

RESEARCH

Open Access



Genomic insights into *Gordonia* diversity: proposal of three novel species (*Gordonia. altitudinis* sp. nov., *Gordonia. ligustrum* sp. nov., and *Gordonia. pistacia* sp. nov.) with distinct environmental and biomedical traits

Cong-Jian Li^{1,2}, Jing Zhang¹, Jing Su¹, Li-Yan Yu¹, Hua-Hong Chen^{2*} and Yu-Qin Zhang^{1*}

Abstract

Background Rare actinomycetes, particularly *Gordonia* spp., are emerging as critical sources of bioactive metabolites and opportunistic pathogens.

Results In this study, we isolated three novel *Gordonia* strains from soil samples and characterized their taxonomic status using a polyphasic taxonomic approach. Phylogenetic analysis of 16S rRNA genes and whole-genome comparisons indicated that strains CPCC 205333^T, CPCC 205515^T, and CPCC 206044^T represent three distinct novel species. The overall genome relatedness indices of average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) between these studied strains and their related type strains of the genus *Gordonia* were all below the established thresholds for species delineation, confirming the classification of these three as novel species, for which we propose the names *Gordonia altitudinis* sp. nov., *Gordonia ligustrum* sp. nov., and *Gordonia pistacia* sp. nov., respectively. Functional annotation revealed their ecological versatility, with *Gordonia* spp. contributing significantly to soil microbiome functionality through plant growth-promoting traits (e.g., nitrogen fixation, siderophore production) and biosynthetic gene clusters (BGCs), while also harboring virulence factors. Pan-genomic analysis of 225 *Gordonia* strains delineated an open gene pool ($\alpha = 0.82$; 22% fluidity), reflecting adaptive plasticity. Core genomes were enriched in conserved metabolic pathways, whereas accessory and strain-specific genes showed niche-driven functional diversification, suggesting ecological specialization.

Conclusion These findings expand the genomic and functional understanding of *Gordonia*, highlighting its dual role in environmental resilience and pathogenicity, with potential applications in biotechnology and microbiome engineering.

Keywords *Gordonia*, Rare actinomycetes, Pangenome analysis, Comparative analysis

*Correspondence:
Hua-Hong Chen
chenhh@wo.com.cn
Yu-Qin Zhang
yzhang@imb.pumc.edu.cn

¹College of Resources, Environmental Sciences and Chemistry, Chuxiong Normal University, Chuxiong 675000, Yunnan, China

²Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100050, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Rare rhizo-*Actinomycetes* are a non-*Streptomyces* group of *Actinomycetes* abundant in the rhizosphere of plants. These microorganisms have the capacity to stimulate plant growth in a number of direct and indirect ways, including mineral nutrients, regulating phytohormone levels, and biocontrol of phytopathogen [1, 2]. Several genomic analyses reveal conserved gene clusters encoding chitinases, β -1,3-glucanases, and laccases that synergistically degrade phytopathogen cell walls while releasing lignin-derived aromatic metabolites [3]. Notably, genera like *Micromonospora* and *Actinoplanes* exhibit lignocellulose-degrading enzyme systems (e.g., dye-decolorizing peroxidases, aryl-alcohol oxidases). This metabolic versatility is further evidenced by hydrocarbon degradation pathways (alkB hydroxylases) and xenobiotic detoxification networks (cytochrome P450 monooxygenases) in *Nocardia* and *Rhodococcus* [4, 5].

The genus *Gordonia*, a member of rare *Actinobacteria*, stands out for its functions of biotransformation, biodegradation, and active substance synthesis. *Gordonia* (originally '*Gordona*'), belonging to the family *Gordoniaceae*, was initially proposed by Tsukamura in 1971 [6], with *Gordonia bronchialis* as the type species. At the time of writing, there are 55 species with validly published and correct name belonging to the genus *Gordonia* (data from LPSN, <https://lpsn.dsmz.de/genus/gordonia>, accessed on May 27, 2025). Members of the genus *Gordonia* are Gram-positive, non-motile that may appear as short-rods, cocci or filamentous forms in wastewater treatment environments. They are distinguished by a chemotaxonomic profile including A1 γ -type peptidoglycan with long-chain mycolic acids, meso-diaminopimelic acid, glycolated muramic acid, and arabinogalactan cell wall polysaccharides [7]. Initially, *Gordonia* were isolated from patients with pulmonary disease [8]. Nowadays, novel species of this genus have been found in various environmental sources, such as soil [9, 10], foul water [11], wastewater [12, 13], oil-producing well [14], and plant-associated source, such as rice stem [15]. Some members of the genus *Gordonia* have attracted much attention due to their ability to degrade toxic environmental pollutants such as benzothiophene and hydrocarbons, and to produce diverse bioactive compounds [7, 13, 16–22].

In this study, three *Gordonia* strains were isolated and identified from the soil samples collected from various regions of Yunnan Province, China. Utilizing a polyphasic approach involving phenotypic and phylogenetic analyses, alongside whole-genome comparisons, we have conclusively verified that these strains indeed belong to three novel species within the genus *Gordonia*. Furthermore, we investigated the genetic basis of plant growth promotion and human infection across *Gordonia* strains

and functional diversity based on the pangenome of *Gordonia* strains.

Materials and methods

Strain isolation

Strains CPCC 205333^T (original number CS487^T), and CPCC 205515^T (original number I21A-00135^T), and CPCC 206044^T (original number A305^T) were isolated from ecologically distinct soil environments in Yunnan Province, southwestern China. Specifically, strain CPCC 205333^T was isolated from a montane soil in the Cangshan Mountain. Strain CPCC 205515^T and CPCC 206044^T was recovered from the rhizosphere soil of *Pistacia weinmanniifolia* in Honghe Hani and Yi autonomous, and *Ligustrum lucidum* in Chuxiong Yi autonomous, respectively. Soils were suspended in PBS buffer (1 g of soil was suspended in 9 mL of 1X PBS buffer), and serially diluted to a concentration of 10⁻⁴ and plated on chitin agar medium (2 g L⁻¹ chitin, 0.7 g L⁻¹ K₂HPO₄, 0.3 g L⁻¹ KH₂PO₄, 0.5 g L⁻¹ MgSO₄·7H₂O, 0.1 g L⁻¹ FeSO₄, 0.001 g L⁻¹ ZnSO₄, 0.001 g L⁻¹ MnCl₂ and 15 g L⁻¹ agar) with nystatin (50 mg L⁻¹), novobiocin (30 mg L⁻¹), and nalidixic acid (30 mg L⁻¹). After culture at 28°C for 3 weeks, bacterial colonies were purified according to the morphology of the colonies. The purified cultures were maintained as glycerol suspensions (20%, v/v) at -80°C.

Phylogenetic analysis based on 16S rRNA gene

The purified strains were incubated in ISP2 media at 28°C for 7 days for DNA extraction. The genomic DNA of newly isolated strains was extracted using a genomic DNA extraction kit (Tiangen, China). The 16S rRNA genes were amplified by PCR using forward primer 27F (5'-AGA GTT TGA TCC TGG CT-3') and reverse primer 1492R (5'-GGT TAC CTT GTT ACG ACT T-3'), according to the method previously described by Li et al. [23]. The sequences were quality-checked and assembled using Seqman program [24]. The EzBioCloud (<https://www.ezbiocloud.net/>) platform was used to analysis the 16S rRNA gene sequence similarities to closely related species [25]. Moreover, top hits were selected as closest neighbours to determine the precise taxonomic positions for these novel strains. The phylogenetic trees using neighbor-joining (NJ) [26], maximum-likelihood (ML) [27] and Maximum parsimony (MP) [28] methods were conducted on the MEGA software package (version 7.0), with a bootstrap value of 100 resamplings [29].

Genomic sequencing, assembly and annotation

Three novel *Gordonia* isolates were incubated under the same growth conditions with 16S rRNA gene sequencing. The total DNA of the were extracted using the QIAamp DNA MiniKit (Qiagen) according to the manufacturer's protocol. Whole-genome sequencing was performed

at the Beijing Genomics Institute (Beijing, China) using the Illumina HiSeq 4000 platform (Illumina, San Diego, CA, USA), producing 2×150 bp paired-end reads. After adaptors and low-quality bases were trimmed using Fastp v1.0.1 tool, the pair-end reads were assembled by SPAdes v3.10 [30] with default settings. The quality of assemblies was evaluated using CheckM2 [31]. All assemblies passed the quality control threshold (Completeness > 90% and Contamination < 5%).

