caniferolide A has been used to treat Alzheimer's disease [90]. Further, the *N. vaccinii* strain has the genetic capacity to synthesise cremimycin MJ635-86F5 and tetarimycins A and B, these antibiotics are produced by *Streptomyces* strains and show activity against MRSA strains [91, 92]. In addition, the strain contains a presumptive biocluster associated with the expression of leinamycin (2% gene similarity), a betalactone terpene produced by *Streptomyces atroolivaceus* which shows antibacterial and antitumor activity [93].

It can be concluded that strain ncl2^T and *N. vaccinii* NBRC 15928^T have genomes rich in NP-BGCs, notably ones predicted to express for putatively novel polyketide and non-ribosomal peptide compounds thereby providing further evidence that nocardiae are a potentially prolific source of new bioactive compounds [37]. It is particularly interesting that these strains have the capacity to synthesise antifungal and antibiotics given their association with plant tissues. Clearly, nocardiae should feature more prominently in natural product discovery campaigns.

Plant growth promoting properties

Comparative genome mining of strain ncl2^T and the type strains of *N. jiangxiensis*, *N. miyunensis* and *N. vaccinii*, its closest phylogenomic neighbours, revealed the presence of genes associated with direct (eg. phosphate solubilization, phytohormone production) and indirect (eg. lytic enzyme and siderophore production) mechanisms that promote plant growth. *Nocardia casuarinae* BMG51109^T and *N. pseudobrasiliensis* DSM 44290^T were included in these analyses to represent taxa isolated from plant and clinical sources, respectively [27, 54].

Microbes have a pivotal role in making phosphorus available to plants [94] either enzymatically [95] or by producing organic acids and siderophores and other molecules that solubilize inorganic phosphate [96, 97].

The genome of all of the strains, apart from that of the *N. pseudobrasiliensis* DSM 44290^T, contained genes associated with phosphate regulation and metabolism (Table S3). These included gene *ppx-gppA*, which is responsible for the solubilization of inorganic polyphosphate [98] and gene *pstS* which encodes for phosphate binding protein PstS that is involved in the production of the phosphate ABC transporter [99]. The *pstS* gene was not detected in the genome of the clinical isolate thereby suggesting a possible correlation between the environmental origin of the other strains, namely soil and plant tissues, and phosphate metabolism. The genome of all of the strains contained gene *senX3* which is associated with the production of histidine kinase, a high affinity phosphate transporter which has a role in controlling the phosphate regulon [100].

Phytohormones have a central role in plant growth [101], notably indole-3-acetic acid (IAA) and ethylene; the levels of these and other auxins in plants can be regulated by soil microorganisms able to synthesize them [102]. The genome of all of the strains contained genes encoding for indole-3-glycerol phosphate synthase, the precursor of IAA in the tryptophan biosynthetic pathway in plants [103]. They also contained genes encoding for other components of this pathway, including anthranilate phosphoribosyl transferase (*trpD*), anthranilate synthase (*trpE*), and aminase (*trpA* and *B*) [104]. Similarly, gene *trpF*, which is associated with the synthesis of anthranilate phosphoribosyl transferase, was present in the genomes of all of the strains, apart from *N. pseudobrasiliensis* DSM 44290^T. Genes *pxl* and *aad*, which encode for pyridoxine 4-dehydrogenase and aryl-alcohol dehydrogenase (NADP (+)) and are involved in auxin signaling pathways, were found in the genomes of strain ncl2^T, *N. jiangxiensis* NBRC 101359^T, *N. miyunensis* NBRC 108239^T and *N. casurinae* BMG51109^T (Table S3). In contrast, the genomes of all of the strains contained genes associated with tricarboxylic acid biosynthesis, as shown in Table S3. However, only the genome of strain ncl2^T contains gene *acc* that encodes for 1-aminocyclopropane-1-carboxylate deaminase, an ACC deaminase which reduces toxicity due to high levels of ethylene in plants caused by plant growth promoting rhizobacteria. This enzyme also regulates ethylene levels produced by the plant by converting ACC to ammonia and α -ketobutyrate [105, 106].

Plant growth promoting microorganisms can also enhance plant growth by modulating biotic stress as they can decrease, neutralize or prevent infections caused by phytopathogens by synthesizing antibiotics and lytic enzymes [107]. The genomes of all of the strains were equipped with genes associated with the production of chitinases and glucoamylases, as shown in Table S3. They

also contained genes involved in the biosynthesis of antibiotics, as exemplified by *fabG*, *bacC2* and *hdhA* which express for 3-oxoacyl-[acyl-carrier-protein] reductase, bacitracin synthase and 7- α -hydroxysteroid dehydrogenase which play a role in the biosynthesis of pentalenolactone, bacitracin and clavulanic acid, respectively [108–110]. Further, the genomes of all of the strains, apart from *N. pseudobrasiliensis* DSM 44290^T, contained gene *auaJ* which encodes for the epoxidase LasC that is involved in the synthesis of lasalocid, a polyether antibiotic [111]. In contrast, only strain ncl2^T contained gene *tcmO* which expresses tetracenomycin polyketide synthesis 8-O-methyl transferase, a gene is associated with tetracenomycin biosynthesis [112].