Genome based identification and phylogenetic analysis

The overall genome relatedness indices (OGRIs) were used to determine pairwise genome similarity for species delineation. The average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) were calculated using the online ANI calculator [32] and Genome-to-Genome Distance Calculator (<https://ggdc.dsmz.de/ggdc.php/>), respectively. Furthermore, a phylogenomic analysis was performed to decipher the evolutionary relationships among related species. The single-copy orthologroups (SOG, one sequence per genome) were inferred by OrthoFinder v2.0.0 [33] based on the proteomes annotated by Prokka v1.1.4 [34]. For each SOG, a multiple sequence alignment was implemented using ClustalO [35], and the ambiguous sites were trimmed by trimAI [36] with the option "automated1". A maximum-likelihood species tree was constructed based on a concatenation of these alignments using IQ-tree [37] with 1000 ultrafast bootstrap replicates, and the best model was selected by ModelFinder [38].

Morphological and physiological characterization

Bacterial suspensions were adjusted to an optical density at 600 nm of 0.5, and 100 μ L of this suspension was used to inoculate for each phenotypic test. Growth on various culture media was tested by using International *Streptomyces* Project 2 (ISP2) agar, peptone yeast glucose (PYG) medium, tryptone soy agar (TSA; Difco), Luria–Bertani agar (LB; Difco), Reasoner's 2A agar (R2A; Difco), glucose yeast malt agar (GYM: 4.0 g L⁻¹ yeast extract, 10.0 g L⁻¹ malt extract, 4.0 g L⁻¹ dextrose, 2.0 g L⁻¹ calcium carbonate, 20.0 g L⁻¹ agar) and nutrient agar (NA; Difco) for up to 7 days at 28 °C. The growth temperature was tested at 4, 10, 15, 20, 25, 28, 30, 37, 40, 42, and 45 °C using TSA medium. NaCl tolerance was determined in tryptone soy broth (TSB) supplemented with 0, 1, 3, 5, 7, and 10% (w/v) NaCl. The pH tolerance was tested in TSB incubated at 28 °C. The pH was adjusted to pH 4 to 13 (1 pH unit interval incrementing) with the following buffer systems: citric acid/sodium citrate (0.1 M, pH 4.0–5.0), KH₂PO₄/NaOH (0.1 M, pH 6.0–8.0), Na₂CO₃/NaHCO₃ (0.1 M, pH 9.0–10.0), Na₂HPO₄/NaOH (0.05 M/0.1 M, pH 11.0), and KCl/NaOH (0.2 M, pH 12.0–13.0). The Gram reaction was determined by the standard Gram-stain method and

observed using light microscopy (Zeiss Axio Scope, A1 Vario). To observe cell motility, the strain was grown on ISP2 agar at 28 °C for 4 days, flooded with 0.1 M potassium phosphate buffer (pH 7) at room temperature for 10–30 min before observation with a Light microscope. Catalase activity was determined by bubble production in a solution of 3% (v/v) hydrogen peroxide and oxidase activity was tested using API oxidase reagent (bioMérieux) according to the manufacturer's instructions. Metabolic characters were tested with Biolog GEN III MicroPlate, API 50CH, and API ZYM test kits (bioMérieux) according to the manufacturer's instructions after incubation at 28 °C for 48–96 h. Other physiological tests, including H₂S production and the hydrolysis of gelatin, cellulose, and starch, were examined according to previously described procedures [39].

Chemotaxonomic analysis

Biomass for chemotaxonomic characterization of the three strains was collected from 3-day-old cultures growing in shake flasks on a rotary shaker (150 r.p.m.) using TSB medium at 28 °C. Polar lipids were extracted and examined by two-dimensional TLC and identified according to the method described by Minnikin et al. [40]. The menaquinones were extracted [41] and analyzed by High Performance Liquid Chromatography (HPLC) [42]. Cellular fatty acids were extracted and analyzed using the Sherlock Microbial Identification System (MIDI) with the MIDI Sherlock Version 6.0 and ACTIN1 database [43].

Pangenome analysis of *Gordonia* strains

To provide a comprehensive pangenome framework for assessing functional gene content variation and evolutionary trends, we conducted a genome dataset covering 303 *Gordonia* strains, which included 3 strains from this analysis and additional genomes from the NCBI assembly database. All infant-associated genome assemblies were subjected to a contamination check via CheckM2, and contaminated (contamination > 5%) and/or incomplete (completeness < 90%) genome assemblies were excluded from further analysis. Genomes with an average nucleotide identity (ANI) > 99.95% and alignment fraction > 50% were considered redundant; one representative genome from each redundant group was retained. The remaining genomes constituted a representative *Gordonia* genome dataset ($n = 225$).

Pangenome of *Gordonia* genomes based on Prokka annotation of coding sequences (general feature format) was next generated using Orthofinder with default options. Rarefaction curves and Heaps' law α values were calculated using the R package micropan [44]. Maximum-likelihood tree was constructed on a concatenation of 259 single copy alignments using IQ-TREE v2.0.5 with

ultra-fast bootstrap replicates -B 1000 and automatic evolutionary model selection option -m TEST (best fit model for the trees was determined as LG+I+G4 according to the bayesian information criterion).

The amino acid sequences of these genes were aligned against Cluster of Orthologous Groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases to obtain their corresponding annotations using eggNOG Mapper [45] with “-id 30 -subject_cov 40 -query_cov 40 -evalue 0.001”. Virulence factor (VF) annotations were conducted with DIAMOND (v.2.0.14) against the VFDB database with an e-value $< 1e^{-5}$, subject and query coverage $> 40\%$. PGPFinder pipeline was used to search and compare genomes according to their plant growth-promoting traits (PGPT) [46]. Genome mining of bio-synthetic gene clusters (BGCs) was performed using antiSMASH 7.0 [47].

Results and discussion

Isolation and identification of *Gordonia* strains

Strains CPCC 205333^T, CPCC 205515^T, and CPCC 206044^T were isolated from three soil samples in Yunnan, China. Full-length 16S rRNA gene sequences of these strains were obtained and subjected to comparative analysis using multiple alignment tools and phylogenetic tree construction methods. The highest 16S rRNA sequence identities for isolates CPCC 205333^T

and CPCC 206044^T were with the type strains of the species *G. effusa* DSM 44810^T (98.21%) and *G. rhizosphaera* JCM 10426^T (98.68%), respectively. CPCC 205515^T shared highest 16S rRNA sequence similarity with *G. mangrovi* HNM0687^T and *G. bronchialis* DSM 43247^T (both of 98.48%). The observed similarity values fall substantially below the generally accepted 98.7% cutoff criterion for species delineation [48, 49]. In the phylogenetic tree using neighbor-joining analysis of 16S rRNA gene sequences, the isolates constituted distinct monophyletic clusters compared to their phylogenetic relatives in the genus *Gordonia* (Fig. 1). Notably, congruent topological structures recovered from ML and MP trees consistently demonstrated substantial evolutionary distances from species with validly published and correct name (Fig. 1). Moreover, CPCC 205333^T and CPCC 206044^T clustered closely with *G. effusa* DSM 44810^T and *G. rhizosphaera* JCM 10426^T, respectively. In contrast, CPCC 205515^T did not cluster with *G. bronchialis* DSM 43247^T but instead grouped as a sister lineage to *G. mangrovi* HNM0687^T. Consequently, these three strains (*G. effusa* DSM 44810^T, *G. rhizosphaera* JCM 10426^T, and *G. mangrovi* HNM0687^T) were selected as reference strains and integrated genomic analyses were subsequently conducted to clarify the exact taxonomic status of three novel strains.

Comparative genome and phylogenomic analysis for three novel strains

The whole-genome sequences of isolates CPCC 205333^T, CPCC 205515^T and CPCC 206044^T have been deposited in GenBank under accession numbers JBCLVJ0000000000, JBCLVJ0000000000, JBCLVJ0000000000, respectively. The genomes sizes of these novel strains range from 4.5–6.2 Mbp with G + C content of 62.4–66.5%. Detailed genomic characteristics including coding sequences and rRNA operon counts are presented in Supplementary Table S1.

The OGRIs values, including ANI and dDDH, were calculated between the strains with their closely related type strains. The dDDH estimates between the novel strains CPCC 205333^T, CPCC 205515^T, and CPCC 206044^T and their phylogenetically neighbors were 19.3–21.9%, which was lower than 70%, the threshold for species delineation [50]. Corresponding ANI values calculated through OrthoANI (73.1–82.8%) were consistently below the 95% species boundary threshold [50].

To establish precise phylogenetic relationships of strains CPCC 205333^T, CPCC 205515^T and CPCC 206044^T, a curated dataset comprising 49 relatives of validated genomes was determined and subjected to phylogenomic analysis. The resulting phylogenomic tree exhibited strong topological congruence with 16S rRNA gene phylogeny while resolving previously ambiguous branching patterns with 100% bootstrap support at key nodes (Fig. 2). Strain CPCC 205333^T and *G. effusa* NBRC

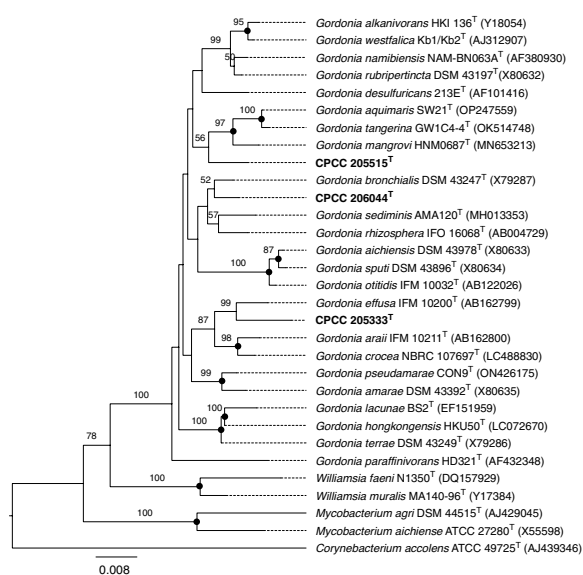


Fig. 1 Neighbour-joining phylogenetic tree based on 16S rRNA gene sequences showing the relationships between strain CPCC 205333^T, CPCC 205515^T and CPCC 206044^T with other representatives of the order *Mycobacteriales*. Numbers at nodes represent bootstrap percentages ($> 50\%$) based on 100 replicates. Filled circles indicate that the corresponding nodes were also recovered in phylogenetic trees generated using Maximum Likelihood and Maximum Parsimony methods. *Corynebacterium accolens* ATCC 49725^T (GenBank accession no. AJ439346) was used as an outgroup. Bar, 0.008 substitutions per nucleotide position

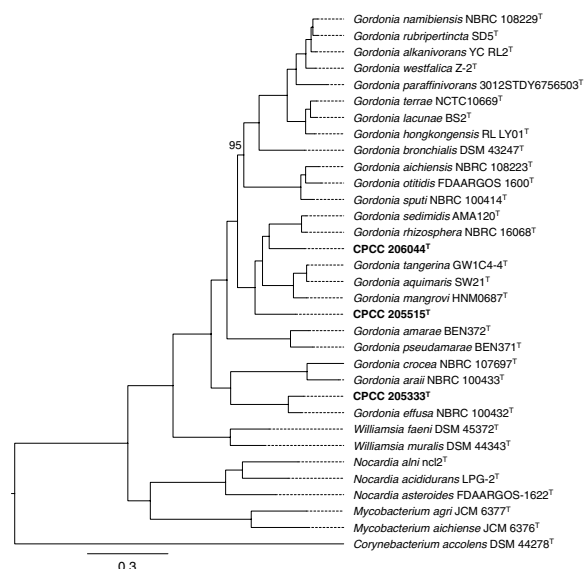


Fig. 2 The phylogenetic tree based on the genome sequences of CPCC 205333^T, CPCC 205515^T and CPCC 206044^T with other representatives of the order Mycobacteriales. The type strain of *Corynebacterium accolens* was used as an outgroup. The bold text represents novel species in this study. The best-fit substitution model was LG+G4 according to Bayesian Information Criterion (BIC). Bootstrap values that equal to 100% are not shown. The scale bar denotes 0.3 substitutions per nucleotide position

100432^T formed a clade inside the phylogenomic tree, and it was distant from the other members of *Gordonia*. In addition, CPCC 205333^T was placed independently of the other *Gordonia* species and formed a monophyletic clade with a bootstrap value of 100%. Similarly, strain CPCC 206044^T also exhibited a distinct phylogenetic placement, forming its own monophyletic clade.

Consequently, phylogenetic analysis and genomic relatedness indicated that strains CPCC 205333^T, CPCC 205515^T and CPCC 206044^T should be classified as three novel species of the genus *Gordonia*.

Phenotypic characteristics and chemotaxonomy

Based on the criterion of polyphasic taxonomy, three candidates of novel species were investigated for phenotypic characterization. Both strains displayed Gram-stain-positive attribute, lacked motility and gliding capabilities, exhibited aerobic behaviour, and featured a rod-shaped morphology. Notably, they were catalase positive and oxidase negative. The phenotypic attributes that distinguished these three strains from their most closely related species (CPCC 205333^T vs. *G. effusa* DSM 44810^T; CPCC 205515^T vs. *G. mangrovi* HNM0687^T; and CPCC 206044^T vs. *G. rhizosphaera* JCM 10426^T) are summarized in Table 1. Strain CPCC 205333^T could not utilize D-fructose, nor grow at and above 7% NaCl, pH 9.0, or 37 °C, distinguishing it from *G. effusa* DSM 44810^T. CPCC 205515^T could be differentiated from *G. mangrovi* HNM0687^T by its acid production form D-mannitol,

D-sucrose and D-trehalose, while it did not utilize α-D-Glucose, D-Mannose, D-Fructose, D-Mannitol D-Arabitol, and myo-Inositol. Strains CPCC 206044^T could grow at 4 °C and utilized D-Mannose, D-Galactose, D-Arabitol and myo-Inositol, while *G. rhizosphaera* JCM 10426^T did not. The formal proposal of these new species is in the Species Descriptions section.

The chemotaxonomic characteristics of the isolates were consistent with those typically observed in the genus *Gordonia*. All isolates exclusively contained MK-9 (H₂) as the predominant respiratory quinone. The polar Lipid profiles of strains CPCC 205333^T, CPCC 205515^T and CPCC 206044^T were dominated by diphosphatidylglycerol, phosphatidylglycerol, phosphatidylethanolamine, and phosphatidylinositol mannoside (Supplementary Fig. 1). Differentially, phosphatidylglycerol was detected in the *G. mangrovi* HNM0687^T. Moreover, the primary fatty acid constituents (>10%) identified in these strains were C_{16:0}, C_{18:0} 10-methyl, and C_{18:1}ω9c. While C_{17:0} was detected as a major fatty acid (>10%) exclusively in *G. rhizosphaera* JCM 10426^T, while it was either absent or present only in trace amounts (<2%) in the other strains. Detailed cellular fatty acid compositions (%) based on FAME analysis for these isolates, along with their closely related type strains, are presented in Supplementary Table S2.

Phylogenetic and functional genome analysis of *Gordonia* spp.

Here, a comprehensive genomic investigation was conducted on 225 *Gordonia* strains representing diverse ecological niches, including aquatic (14 strains), wastewater (28 strains), soil (75 strains), sediment (21 strains), host-associated habitats (54 strains), and others (33 strains) (Supplementary Table S3). The genome sizes of 255 strains spanned 3.19–7.33 Mb, and protein-coding gene counts ranged from 3,008 to 6,655 per strain. Notably, genome sizes differed significantly between habitats (Kruskal–Wallis test, *P* value = 1e⁻⁵), with soil-derived and general aquatic *Gordonia* strains exhibiting larger genomes than those originating from wastewater treatment systems and host-associated niches (Fig. 3A and Supplementary Figure S2). Core genome-based phylogenetic analysis of 225 *Gordonia* strains from different habitats was conducted to examine their evolutionary relationships. Type strains of validly published *Gordonia* species sharing ≥95% ANI, forming strongly supported monophyletic clades (bootstrap support = 100%) across the maximum likelihood phylogeny reconstructed from concatenated single-copy orthologues (Fig. 3A). However, beyond these species-level groupings, strains isolated from similar environments were distributed across distinct phylogenetic branches, suggesting phylogenetic

Table 1 Physiological characteristics of three novel *Gordonia* strains and their closely related type strains