It can be concluded that while strain ncl2^T is most closely related to the type strains of *N. jiangxiensis*, *N. miyunensis* and *N. vaccinii*, it can be distinguished from them as it forms a distinct branch in the phylogenomic tree, has a distinct fatty acid profile and shares low ANI and dDDH values with them. Genomic features, notably genome size and CDS numbers, show that the strain is most closely related to *N. vaccinii* NBRC15992^T, but can be distinguished from the latter by a wealth of chemotaxonomic, genomic and phenotypic data. It is, therefore, proposed that strain ncl2^T should be recognized as a new species within the genus *Nocardia* for which the name *Nocardia alni* sp. nov. is proposed.

Description of *Nocardia alni* sp. nov.

Nocardia alni (al'ni L. gen. fem. n. *alni*, of *Alnus* (a genus name), referring to the source of the strain, a root nodule of *Alnus glutinosa*)

Aerobic, Gram-stain-positive, nonmotile actinobacterium that forms an extensively branched substrate mycelium and aerial hyphae which fragment into coccoid to rod-like elements. Beige pink aerial hyphae are formed on DSMZ 65, yeast extract-malt extract and tryptic soy agar. Grows from pH 5–7.5 (optimally at pH 7), from 28 to 37°C and in the presence of up to 8% w/v sodium chloride. Produces niacin, reduces potassium tellurite, but is arylsulfatase negative after 3 days. Additional phenotypic properties are shown in Table 3. The diamino acid of the peptidoglycan is meso-A₂pm, the whole cell sugars are arabinose, galactose and glucose and the predominant fatty acids are C_{16:0} and C_{18:1}ω9c. Mycolic acids have 42 to 62 carbon atoms and the polar lipids are diphosphatidylglycerol, phosphatidylethanolamine and phosphatidylinositol, unidentified phosphoglycolipid, phospholipids, lipids, an aminolipid and a glycolipid. The genome size is 9.93 Mbp and the *in-silico* G+C content 67.0%. The genome is rich in biosynthetic gene clusters predicted to encode for new specialised metabolites, notably antibiotics, and with

genes with the capacity to produce products that promote plant growth.

The type strain ncl2^T (DSM 110931^T = CECT 30122^T) was isolated from a root nodule of *Alnus glutinosa* growing in Leazes Park, Newcastle upon Tyne, UK. The GenBank accession number for the 16S rRNA gene and whole genome sequence of the strain were MZ014381 and JAGPOX000000000, respectively.

Conclusions

Novel endophytic nocardiae are being isolated from rhizospheric soil [19, 113], plant roots and stems [26] and from nodules of actinorhizal plants [28–30] as in the case of *N. alni*. Nodular tissues are rich in carbohydrates hence they are excellent habitats for bacteria, including actinobacteria [114, 115]. Filamentous actinobacteria are associated with actinorhizal and legume root nodules, notably novel *Micromonospora* species [116]. The present study suggests that nocardiae, like micromonosporae, have the potential to promote plant growth though eco-physiological studies are needed to establish their interactions with plants, notably their role in root nodules of actinorhizal plants.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-08257-y>.

Additional file 1: Table S1. Fatty acids profiles of strain ncl2^T and the type strain of *N. vaccinii*, its closest phylogenomic neighbour. **Table S2.** Putative natural product biosynthetic gene clusters detected in the draft genome of strain ncl2^T and *N. vaccinii* NBRC 15922^T using Antismash 5.0.0 webserver. **Figure S1.** Culture of isolate ncl2^T grown on medium DSMZ 65 after 10 days incubation at 28°C. **Figure S2.** Statistical comparisons of genomic properties of: 1, strain ncl2^T, 2, *N. vaccinii* NBRC 15922^T, 3, *N. casuarinae* BMG51109^T, 4, *N. jiangxiensis* NBRC 101359^T, 5, *N. miyunensis* NBRC 108239^T and 6, *N. pseudobrasiliensis* DSM 44290^T. **Figure S3.** Frequency plot of POGs showing major COG categories. **Figure S4.** Functional categories found in the (a) COG/EggNog and (b) SEED analyses. **Figure S5.** Two-dimensional TLC plates of polar lipids extracted from isolate ncl2^T (a) and *N. vaccinii* (b) stained with molybdatophosphoric acid (SigmaP1518). Key: DPG, diphosphatidylglycerol; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; GPL, glycerophospholipid; AL1-2, aminolipids, GL1-5, glycolipids and L1-3, lipid. Solvent1: chloroform: methanol: distilled water (65:25:4); solvent 2: chloroform: glacial acetic acid: methanol: distilled water (80:12:15:4).

Additional file 2: Table S3. Pairwise ortholog matrix between the genome sequences of isolate ncl2^T (1) and *N. vaccinii* NBRC 15922^T (2), *N. casuarinae* BMG51109^T (3), *N. jiangxiensis* NBRC 101359^T (4), *N. miyunensis* NBRC 108239^T (5) and *N. pseudobrasiliensis* DSM 44290^T (6).