Characteristics	1	2	3	4	5*	6
Growth Temperature (°C)	4–28	10–42	4–42	4–37	20–40	20–37
Growth pH	6.0–8.0	5.0–11.0	6.0–11.0	6.0–9.0	5.0–10.0	6.0–12.0
NaCl tolerance (%)	5	10	5	7	8	10
Utilization of						
α-D-Glucose	-	-	+	-	+	+
D-Mannose	-	-	+	-	+	-
D-Fructose	-	-	+	+	+	+
D-Galactose	-	-	+	-	-	-
D-Sorbitol	-	-	+	-	-	-
D-Mannitol	-	-	+	-	+	+
D-Arabitol	-	-	+	-	+	-
myo-Inositol	-	-	+	-	+	-
Activities of						
Alkaline phosphatase	+	+	+	+	ND	-
Esterase(C4)	-	+	+	-	ND	-
Lipase (C14)	+	+	+	+	ND	-
Valine arylamidase	+	+	+	+	ND	-
Cystine arylamidase	+	+	+	+	ND	-
Trypsin	+	+	-	+	ND	+
α-chymotrypsin	+	+	-	+	ND	+
α-galactosidase	+	+	-	+	ND	-
β-glucosidase	+	+	-	+	ND	+
Acid production from						
glycerol	-	+	+	-	+	-
D-glucose	-	+	+	-	+	-
D-fructose	+	+	+	+	+	-
D-mannose	-	+	+	-	+	-
inositol	-	+	+	-	+	-
D-mannitol	+	+	+	+	-	-
D-sucrose	-	+	+	-	-	-
D-trehalose	-	+	+	-	-	-
Major fatty acids (> 10%)	C _{16:0} , C _{18:0} 10-methyl, C _{18:1} ω9c, Summed Feature 4	C _{16:0} , C _{18:0} 10-methyl, C _{18:1} ω9c, Summed Feature 4	C _{16:0} , C _{18:0} 10-methyl, C _{18:1} ω9c, Summed Feature 3	C _{16:0} , C _{18:0} 10-methyl, C _{18:1} ω9c, Summed Feature 4	C _{16:0} , C _{18:0} , C _{18:0} 10-methyl, Summed Feature 3	C _{16:0} , C _{17:0} , C _{18:0} 10-methyl, C _{18:1} ω9c, Summed Feature 4

Strains: 1, CPCC 205333^T; 2, CPCC 205515^T; 3, CPCC 206044^T; 4, *Gordonia effusa* DSM 44810^T; 5, *Gordonia mangrovi* HNM0687^T; 6, *Gordonia rhizosphaera* JCM 10426^T. +, positive; -, negative activities; ND, not determined. Summed Feature 4: C_{15:0}iso 2-OH/trans-C_{16:1}ω9c, Summed Feature 3: C_{16:1}ω7c/C_{16:1}ω6c

* Data taken from Xie et al., 2020 [51]

distance between strains was not related to their ecological origin or genome size.

As an emerging dual-functional microorganism with significant roles in both soil ecosystems and clinical infections, we conduct a subsystems analysis for *Gordonia* strains, including VFs, PGPTs, and BGCs. VF profiling analysis identified an average of 77.61 ± 5.28 VFs per strain, with 100% of isolates encoding conserved modules for immune modulation, adhesion and stress survival VFs (Fig. 3B). These findings were consistent with previous case reports of clinical infections caused by *Gordonia* spp. [52–54]. In the PGPT analysis, over 40 functional

traits assigned to 7 categories associated with plant growth promotion were identified across the *Gordonia* strains (Fig. 3B). Notably, A large number of genes related to biofertilization (phosphate solubilization, nitrogen and iron acquisition), bioremediation (xenobiotics biodegradation, heavy metal detoxification), plant colonization (plant derived substrate usage and bacterial fitness), and stress response (neutralizing abiotic stress) were consistently enriched in genomes of all *Gordonia* strains. Moreover, using antiSMASH, we identified a total of 4,024 BGCs in the 225 genomes (Fig. 3B). The predominant BGC types in our dataset are annotated as: nonribosomal

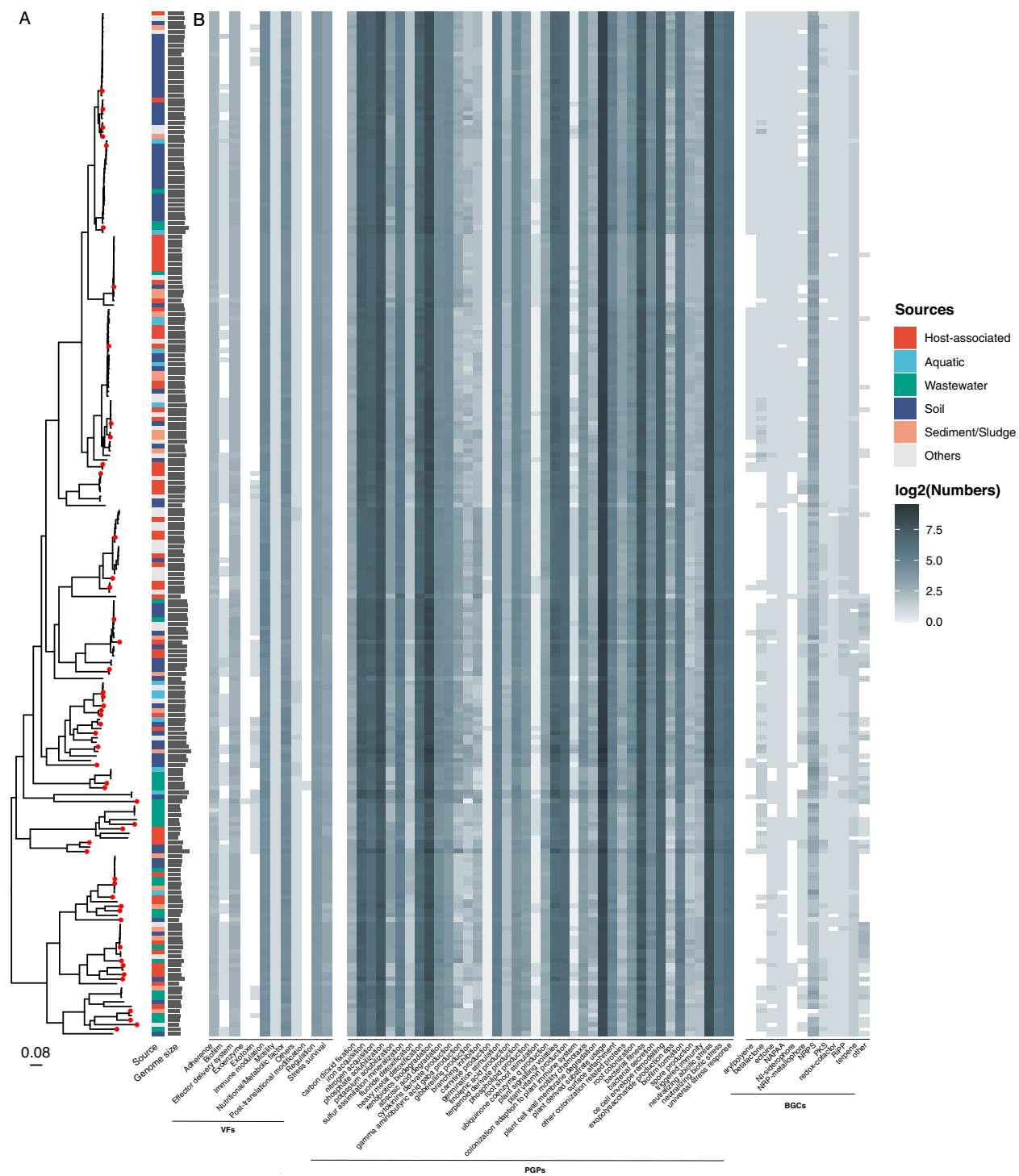


Fig. 3 Genomic features and phylogenomics of *Gordonia* strains. **A** The phylogenetic tree based on the genome sequences of 225 *Gordonia* strains. Red filled circle on the node represent type strains. Color strip and bar outside the tree are depicted by isolated sources and genomic size, respectively. Bootstrap values that are equal to 100% are not shown. The scale bar denotes 0.08 substitutions per amino acid position. **B** Gene copies or numbers associated with virulence factors, plant growth promotion traits, and biosynthetic gene clusters across *Gordonia* strains

peptide synthetase (NRPS) (1,596 BGCs), terpene (447 BGCs), polyketide synthase (PKS) (282 BGCs), ribosomally synthesized and post-translationally modified peptide (RiPP) (267 BGCs).