Acknowledgments

The authors are indebted to Mrs Gabriele Pötter (Leibniz Institute DSMZ–German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany) for help with the chemotaxonomic analyses. Genome sequencing was provided by MicrobesNG (<http://www.microbesng.com>).

Authors' contributions

IN conceived and designed the study. IN, SMH, IB, JC, MG were responsible for analysing the data and preparing the draft paper. All authors read and approved the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. No external funding was received for this study

Availability of data and materials

The 16S rRNA gene and whole genome sequence of strain ncl2^T are available in the GenBank under accession numbers MZ014381 and JAGPOX000000000, respectively.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflict of interest.

Author details

¹Department of Microorganisms, Leibniz Institute DSMZ–German Collection of Microorganisms and Cell Cultures, 38124 Braunschweig, Germany. ²Department of Integrative Biology and Physiology, University of California Los Angeles, 610 Charles E. Young Drive East, Los Angeles CA 90095, USA. ³School of Biological Sciences & Institute of Molecular Biology and Genetics, Seoul National University, Seoul, Korea. ⁴ChunLab, Inc, Seoul, Korea. ⁵School of Natural and Environmental Sciences, Newcastle University, Ridley Building 2, NE1 7RU Newcastle upon Tyne, UK.

Received: 19 July 2021 Accepted: 14 December 2021

Published online: 21 January 2022

References

1. Trevisan VI. Generi e le Specie delle Batteriacee. Milano: Zanaboni and Gabuzzi; 1889.
2. Castellani A, Chalmers AJ. Manual of tropical medicine. 3 edn. New York USA: William Wood and Co; 1919.
3. Zhi XY, Li WJ, Stackebrandt E. An update of the structure and 16S rRNA gene sequence-based definition of higher ranks of the class *Actinobacteria*, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa. Int J Syst Evol Microbiol. 2009;59:589–608.
4. Lechevalier MP. The taxonomy of the genus *Nocardia*: some light at the end of the tunnel. In: Goodfellow M, Brownell GH, Serrano JA, editors. The biology of the Nocardiae. London: Academic; 1976. p. 1–38.
5. Goodfellow M, Maldonado LA. Genus I. *Nocardia* Trevisan 1889AL. In: Goodfellow M, Kämpfer P, Busse H-J, Trujillo M-E, Suzuki K-I, Ludwig W, Whitman WB, editors. Bergey's manual of systematic bacteriology, the *Actinobacteria*, 2nd ed. parts A and B. New York: Springer; 2012. p. 376–419.
6. Goodfellow M, Isik K, Yates E. Actinomycete systematics: an unfinished synthesis. Nova Acta Leopold. 1999;312:47–82.
7. Tamura T, Ohji S, Ichikawa N, Hosoyama A, Amazoe A, Hamada M, Komaki H, Shibata C, Matsuzawa T, Gono T, Suzuki K-I, Fujita N. Reclassification of *Nocardia* species based on whole genome sequence and associated phenotypic data. J Antibiot. 2018;71:633–41.
8. Nouioui I, Cortés-Albayay C, Neumann-Schaal M, Vicente D, Cilla G, Klenk H-P, Marimón JM, Ercibengoa M. Genomic virulence features of two novel species *Nocardia barduliensis* sp. nov. and *Nocardia gipuzkoensis* sp. nov., isolated from patients with chronic pulmonary diseases. Microorganism. 2020;8:1517.