Pangenome analyses of *Gordonia*

The 225 *Gordonia* strains had a pangenome size of 17,170 genes. Based on this dataset, the core genome was composed of 1,393 genes (genes shared by more than 99%

strains), the accessory genome was composed of 15,547 genes, and the unique genome (strain-specific genes) was composed of 230 genes (Supplementary Fig. S3). To assess how pan-genome size and core-genome content scale with the number of *Gordonia* strains, we generated rarefaction curves for all 225 genomes and fitted them to a Heap's law model (Fig. 4). The fitted exponent ($\alpha = 0.82$) is < 1 , indicating an open pan-genome whereby each newly added genome contributes novel genes and the curve fails to reach saturation. Moreover, the pangenome fluidity value was 0.22 ± 0.09 , indicating that *Gordonia* genomes differ 22% on average. By and large, moderate openness and fluidity of the pangenome were typically associated with *Gordonia* inhabiting multiple environments and showing evidence of the exchange of genetic material [55].

To investigate the functional divergence in three different pangenome sets (Core, accessory, and unique genomes), we compared the frequency of genes assigned to KEGG pathways and COG functional categories using eggNOG-mapper, and Fisher's exact tests were conducted between each frequency of KEGG pathway and COG category within each pangenome set. As expected, COG J (translation, ribosomal structure and biogenesis) was significantly enriched in the core genome (adjusted $P < 0.05$ & Odd ratio > 1) (Fig. 5A). The core genes were also significantly assigned (p adjust < 0.05) to metabolics (F: nucleotide transport and metabolism and H: coenzyme transport and metabolism) and cellular processes and signaling (D: Cell cycle control, cell division, chromosome partitioning and O: Posttranslational modification, protein turnover, chaperones). However, with the

exception of COG S (Function unknown), no COGs were found with either frequent significant enrichment or odd ratio > 1 in the unique and accessory genomes (Supplementary Table S4).

In the similar analysis of KEGG pathways, core genome enriched various metabolism pathways (Odd Ratio > 1 and adjust P value < 0.05), including central and secondary metabolites pathway (Fig. 5B). Additionally, two translation-related pathways (ribosome and aminoacyl-tRNA biosynthesis) were also detected to significantly enrich in core genomes, consistent with the J COG previously found enriched in core genomes. In contrast, no pathways were found to be significantly enriched in a majority of accessory or unique genomes (Supplementary Table S5). Overall, this functional analysis suggests that the core genomes of *Gordonia* spp. were likely enriched for metabolic and translational functions, while non-core genes might draw from a wider variety of relatively niche functions.

Species descriptions

Description of *Gordonia altitudinis* sp. nov.

Gordonia altitudinis (al.ti.tu'di.nis. L. gen. fem. n. *altitudinis*, of a high place).

Cells are Gram-stain-positive, non-motile, aerobic and rod-shaped. Colonies on TSA are off-white, marginal irregular, opaque and rough, approximately 1 mm in diameter after growing for 3 days at 28°C and pH 7.0. Growth occurs between 4°C and 28°C with optimum growth at 28°C. The pH range for growth is pH 6.0–8.0 (optimally at pH 7.0). NaCl is not required for growth, but growth occurs in the presence of 0–5% NaCl (w/v).

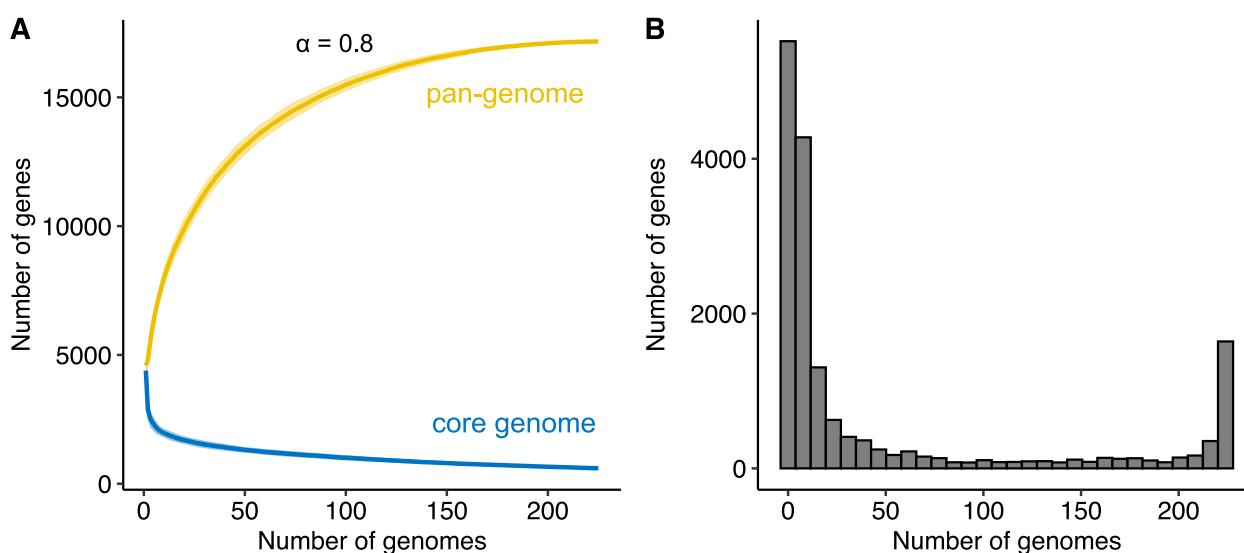


Fig. 4 The pan-genome of genus *Gordonia*. **A** Curves depicted the correlation between genome numbers and pan (blue) and core (yellow) genome sizes. The alpha exponent of Heap's Law was used to infer whether a pangenome is open or closed. Thus, if α (alpha) ≤ 1 , the pan-genome is open. In contrast, $\alpha > 1$ represents a closed pan-genome. This coefficient was computed using the *heaps* function of the micropan R package (See Methods). **B** Pangenome structure of genus *Gordonia*