9. Zhang X, Zhang L, Yu XY, Zhang J, Jiao Y, Ju H, Wang X, Zhao J, Xiang W. *Nocardia bovisericoris* sp. nov., an actinobacterium isolated from cow dung. *Int J Syst Evol Microbiol*. 2021;71:004699.
10. Conville PS, Brown-Elliott BA, Smith T, Zelazny AM. The complexities of *Nocardia* taxonomy and identification. *J Clin Microbiol*. 2018;56:e01419-17.
11. Kosova-Maali D, Bergeron E, Maali Y, Durand T, Gonzalez J, Mounié D, Sandoval H, Trujillo M, Boiron P, Salinas-Carmona M-C, Rodriguez-Nava V. High intraspecific genetic diversity of *Nocardia brasiliensis*, a pathogen responsible for cutaneous nocardiosis found in France: phylogenetic relationships by using *sod* and *hsp65* genes. *Biomed Res Int*. 2018;20:7314054.
12. Mehta HH, Shamoo Y. Pathogenic *Nocardia*: A diverse genus of emerging pathogens or just poorly recognized? *Plos Pathog*. 2020;16:1–7.
13. Demaree JB, Smith NR. *Nocardia vaccinii* n. sp. causing galls on blueberry plants. *Phytopathology*. 1952;42:249–52.
14. Cui Q, Wang L, Huang Y, Liu Z, Goodfellow M. *Nocardia jiangxiensis* sp. nov. and *Nocardia miyunensis* sp. nov., isolated from acidic soils. *Int J Syst Evol Microbiol*. 2005;56:1921–5.
15. Golinska P, Wang D, Goodfellow M. *Nocardia aciditolerans* sp. nov., isolated from a spruce forest soil. *Antonie van Leeuwenhoek*. 2013;103:1079–88.
16. Sazak A, Sahin N, Camas M. *Nocardia goodfellowii* sp. nov. and *Nocardia thraciensis* sp. nov., isolated from soil. *Int J Syst Evol Microbiol*. 2012;62:1228–34.
17. de Barros Albuquerque EVS, Manfio GP, Maitan VR, Mendes Bataus LA, Kim SB, Maldonado LA, Goodfellow M. *Nocardia cerradoensis* sp. nov., a novel isolate from Cerrado soil in Brazil. *Int J Syst Evol Microbiol*. 2003;53:29–33.
18. Fang BZ, Han MX, Zhang LY, Jiao JY, Zhang XT, Zhang ZT, Wang Y, Nie GX, Li WJ. *Nocardia aurea* sp. nov., a novel actinobacterium isolated from a karstic subterranean environment. *Int J Syst Evol Microbiol*. 2019;69:159–64.
19. Wang Y, Liu W, Feng WW, Zhou XQ, Bai JL, Yuan B, Ju XY, Cao CL, Huang Y, Jiang JH, LV A-J, Qin S. *Nocardia rhizosphaerae* sp. nov., a novel actinomycete isolated from the coastal rhizosphere of *Artemisia* Linn., China. *Antonie Van Leeuwenhoek*. 2015;108:31–9.
20. Ding P, Bai JL, Wang TT, Sun Y, Cao CL, Jiang JH, Qin S. *Nocardia rhizosphaerihabitans* sp. nov., a novel actinomycete isolated from a coastal soil. *Int J Syst Evol Microbiol*. 2018;68:192–7.
21. Nikou MM, Ramezani M, Amoozegar MA, Rasooli M, Harirchi S, Shahzadeh Fazeli SA, Schumann P, Sproer C, Ventosa A. *Nocardia halotolerans* sp. nov., a halotolerant actinomycete isolated from saline soil. *Int J Syst Evol Microbiol*. 2015;65:3148–54.
22. Bai JL, Wang Y, Qin S, Ding P, Xing K, Yuan B, Cao CL, Huang Y, Zhang YQ, Jiang JH. *Nocardia jiangsuensis* sp. nov., an actinomycete isolated from coastal soil. *Int J Syst Evol Microbiol*. 2019;66:4633–8.
23. Zhang H, Zhang W, Jin Y, Jin M, Yu X. Comparative study on the phylogenetic diversity of culturable actinobacteria isolated from five marine sponge species. *Antonie Van Leeuwenhoek*. 2008;93:241–8.
24. Thawai C, Rungjindamai N, Klanbut K, Tanasupawat S. *Nocardia xestospongiae* sp. nov., isolated from a marine sponge in the Andaman Sea. *Int J Syst Evol Microbiol*. 2017;67:1451–6.
25. Benndorf R, Schwitalla JW, Martin K, de Beer ZW, Vollmers J, Kaster AK, Poulsen M, Beemelmanns C. *Nocardia macrotermidis* sp. nov. and *Nocardia aurantia* sp. nov., isolated from the gut of the fungus-growing termite *Macrotermes natalensis*. *Int J Syst Evol Microbiol*. 2020;70:5226–34.
26. Xing K, Qin S, Fei SM, Lin Q, Bian GK, Miao Q, Wang Y, Cao CL, Tang SK, Jiang JH, Li W-J. *Nocardia endophytica* sp. nov., an endophytic actinomycete isolated from the oil-seed plant *Jatropha curcas* L. *Int J Syst Evol Microbiol*. 2011;61:1854–8.
27. Ghodhbane-Gtari F, Nouioui I, Salem K, Ktari A, Montero-Calasanz MDC, Tisa LS, Klenk H-P, Gtari M. *Nocardia casuarinae* sp. nov., an actinobacterial endophyte isolated from root nodules of *Casuarina glauca*. *Antonie van Leeuwenhoek*. 2014;105:1099–106.
28. Ghodhbane-Gtari F, Essoussi I, Chattaoui M, Chouaib B, Jaouani A, Daffonchio D, Boudabous A, Gtari M. Isolation and characterization of non-*Frankia* actinobacteria from root nodules of *Alnus glutinosa*, *Casuarina glauca* and *Elaeagnus angustifolia*. *Symbiosis*. 2010;50:51–7.
29. Ghodhbane-Gtari F, Nouioui I, Hezbri K, Lundstedt E, D'Angelo T, McNutt Z, Laplaze L, Gherbi H, Vaissayre V, Svistoonoff S, ben Ahmed H, Boudabous A, Tisa LS. The plant-growth-promoting actinobacteria of the genus *Nocardia* induces root nodule formation in *Casuarina glauca*. *Antonie van Leeuwenhoek*. 2019;112:75–90.
30. Komaki H, Ichikawa N, Hosoyama A, Takahashi-Nakaguchi A, Matsuzawa T, Suzuki KI, Fujita N, Gono T. Genome based analysis of type-I polyketide synthase and non-ribosomal peptide synthetase gene clusters in seven strains of five representative *Nocardia* species. *BMC Genomics*. 2014;5:323.
31. Dhakal D, Rayamajhi V, Mishra R, Sohng JK. Bioactive molecules from *Nocardia*: diversity, bioactivities and biosynthesis. *J Ind Microbiol Biotechnol*. 2019;46:385–407.
32. Sun W, Zhang YQ, Huang Y, Zhang YQ, Yang ZY, Liu ZH. *Nocardia jinanensis* sp. nov., an amicoumacin B-producing actinomycete. *Int J Syst Evol Microbiol*. 2009;59:417–20.
33. Nemoto A, Hoshino Y, Yazawa K, Ando A, Mikami Y, Komaki H, Graefe U. Asterobactin, a new siderophore group antibiotic from *Nocardia asteroides*. *J Antibiot*. 2002;55:593–7.
34. Komaki H, Nemoto A, Tanaka Y, Takagi H, Yazawa K, Mikami Y, Shigemori H, Kobayashi J, Ando A, Nagata Y, Brasilicardin A. A new terpenoid antibiotic from pathogenic *Nocardia brasiliensis*: fermentation, isolation and biological activity. *J Antibiot (Tokyo)*. 1999;52:13–9.
35. Hosoda J, Konomi T, Tani N, Aoki H, Imanaka H. Isolation of new nocardins from *Nocardia uniformis* subsp *tsuyamanensis*. *Agri Biol Chem*. 1977;41:2013–20.
36. Kinoshita N, Homma Y, Igarashi M, Ikeno S, Hori M, Hamada M. *Nocardia vinacea* sp. nov. *Actinomycetologica*. 2001;15:1–5.
37. Männle D, McKinnie SMK, Mantri SS, Steinke K, Lu Z, Moore BS, Ziemert N, Kayser L. Comparative genomics and metabolomics in the genus *Nocardia*. *mSystems*. 2020;5:e00125–20.
38. Zoropogui A, Pujic P, Normand P, Barbe V, Belli P, Graindorge A, Roche D, Vallenet D, Mangenot S, Boiron P, Rodriguez-Nava V, Ribun S, Richard Y, Cournoyer B, Blaha D. The *Nocardia cryiacigeorgica* GUH-2 genome shows ongoing adaptation of an environmental actinobacteria to a pathogen's lifestyle. *BMC Genomics*. 2013;14:286–304.
39. Gtari M, Brusetti L, Gharbi S, Diego M, Boudabous A, Daffonchio D. Isolation of *Elaeagnus*-compatible *Frankia* from soils collected in Tunisia. *FEMS Microbiol Lett*. 2004;234:349–55.
40. Shirling E, Gottlieb D. Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol*. 1966;16:313–40.
41. Manual of microbiological methods. Society of American Bacteriologists. New York: McGraw-Hill; 1957.
42. Stanek JL, Roberts GD. Simplified approach to identification of aerobic actinomycetes by thin-layer chromatography. *Applied Microbiol*. 1974;28:226–31.
43. Lechevalier MP, Lechevalier H. Chemical composition as a criterion in the classification of aerobic actinomycetes. *Int J Syst Bacteriol*. 1970;20:435–43.
44. Minnikin D, O'Donnell A, Goodfellow M, Alderson G, Athalye M, Schaal A, Parlett J. An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *J Microbiol Methods*. 1984;2:233–41.
45. Miller LT. Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids. *J Clin Microbiol*. 1982;16:584–6.
46. Kuykendall LD, Roy MA, O'Neill JJ, Devine TE. Fatty acids, antibiotic resistance and deoxyribonucleic acid homology groups of *Bradyrhizobium japonicum*. *Int J Syst Evol Microbiol*. 1988;38:358–61.
47. Sasser M. Identification of bacteria by gas chromatography of cellular fatty acids. MIDI Inc Technical Notes. 1990;101:1.
48. Minnikin DE, Goodfellow M. In: Goodfellow M, Brownell GH, Serrano JA, editors. The Biology of the Nocardia Lipid composition in the classification and identification of nocardiae and related taxa. London: Academic Press; 1976. p. 160–219.
49. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formisma K, Gerdes SY, Glass E, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*. 2008;9:75.
50. Yoon S-H, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J. Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene