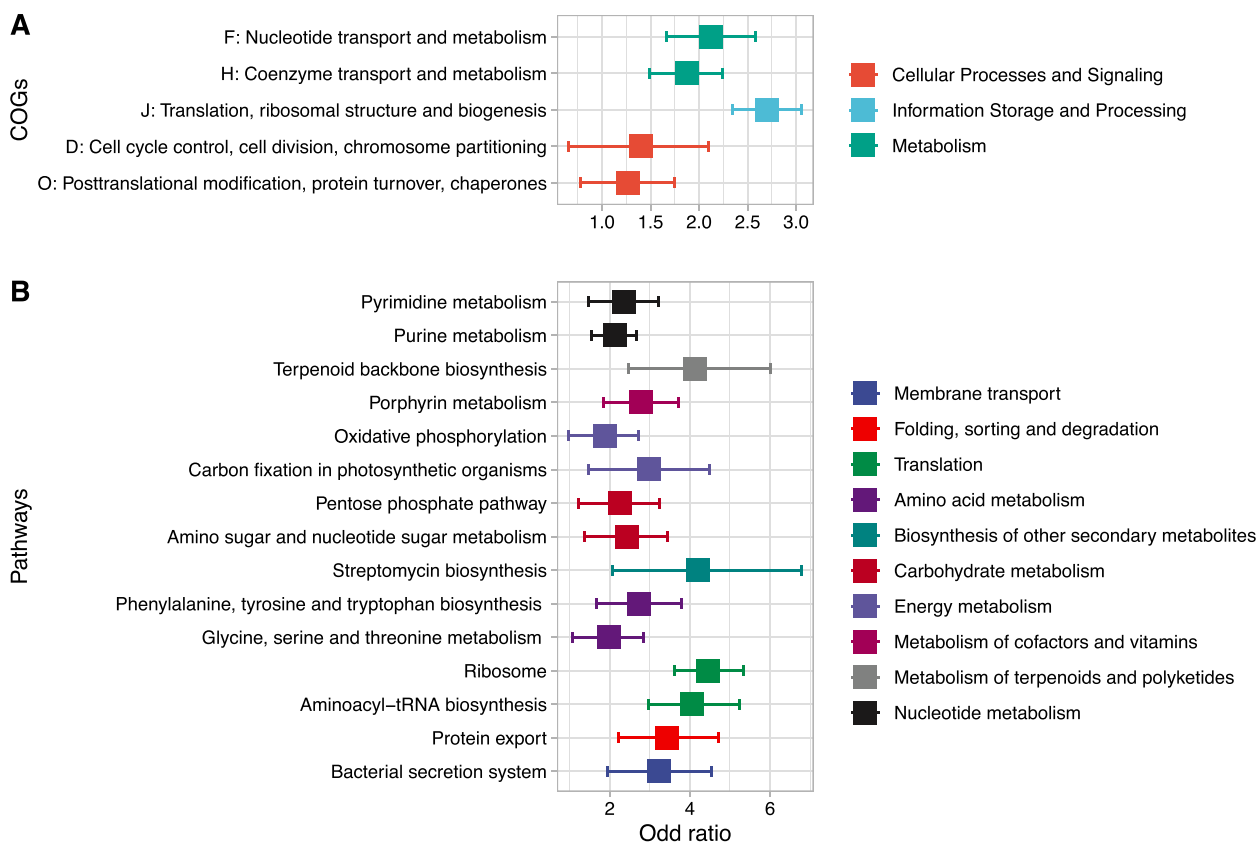


Fig. 5 Functional enrichments by COG functional category and KEGG pathway in the core and accessory genomes. The distribution of odds ratios is shown. Only COGs and KEGG pathways showing positive enrichment (Fisher's exact test, adjusted P under Bonferroni correction < 0.05) and with mean odd ratios > 1 are shown. COG "S: Function unknown" is not shown

Positive for catalase activity and hydrolysis of gelatin, but negative for oxidase reaction, H_2S production, hydrolysis of starch and nitrate reduction. Positive for alkaline phosphatase, lipase (C14), leucine aminopeptidase, valine aminopeptidase, cystine aminopeptidase, trypsin, α -chymotrypsin, α -galactosidase, β -galactosidase, β -glucuronidase, α -glucosidase, β -glucosidase, N-acetyl- β -glucosaminidase, α -mannosidase and α -fucosidase. Acetic acid, D-glucuronic acid, glucuronamide, α -ketoglutaric acid and pectin were utilized as sole source of carbon. Acid production from carbohydrates are D-fructose, D-mannitol, starch, and 2-ketogluconate. The major respiratory quinone is MK-9 (H_2). The primary polar lipid includes diphosphatidylglycerol, phosphatidylglycerol, phosphatidylethanolamine, phosphatidylinositol mannoside. The major fatty acids are $C_{16:0}$, summed feature 4 (iso- $C_{15:0}$ 2-OH/trans- $C_{16:1}\omega 9c$), $C_{18:0}$ 10-methyl, $C_{18:1}\omega 9c$, $C_{14:0}$ and $C_{19:0}$.

The type strain, CPCC 205333^T (=CS487^T=KCTC 49415^T), was isolated from a soil sample collected from Cangshan Mountain, Yunnan province, China. The genomic DNA G + C content of the type strain is 62.4% and genome size is 4.5 Mbp. The 16S rRNA gene and whole genome sequence of strain CPCC 205333^T

are publicly available through the accession numbers PP657641 and JBCLVI000000000, respectively.

Description of *Gordonia ligustrum* sp. nov.

Gordonia Ligustrum (li.gus'trum. L. gen. fem. n. *ligustrum*, of *Ligustrum*, referring to the botanical genus name of *Ligustrum lucidum*, from the rhizosphere of which the strain was isolated).

Cells are Gram-stain-positive, non-motile, aerobic and rod-shaped. Colonies on TSA are orange, convex, and opaque after growing for 3 days at 28°C and pH 7.0. Growth occurs between 10°C and 42°C with optimum growth at 15–37°C. The pH range for growth is pH 5.0–11.0 (optimally at pH 7.0–8.0). NaCl is not required for growth, but growth occurs in the presence of 0–10% NaCl (w/v). Positive for catalase activity and nitrate reduction, but negative for oxidase reaction, H_2S production, hydrolysis of starch and hydrolysis of gelatin. Positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), leucine aminopeptidase, valine aminopeptidase, cystine aminopeptidase, trypsin, α -chymotrypsin, acid phosphatase, naphthol-AS-B1-phosphohydrolase, α -galactosidase, β -galactosidase, β -glucuronidase, α -glucosidase and β -glucosidase. Acetic

acid, D-aspartic acid, N-acetyl neuraminic acid,, propionic acid can be used as the sole carbon. Acid production from carbohydrates are amygdalin, arbutin, D-arabinose, D-fructose, D-galactose, D-glucose, D-mannitol, D-mannose, D-ribose, D-sorbitol, D-trehalose, D-xylose, erythritol, glycerol, inositol, methyl α -D-glucopyranoside, N-acetylglucosamine, salicin, xylitol. The respiratory quinone is MK-9 (H_2). The cellular polar Lipid system includes diphosphatidylglycerol, phosphatidylglycerol, phosphatidylethanolamine, phosphatidylinositol, and phosphatidylinositol mannoside.. The major fatty acids are summed feature 4 (iso- $C_{15:0}$ 2-OH/trans- $C_{16:1}\omega 9c$), $C_{16:0}$, $C_{18:0}$ 10-methyl, $C_{18:1}\omega 9c$ and $C_{14:0}$.

The type strain, CPCC 205515^T (=I21A-00135^T=KCTC 59069^T), was isolated from a rhizosphere soil sample of *Ligustrum lucidum* in Chuxiong Yi autonomous prefecture, Yunnan province, China. The genomic DNA G+C content of the type strain is 66.5% and genome size is 6.2 Mbp. The 16S rRNA gene and whole genome sequence of strain CPCC 205515^T are publicly available through the accession numbers PP657644 and JBCLVJ0000000000, respectively.

Description of *Gordonia pistacia* sp. nov.

Gordonia pistacia (pis.ta'cia. L. gen. fem. n. *pistacia*, of *Pistacia*, referring to the botanical genus name of *Pistacia weinmannifolia*, from the rhizosphere of which the strain was isolated).

Cells are Gram-stain-positive, non-motile, aerobic and rod-shaped. Colonies on TSA are cream-coloured, convex, smooth and opaque after growing for 3 days at 28°C and pH 7.0. Growth occurs between 4°C and 42°C with optimum growth at 20–37°C. The pH range for growth is pH 6.0–11.0 (optimally at pH 7.0–8.0). NaCl is not required for growth, but growth occurs in the presence of 0–5% NaCl (w/v). Positive for catalase activity and nitrate reduction, but negative for oxidase reaction, H_2S production, hydrolysis of starch and hydrolysis of gelatin. Positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), leucine aminopeptidase, valine aminopeptidase, cystine aminopeptidase, acid phosphatase, naphthol-AS-B1-phosphohydrolase and α -glucosidase. D-Arabitol, D-fructose, D-galactose, D-glucose, D-mannitol, D-mannose, D-sorbitol, and myo-Inositol can be used as the sole carbon sources for energy and growth. Acid produced from 2-ketogluconate, D-fructose, and D-mannitol. The respiratory quinone is MK-9 (H_2). The cellular polar lipid system includes diphosphatidylglycerol, phosphatidylglycerol, phosphatidylethanolamine and phosphatidylinositol mannoside. The major fatty acids are $C_{16:0}$, summed feature 3 ($C_{16:1}\omega 7c/C_{16:1}\omega 6c$), $C_{18:1}\omega 9c$, $C_{18:0}$ 10-methyl and $C_{14:0}$.

The type strain, CPCC 206044^T (=A305^T=KCTC 59071^T) was isolated from a rhizosphere soil sample of

Pistacia weinmannifolia in Honghe Hani and Yi autonomous prefecture, Yunnan province, China. The genomic DNA G+C content of the type strain is 66.5% and genome size is 5.8 Mbp. The 16S rRNA gene and whole genome sequence of strain CPCC 206044^T are publicly available through the accession numbers PP657647 and JBCLVL0000000000, respectively.