- sequences and whole-genome assemblies. *Int J Syst Evol Microbiol.* 2017;67:1613–7.
51. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun.* 2019;10:2182.
 52. Lee I, Kim YO, Park S-C, Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *IJSEM.* 2016;66:1100–3.
 53. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinform.* 2013;14:60.
 54. Ruimy R, Riegel P, Carlotti A, Boiron P, Bernardin G, Monteil H, Wallace RJJ, Christen R. *Nocardia pseudobrasilensis* sp. nov., a new species of *Nocardia* which groups bacterial strains previously identified as *Nocardia brasilensis* and associated with invasive diseases. *Int J Syst Bacteriol.* 1996;46:259–64.
 55. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. The SEED and the Rapid Annotation of Microbial Genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 2014;42:D206–14 (Database issue).
 56. Ha S-M, Kim CK, Roh J, Byun J-H, Yang S-J, Choi S-B, Chun J, Yong D. Application of the Whole Genome-Based Bacterial Identification System, TrueBac ID, Using Clinical Isolates That Were Not Identified With Three Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) Systems. *Ann Lab Med.* 2019;39:530–6.
 57. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu X-W, Meyer SD, Trujillo ME. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol.* 2018;68:461–6.
 58. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics.* 2010;26:2460–1.
 59. Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY, Haley BJ, Taviani E, Jeon Y, Kim DW, Lee J, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Munk AC, Chertkov O, Meincke L, Saunders E, Walters RA, Huq A, Nair GB, Colwell RR. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *PNAS.* 2009;106:15442–7.
 60. Ha S-M, Chailita M, Yang S-J, Yoon S-H, Cho K, Seong WK, Hong S, Kim J, Kwak H.-S., Chun J. Comparative genomic analysis of the *Vibrio cholerae* outbreak in South Korea. *Front Public Heal.* 2016;2019:7.
 61. Powell S, Forslund K, Szklarczyk D, Trachana K, Rot A, Huerta-Cepas J, Gabaldón T, Rattei T, Creevey C, Kuhn M, Jensen LJ, Mering C, Bork P. eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Res.* 2014;42:D231–9.
 62. Vaas LAI, Sikorski J, Hofner B, Buddruhs N, Fiebig A, Klenk H-P, Göker M. opm: An R package for analysing OmniLog[®] Phenotype MicroArray Data. *Bioinformatics.* 2013;29:823–1824.
 63. Vaas LAI, Sikorski J, Michael V, Göker M, Klenk H-P. Visualization and curve-parameter estimation strategies for efficient exploration of phenotype microarray kinetics. *PlosOnes.* 2012;7:e34846.
 64. Kent PT, Kubica GP. *Public Health Mycobacteriology a Guide for the Level III Laboratory.* Atlanta GA USA: CDC Prevention; 1985.
 65. Kilburn JO, Silcox VA, Kubica G. Differential identification of mycobacteria. V. The tellurite reduction test. *Am Rev Res Dis.* 1969;99:94–100.
 66. Tsukamura M. Numerical analysis of the taxonomy of nocardiae and rhodococci. *Microbiol Immunol.* 1982;26:1101–19.
 67. Kageyama A, Suzuki S, Yazawa K, Nishimura K, Kroppenstedt RM, Mikami Y. *Nocardia aobensis* sp. nov., isolated from patients in Japan. *Microbiol Immunol.* 2004;48:817–22.
 68. Kämpfer P, Huber B, Buczolits S, Thummes K, Grun-Wollny, Busse HJ. *Nocardia acidivorans* sp. nov., isolated from soil of the island of Stromboli. *Int J Syst Evol Microbiol.* 2007;57:1183–7.
 69. Zhao GZ, Li J, Zhu WY, Klenk HP, Xu LH, Li WJ. *Nocardia artemisiae* sp. nov., an endophytic actinobacterium isolated from a surface-sterilized stem of *Artemisia annua* L. *Int J Syst Evol Microbiol.* 2011;61:2933–7.
 70. Nouioui I, Carro L, García-López M, Meier-Kolthoff JP, Woyke T, Kyripides NC, Pukall R, Klenk H-P, Goodfellow M, Göker M. Genome-based taxonomic classification of the phylum *Actinobacteria*. *Front Microbiol.* 2018;9:2007.
 71. Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP, Truper HG. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol.* 1987;37:463–4.