Conclusions

In this work, we isolated and characterised three novel *Gordonia* strains from the soil samples. Based on physiological, biochemical, and taxogenomics characteristics, strains CPCC 205333^T, CPCC 205515^T, and CPCC 206044^T represent three novel species of the genus *Gordonia* for which the names *Gordonia altitudinis* sp. nov., *Gordonia ligustrum* sp. nov., and *Gordonia pistacia* sp. nov. were proposed. Functional genome analysis revealed that *Gordonia* spp. have genetically conserved dual functional capacities: beneficial to plant and pathogens to human. They harbored extensive plant growth-promoting (PGP) traits specializing in nutrient acquisition (nitrogen fixation, potassium solubilization, and iron uptake) and abiotic stress tolerance, while simultaneously having virulence-associated modules for host immune modulation and biofilm-mediated adhesion. Moreover, the open pangenome architecture ($\alpha = 0.82$) with moderate fluidity (0.22 ± 0.09) reflects widely ecological adaptation for *Gordonia*. Functional enrichments suggested core genomes of *Gordonia* enriched essential translational and metabolic pathways, while the accessory and unique genome provided genomic plasticity for niche-specific adaptations, as evidenced by the absence of conserved functional signatures in non-core compartments.

Competing interests

The authors declare no competing interests.

Abbreviations

NJ	Neighbor-joining
ML	Maximum-likelihood
MP	Maximum parsimony
MEGA	Molecular Evolutionary Genetics Analysis
OGRIs	Overall genome relatedness indices
ANI	Average nucleotide identity
dDDH	Digital DNA-DNA hybridization
SOG	Single-copy orthologs
ISP2	International Streptomyces Project 2 agar
PYG	Peptone yeast glucose
TSA	Tryptone soy agar
TSB	Tryptone soy broth
LB	Luria-Bertani
R2A	Reasoner's 2A agar
GYM	Glucose yeast malt agar
NA	Nutrient agar
HPLC	High Performance Liquid Chromatography
MIDI	Microbial Identification System
COG	Cluster of Orthologous Groups
KEGG	Kyoto Encyclopedia of Genes and Genomes
VF	Virulence factor
PGPT	Plant growth-promoting traits

BGCs Biosynthetic gene clusters
 NRPS Nonribosomal peptide synthetase
 PKS Polyketide synthase
 RiPP Ribosomally synthesized and post-translationally modified peptide

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12866-025-04362-0>.

Supplementary material 1.

Authors' contributions

C-JL carried out the experiments and data analysis; JZ, JS and L-YY collected the environmental samples and carried out partial data analysis; H-HC and Y-QZ conceived the research, analysed the data and prepared the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by National Natural Science Foundation of China (32170021), CAMS Innovation Fund for Medical Sciences (CIFMS, 2021-I2M-1-055) and the National Infrastructure of Microbial Resources (NIMR-2024-4).

Data availability

The 16S rRNA gene and the whole genomic sequences of strains CPCC 205333 T, CPCC 205515 T and CPCC 206044 T have been deposited in the NCBI Assembly and NCBI GenBank database, respectively. The accession numbers of 16S rRNA gene and the whole genome sequence of strain CPCC 205333 T are PP657641 and JBCLV1000000000, respectively; the accession numbers of 16S rRNA gene and the whole genome sequence of strain CPCC 205515 T are PP657644 and JBCLVJ0000000000, respectively; the accession numbers of 16S rRNA gene and the whole genome sequence of strain CPCC 206044 T are PP657647 and JBCLV1000000000, respectively.

Declarations

Ethics approval and consent to participate

This research did not contain any studies with animals performed by any of the authors.

Received: 13 May 2025 / Accepted: 26 August 2025

Published online: 28 October 2025

References

- Imade EE, Babalola OO. Biotechnological utilization: the role of zea mays rhizospheric bacteria in ecosystem sustainability. *Appl Microbiol Biotechnol*. 2021;105(11):4487–500. <https://doi.org/10.1007/s00253-021-11351-6>.
- Hartman K, Tringe SG. Interactions between plants and soil shaping the root microbiome under abiotic stress. *Biochem J*. 2019;476(19):2705–24. <https://doi.org/10.1042/BCJ20180615>.
- Olanrewaju OS, Glick BR, Babalola OO. Mechanisms of action of plant growth promoting bacteria. *World J Microbiol Biotechnol*. 2017;33(11):197. <https://doi.org/10.1007/s11274-017-2364-9>.
- Mitra D, Mondal R, Khoshru B, Senapati A, Radha TK, Mahakur B, et al. Actinobacteria-enhanced plant growth, nutrient acquisition, and crop protection: advances in soil, plant, and microbial multifactorial interactions. *Pedosphere*. 2022;32(1):149–70. [https://doi.org/10.1016/S1002-0160\(21\)60042-5](https://doi.org/10.1016/S1002-0160(21)60042-5).
- Menendez E, Carro L. Actinobacteria and Their Role as Plant Probiotics. In: Giri B, Prasad R, Wu Q-S, Varma A, editors. *Biofertilizers for Sustainable Agriculture and Environment*. Cham: Springer International Publishing; 2019. p. 333–51.
- Tsukamura M. Proposal of a new genus, *Gordonia*, for slightly acid-fast organisms occurring in sputa of patients with pulmonary disease and in soil. *J Gen Microbiol*. 1971;68(1):15–26. <https://doi.org/10.1099/00221287-68-1-15>.
- Kim SB, Brown R, Oldfield C, Gilbert SC, Goodfellow M. *Gordonia desulfuricans* sp. nov., a benzothiophene-desulphurizing actinomycete. *Int J Syst Bacteriol*. 1999;49(4):1845–51. <https://doi.org/10.1099/00207713-49-4-1845>.
- Iida S, Taniguchi H, Kageyama A, Yazawa K, Chibana H, Murata S, et al. *Gordonia otitidis* sp. nov., isolated from a patient with external otitis. *Int J Syst Evol Microbiol*. 2005;55(Pt 5):1871–6. <https://doi.org/10.1099/ijs.0.63282-0>.
- Park S, Kang SJ, Kim W, Yoon JH. *Gordonia hankookensis* sp. nov., isolated from soil. *Int J Syst Evol Microbiol*. 2009;59(Pt 12):3172–5. <https://doi.org/10.1099/ijs.0.011585-0>.
- Kim SB, Brown R, Oldfield C, Gilbert SC, Ilarionov S, Goodfellow M. *Gordonia amicalis* sp. nov., a novel dibenzothiophene-desulphurizing actinomycete. *Int J Syst Evol Microbiol*. 2000;50(6):2031–6. <https://doi.org/10.1099/00207713-50-6-2031>.
- Linos A, Berekaa MM, Steinbuechel A, Kim KK, Sproer C, Kroppenstedt RM. *Gordonia westfalica* sp. nov., a novel rubber-degrading actinomycete. *Int J Syst Evol Microbiol*. 2002;52(Pt 4):1133–9. <https://doi.org/10.1099/00207713-52-4-1133>.
- Li SH, Jin Y, Cheng J, Park DJ, Kim CJ, Hozzein WN, et al. *Gordonia jinhuensis* sp. nov., a novel actinobacterium, isolated from a VBNC (viable but non-culturable) state in pharmaceutical wastewater. *Antonie Van Leeuwenhoek*. 2014;106(2):347–56. <https://doi.org/10.1007/s10482-014-0207-3>.
- Kampfer P, Martin K, Dott W. *Gordonia phosphorivorans* sp. nov., isolated from a wastewater bioreactor with phosphorus removal. *Int J Syst Evol Microbiol*. 2013;63(Pt 1):230–5. <https://doi.org/10.1099/ijs.0.037093-0>.
- Xue Y, Sun X, Zhou P, Liu R, Liang F, Ma Y. *Gordonia paraffinivorans* sp. nov., a hydrocarbon-degrading actinomycete isolated from an oil-producing well. *Int J Syst Evol Microbiol*. 2003;53(Pt 5):1643–6. <https://doi.org/10.1099/ijs.0.02605-0>.
- Muangham S, Lipun K, Thamchaipenet A, Matsumoto A, Duangmal K. *Gordonia oryzae* sp. nov., isolated from rice plant stems (*Oryza sativa* L.). *Int J Syst Evol Microbiol*. 2019;69(6):1621–7. <https://doi.org/10.1099/ijsem.0.003368>.
- Drzyzga O, Navarro Llorens JM, Fernandez de Las Heras L, Garcia Fernandez E, Perera J. *Gordonia cholesterolivorans* sp. nov., a cholesterol-degrading actinomycete isolated from sewage sludge. *Int J Syst Evol Microbiol*. 2009;59(Pt 5):1011–5. <https://doi.org/10.1099/ijs.0.005777-0>.
- Kummer C, Schumann P, Stackebrandt E. *Gordonia alkanivorans* sp. nov., isolated from tar-contaminated soil. *Int J Syst Bacteriol*. 1999;49(4):1513–22. <https://doi.org/10.1099/00207713-49-4-1513>.
- Kim KK, Lee KC, Klenk HP, Oh HM, Lee JS. *Gordonia kroppenstedtii* sp. nov., a phenol-degrading actinomycete isolated from a polluted stream. *Int J Syst Evol Microbiol*. 2009;59(Pt 8):1992–6. <https://doi.org/10.1099/ijs.0.005322-0>.
- Yoon JH, Lee JJ, Kang SS, Takeuchi M, Shin YK, Lee ST, et al. *Gordonia nitida* sp. nov., a bacterium that degrades 3-ethylpyridine and 3-methylpyridine. *Int J Syst Evol Microbiol*. 2000;50(3):1203–10. <https://doi.org/10.1099/00207713-50-3-1203>.
- Jin D, Kong X, Jia M, Yu X, Wang X, Zhuang X, et al. *Gordonia phthalatica* sp. nov., a di-n-butyl phthalate-degrading bacterium isolated from activated sludge. *Int J Syst Evol Microbiol*. 2017;67(12):5128–33. <https://doi.org/10.1099/ijsem.0.002430>.
- Linos A, Steinbuechel A, Sproer C, Kroppenstedt RM. *Gordonia polyisoprenivorans* sp. nov., a rubber-degrading actinomycete isolated from an automobile tyre. *Int J Syst Bacteriol*. 1999;49(Pt 4):1785–91. <https://doi.org/10.1099/00207713-49-4-1785>.
- Kim KK, Lee CS, Kroppenstedt RM, Stackebrandt E, Lee ST. *Gordonia sihwensis* sp. nov., a novel nitrate-reducing bacterium isolated from a wastewater-treatment bioreactor. *Int J Syst Evol Microbiol*. 2003;53(Pt 5):1427–33. <https://doi.org/10.1099/ijs.0.02224-0>.
- Li WJ, Xu P, Schumann P, Zhang YQ, Pukall R, Xu LH, et al. *Georgenia ruanii* sp. nov., a novel actinobacterium isolated from forest soil in Yunnan (China), and emended description of the genus *Georgenia*. *Int J Syst Evol Microbiol*. 2007;57(Pt 7):1424–8. <https://doi.org/10.1099/ijs.0.64749-0>.
- Swindell SR, Plasterer TN. SEQMAN. Contig assembly. *Methods Mol Biol*. 1997;70:75–89.
- Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, et al. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol*. 2017;67(5):1613–7. <https://doi.org/10.1099/ijsem.0.001755>.
- Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4(4):406–25. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>.
- Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol*. 1981;17(6):368–76. <https://doi.org/10.1007/BF01734359>.
- Xia X. Maximum Parsimony Method in Phylogenetics. In: Xia X, editor. *Bioinformatics and the Cell: Modern Computational Approaches in Genomics*,