 72. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Ac Res.* 2019;47:W81–7.
 73. Leet JE, Schroeder DR, Golik J, Matson JA, Doyle TW, Hill SE, Lee MS, Whitney JL, Krishnan BS. Himastatin a new antitumor antibiotic from *Streptomyces hygroscopicus*. III Structural Elucidation *Antibio.* 1996;49:299–311.
 74. Suzukake-Tsuchiya K, Moriya Y, Yamazaki K, Hori M, Hosokawa N, Sawa T, Iinuma H, Naganawa H, Imada C, Hamada M. Screening of antibiotics preferentially active against ras oncogene-expressed cells. *J Antibiot (Tokyo).* 1990;43:1489–96.
 75. Maple PA, Hamilton-Miller JM, Brumfitt W. Comparative in-vitro activity of vancomycin, teicoplanin, ramoplanin (formerly A16686), paldimycin, DuP 721 and DuP 105 against methicillin and gentamicin resistant *Staphylococcus aureus*. *J Antimicrob Chemother.* 1989;23:517–25.
 76. Schoner TA, Gassel S, Osawa A, Tobias NJ, Okuno Y, Sakagibara Y, Shindo K, Sandmann G, Bode HB. Aryl polyenes, a highly abundant class of bacterial natural products are functionally related to antioxidative carotenoids. *Chembiochem.* 2016;17:247–53.
 77. Rao MPN, Xiao M, Li W-J. Fungal and bacterial pigments: secondary metabolites with wide applications. *Front Microbiol.* 2017;8:1113.
 78. Graf R, Ansali S, Buenger J, Pfluecker R, Drillier H. the multifunctional role of ectoine as a natural cell protectant. *Clin Dermatol.* 2008;26:326–33.
 79. Ma SY, Xiao YS, Zhang B, Shao FL, Guo ZK, Zhang JJ, Jiao RH, Sun Y, Xu Q, Tan RX, Ge HM. Amycolamycins A and B, two enediyne-derived compounds from a locust-associated actinomycete. *Org Lett.* 2017;19:6208–11.
 80. Santa Maria KC, Chan AN, O'Neill EM, Li B. Targeted rediscovery and biosynthesis of the farnesyl-transferase inhibitor peptidocinnamin E. *ChemBioChem.* 2019;20:1387–93.
 81. Laureti L, Song L, Huang S, Corre C, Leblond P, Challis GL, Aigle B. Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambifaciens*. *PNAS.* 2011;108:6258–63.
 82. Fu LM, Shinnick TM. Genome-wide exploration of the drug action of capreomycin on *Mycobacterium tuberculosis* using Affymetrix oligonucleotide Gene Chips. *J Infect.* 2007;54:277–84.
 83. Westley JW, Evans RH, Williams T, Stempel A. Structure of antibiotic X-537A. *Chem Commun.* 1970;1970:71–2.
 84. Hochlowski JE, Swanson SJ, Ranfranz LM, Whittern DN, Buko AM, McAlpine JB. Tiacumicins, a novel complex of 18-membered macrolides. II. Isolation and structure determination. *Antibiotics.* 1987;40:575–88.
 85. Vanittanakom N, Loeffler W, Koch U, Jung G. Fengycin—a novel antifungal lipopeptide antibiotic produced by *Bacillus subtilis* F-29-3. *J Antibiot (Tokyo).* 1986;39:888–901.
 86. Espinel-Ingroff AV. *Medical mycology in the United States: a historical analysis (1894-1996).* Springer, pp 62, 2003.
 87. Zofia E, Leszczynska-Bakal H, Pawlak E, Smarzyński T. Gel with nystatin for treatment of lung mycosis. *Pol J Pharmacol Pharm.* 1976;28:49–52.
 88. Cheverette MG, Carlson CM, Ortega HE, Thomas C, Ananiev GE, Barns KJ, Book AJ, Cagnazzo J, Carlos C, Flanigan W, Grubbs KJ, Horn HA, Hoffmann FM, Klassen JL, Knack JJ, Lewin GR, McDonald BR, Muller L, Melo WGP, Pinto-Tomás AA, Schmitz A, Wendt-Pienkowski E, Wildman S, Zhao N, Zhang F, Bugni TS, Andes DR, Pupo MT, Currie CR. The antimicrobial potential of *Streptomyces* from insect microbiomes. *Nat Commun.* 2019;10:516.
 89. Pérez-Victoria I, Oves-Costales D, Lacrete R, Martín J, Sánchez-Hidalgo M, Díaz C, Cautain B, Vicente F, Genilloud O, Reyes F. Structure elucidation and biosynthetic gene cluster analysis of caniferolides A–D, new bioactive 36-membered macrolides from the marine-derived *Streptomyces caniferus* CA-271066. *Org Biomol Chem.* 2019;17(11):2954–71.
 90. Alvario R, Alonso E, Lacrete R, Oves-Costales D, Genilloud O, Reyes F, Alfonso A, Botana LM. caniferolide A, a macrolide from *Streptomyces caniferus*, attenuates neuroinflammation, oxidative stress, amyloid-beta, and tau pathology in vitro. *Mol Pharm.* 2019;16:1456–66.

91. Kallifidas D, Kang H-S, Brady SF. Tetarimycin A, an MRSA-active antibiotic identified through induced expression of environmental DNA gene clusters. *J Am Chem Soc.* 2012;134:19552–5.
92. Igarashi M, Tsuchida T, Kinoshita N, Kamijima M, Sawa R, Sawa T, Nagawana H, Hamada M, Takeuchi T, Yamazaki K, Ishizuka M. Cremimycin, a novel 19-membered macrocyclic lactam antibiotic, from *Streptomyces* sp. *J Antibiot.* 1998;51:123–9.
93. Kara M, Asano K, Kawamoto I, Takiouchi T, Katsumata S, Takahashi K-I, Nakano H. Leinamycin, a new antitumor antibiotic from *Streptomyces*; Producing organism, fermentation and isolation. *J Antibiot.* 1989;42:1768–74.
94. Zhu F, Qu L, Hong X, Sun X. Isolation and characterization of a phosphate-solubilizing halophilic *Bacterium kushneria* sp. YCWA18 from Daqiao Saltern on the Coast of Yellow Sea of China. *Evid Based Comp Altern Med.* 2011;2011:615032.
95. Rossolini GM, Schippa S, Riccio ML, Berlutti F, Macaskie LE, Thaller MC. Bacterial nonspecific acid phosphohydrolases: physiology, evolution and use as tools in microbial biotechnology. *Cell Mol Life Sci.* 1998;54:833–50.
96. Rodriguez H, Fraga R. Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnol Adv.* 1999;17:319–39.
97. Sharma SB, Sayyed RZ, Trivedi MH, Gobi TA. Phosphate solubilizing microbes: sustainable approach for managing phosphorus deficiency in agricultural soils. *Springerplus.* 2013;2:587.
98. Müller WEG, Schröder HC, Wang X. Inorganic polyphosphates as storage for and generator of metabolic energy in the extracellular matrix. *Chem Rev.* 2019;119:12337–74.
99. Willsky GR, Malamy MH. Characterization of two genetically separable inorganic phosphate transport systems in *Escherichia coli*. *J Bacteriol.* 1980;144:356–65.
100. Glover RT, Kriakov J, Garforth S, Baughn AD, Jacobs WR Jr. The Two-component regulatory system senX3-regX3 regulates phosphate-dependent gene expression in *Mycobacterium smegmatis*. *J Bacteriol.* 2007;189:5495–503.
101. Carro L, Nouioui I. Taxonomy and systematics of plant probiotic bacteria in the genomic era. *AIMS Microbiol.* 2017;3:383–412.
102. Arshad M, Frankenberger WT. Microbial production of plant hormones. *Plant Soil.* 1991;133:1–8.
103. Ouyang J, Shao X, Li J. Indole-3-glycerol phosphate, a branchpoint of indole-3-acetic acid biosynthesis from the tryptophan biosynthetic pathway in *Arabidopsis thaliana*. *Plant J.* 2000;24:327–33.
104. Lambrecht JA, Downs DM. Anthranilate phosphoribosyl transferase (TrpD) generates phosphoribosylamine for thiamine synthesis from enamines and phosphoribosyl pyrophosphate. *ACS Chem Biol.* 2013;8:242–8.
105. Glick BR. The enhancement of plant growth by free-living bacteria. *Can J Microbiol.* 1995;41:109–17.
106. Glick BR, Penrose DM, Li J. A model for the lowering of plant ethylene concentrations by plant growth-promoting bacteria. *J Theor Biol.* 1998;190:63–8.
107. Thomashow LS, Weller DM, Bonsall RF, Pierson LS. Production of the antibiotic phenazine-1-carboxylic acid by fluorescent *Pseudomonas* species in the rhizosphere of wheat. *Appl Environ Microbiol.* 1990;56:908–12.
108. You Z, Omura S, Ikeda H, Cane DE. Pentalenolactone biosynthesis: Molecular cloning and assignment of biochemical function to PtlF, a short-chain dehydrogenase from *Streptomyces avermitilis*, and identification of a new biosynthetic intermediate. *Arch Biochem Biophys.* 2007;459:233–40.
109. Konz D, Klens A, Schörgendorfer K, Marahiel MA. The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. *Chem Biol.* 1997;4:927–37.
110. MacKenzie AK, Kershaw NJ, Hernandez H, Robinson CV, Schofield CJ, Andersson I. Clavulanic acid dehydrogenase: structural and biochemical analysis of the final step in the biosynthesis of the β -lactamase inhibitor clavulanic acid. *Biochemistry.* 2007;46:1523–33.
111. Smith L, Hong H, Spencer JB, Leadlay PF. Analysis of specific mutants in the lasalocid gene cluster: Evidence for enzymatic catalysis of a disfavoured polyether ring closure. *ChemBioChem.* 2008;9:2967–75.
112. Fu H, Alvarez MA, Khosla C, Bailey JE. Engineered biosynthesis of novel polyketides: regiospecific methylation of an unnatural substrate by the tcmO O-methyltransferase. *Biocemistry.* 1996;28:6527–32.
113. Vaddavalli R, Peddi S, Kothagauni SY, Ling VR. *Nocardia bhagyanesis* sp. nov., a novel actinomycete isolated from the rhizosphere of *Callistemon citrinus* (Curtis), India. *Antonie van Leeuwenhoek.* 2014;105:443–50.
114. Muresu R, Polone E, Sulas L, Baldan B, Tondello A, Delogo G, Cappuccinelli P, Alberghini S, Benhizia Y, Benhizia H, Benguedouar A, Mori B, Calamassi R, Dazzo FB, Squartini A. Coexistence of predominantly nonculturable rhizobia with diverse, endophytic bacterial taxa within nodules of wild legumes. *FEMS Microbiol Ecol.* 2008;63:383–400.
115. Benito P, Alonso-Vega P, Aguado C, Luján R, Anzai Y, Hirsch AM, Trujillo ME. Monitoring the colonization and infection of legume nodules by *Micromonospora* in co-inoculation experiments with rhizobia. *Sci Rep.* 2017;7:11051.
116. Trujillo ME, Alonso-Vega P, Rodríguez R, Carro L, Cerda E, Alonso P, Martínez-Molina E. The genus *Micromonospora* is widespread in legume root nodules: the example of *Lupinus angustifolius*. *ISME J.* 2010;201(4):1265–128.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