- Proteomics and Transcriptomics. Cham: Springer International Publishing; 2018. p. 327–41.
29. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985;39(4):783–91. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>.
 30. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455–77. <https://doi.org/10.1089/cmb.2012.0021>.
 31. Chklovskii A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods*. 2023;20(8):1203–12. <https://doi.org/10.1038/s41592-023-01940-w>.
 32. Lee I, Ouk Kim Y, Park SC, Chun J. Orthoani: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol*. 2016;66(2):1100–3. <https://doi.org/10.1099/ijsem.0.000760>.
 33. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20(1):238. <https://doi.org/10.1186/s13059-019-1832-y>.
 34. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068–9. <https://doi.org/10.1093/bioinformatics/btu153>.
 35. Sievers F, Higgins DG. Clustal omega for making accurate alignments of many protein sequences. *Protein Sci*. 2018;27(1):135–45. <https://doi.org/10.1002/pro.3290>.
 36. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009;25(15):1972–3. <https://doi.org/10.1093/bioinformatics/btp348>.
 37. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-tree 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020;37(5):1530–4. <https://doi.org/10.1093/molbev/msaa015>.
 38. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017;14(6):587–9. <https://doi.org/10.1038/nmeth.4285>.
 39. Gonzalez C, Gutierrez C, Ramirez C. Halobacterium vallismortis sp. nov. an amyolytic and carbohydrate-metabolizing, extremely halophilic bacterium. *Can J Microbiol*. 1978;24(6):710–5. <https://doi.org/10.1139/m78-119>.
 40. Minnikin DE, O'Donnell AG, Goodfellow M, Alderson G, Athalye M, Schaaf A, et al. An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *J Microbiol Methods*. 1984;2(5):233–41. [https://doi.org/10.1016/0167-7012\(84\)90018-6](https://doi.org/10.1016/0167-7012(84)90018-6).
 41. Collins MD, Pirouz T, Goodfellow M, Minnikin DE. Distribution of menaquinones in actinomycetes and corynebacteria. *J Gen Microbiol*. 1977;100(2):221–30. <https://doi.org/10.1099/00221287-100-2-221>.
 42. Groth I, Schumann P, Rainey FA, Martin K, Schuetze B, Augsten K. Demetria terrigena gen. nov., sp. nov., a new genus of actinomycetes isolated from compost soil. *Int J Syst Bacteriol*. 1997;47(4):1129–33. <https://doi.org/10.1099/00207713-47-4-1129>.
 43. Kroppenstedt RM. Fatty acid and menaquinone analysis of actinomycetes and related organisms. Society of Applied Bacteriology Technical series. 1985:173–99.
 44. Snipen L, Liland KH. micropan: an R-package for microbial pan-genomics. *BMC Bioinformatics*. 2015;16(1):79. <https://doi.org/10.1186/s12859-015-0517-0>.
 45. Cantalapiedra CP, Hernandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38(12):5825–9. <https://doi.org/10.1093/molbev/msab293>.
 46. Pellegrinetti TA, Monteiro GGTN, Lemos LN, Santos RACD, Barros AG, Mendes LW. PGPg_finder: a comprehensive and user-friendly pipeline for identifying plant growth-promoting genes in genomic and metagenomic data. *Rhizosphere*. 2024;30:100905. <https://doi.org/10.1016/j.rhisph.2024.100905>.
 47. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. AntiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res*. 2023;51(W1):W46–50. <https://doi.org/10.1093/nar/gkad344>.
 48. Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol*. 2014;64(Pt 2):346–51. <https://doi.org/10.1099/ijso.0.059774-0>.
 49. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, et al. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol*. 2018;68(1):461–6. <https://doi.org/10.1099/ijsem.0.002516>.
 50. Riesco R, Trujillo ME. Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol*. 2024;74(3). <https://doi.org/10.1099/ijsem.0.006300>.
 51. Xie Y, Zhou S, Xu Y, Wu W, Xia W, Zhang R, et al. *Gordonia mangrovi* sp. nov., a novel actinobacterium isolated from mangrove soil in Hainan. *Int J Syst Evol Microbiol*. 2020;70(8):4537–43. <https://doi.org/10.1099/ijsem.0.004310>.
 52. Johnson JA, Onderdonk AB, Cosimi LA, Yawetz S, Lasker BA, Bolcen SJ, et al. *Gordonia bronchialis* bacteremia and pleural infection: case report and review of the literature. *J Clin Microbiol*. 2011;49(4):1662–6. <https://doi.org/10.1128/JCM.02121-10>.
 53. Ding X, Yu Y, Chen M, Wang C, Kang Y, Li H, et al. Bacteremia due to *Gordonia polyisopenivorans*: case report and review of literature. *BMC Infect Dis*. 2017;17(1):419. <https://doi.org/10.1186/s12879-017-2523-5>.
 54. Ramanan P, Deziel PJ, Wengenack NL. *Gordonia* bacteremia. *J Clin Microbiol*. 2013;51(10):3443–7. <https://doi.org/10.1128/jcm.01449-13>.
 55. Reis AC, Cunha MV. The open pan-genome architecture and virulence landscape of *Mycobacterium bovis*. *Microb Genom*. 2021;7(10). <https://doi.org/10.1099/mgen.0.000664>.

Publisher's Note

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