



OPEN

Genomic insights into a novel species, *Dyella thailandensis* sp. nov., a cellulolytic and xylanolytic bacterium isolated from soil associated with leaf compost

Nahatai Intarapasi¹, Anon Thammasittirong^{1,2}, Sukanya Jeennor³, Pattaraporn Yukphan⁴ & Sutticha Na-Ranong Thammasittirong^{1,2}✉

The efficient enzymatic conversion of lignocellulose into fermentable sugars is essential for producing sustainable biofuels and biochemicals. During a screening for lignocellulose-degrading microorganisms, a novel bacterial strain, designated KULCS107^T, was isolated from soil associated with leaf compost in Nakhon Pathom, Thailand. Phylogenetic analysis based on the 16S rRNA gene sequence placed the strain within the genus *Dyella*, with *Dyella ginsengisoli* Gsoil 3046^T as its closest described relative (99.0% sequence identity). However, genome-wide analyses revealed that strain KULCS107^T represents a distinct species. The digital DNA-DNA hybridization (dDDH) and average nucleotide identity (ANI) values between KULCS107^T and *D. ginsengisoli* Gsoil 3046^T were 36.2% and 88.9%, respectively—both substantially below the established thresholds for species delineation. The complete genome of strain KULCS107^T is 3.98 Mb, with a G+C content of 67.8%, and encodes 3,549 protein-coding genes, 48 tRNA genes, and 6 rRNA genes. Cellulase and xylanase production was confirmed when the strain was grown in carboxymethyl cellulose (CMC) and xylan broth, consistent with genomic predictions revealing numerous carbohydrate-active enzymes (CAZymes). Based on this polyphasic taxonomic evidence, strain KULCS107^T represents a novel species within the genus *Dyella*, for which the name *Dyella thailandensis* sp. nov. is proposed. The type strain is KULCS107^T (=TBRC 20486^T = KCTC 18276^T = NBRC 117363^T).

Keywords *Dyella thailandensis*, Lignocellulose degradation, CAZyme, Polyphasic taxonomy, Genomics

Lignocellulosic biomass is the most globally abundant renewable carbon source, primarily composed of cellulose, hemicellulose, and lignin. Its complex and recalcitrant structure poses a major challenge to efficient bioconversion. Conventional chemical and physical methods, such as acid or alkaline hydrolysis and steam explosion, have long been employed to deconstruct lignocellulose and enhance saccharification efficiency. However, these approaches often require harsh conditions, high energy input, and may generate inhibitory by-products^{1,2}. In contrast, enzymatic hydrolysis provides a milder and environmentally friendly alternative, employing microorganisms and their enzymes to depolymerize lignocellulose into fermentable sugars, thereby supporting the sustainable production of biofuels and value-added biochemicals³. Although certain microorganisms possess the intrinsic ability to degrade native lignocellulosic materials, efficient bioconversion at laboratory or industrial scales generally requires pretreatment to improve substrate accessibility and enzymatic efficiency². Consequently, the isolation and characterization of novel lignocellulose-degrading microorganisms are of great importance. The

¹Department of Science and Bioinnovation, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand. ²Microbial Biotechnology Unit, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand. ³Industrial Bioprocess Technology Research Team, Functional Ingredients and Food Innovation Research Group (IFIG), National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani 12120, Thailand. ⁴Microbial Diversity and Utilization Research Team, Thailand Bioresource Research Center (TBRC), National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani 12120, Thailand. ✉email: sutticha.n@ku.ac.th

discovery of new species with these capabilities not only uncovers untapped genetic resources, advancing our understanding of microbial diversity in carbon cycling, but also provides promising and robust biocatalysts for future industrial applications.

The genus *Dyella* was first proposed in 2005, with *Dyella japonica* designated as the type species⁴. As of August 2025, the genus *Dyella* comprises 36 recognized species. This total, which includes 35 species with validly published names and one effectively published species (*Dyella sedimenti*) (<https://lpsn.dsmz.de/genus/dyella>), is consistent with the 36 reference genomes available in the NCBI database. Members of *Dyella* are ubiquitous in soil, having been isolated from diverse environments including forest^{5,6}, cultivated^{7,8}, and grassland soils⁹, as well as specialized habitats such as cliff soils¹⁰. However, some *Dyella* strains have been recovered from non-soil environments, including rock surfaces¹¹ and winery sediment¹². Phenotypically, species in this genus are characterized typically as Gram-stain-negative and aerobic. Cells are rod-shaped and colonies are yellow-pigmented. The major respiratory quinone is ubiquinone-8^{5,12}. Several *Dyella* strains have demonstrated diverse metabolic capabilities relevant to both environmental and applied microbiology, including tolerance to activated aluminum ions (Al³⁺) and inhibition of the fungal phytopathogen *Fusarium oxysporum* f. sp. *melonis*⁷, and degradation of pollutants such as biphenyl¹³ and triclosan¹⁴. In addition, some strains produce the food processing enzyme isoamylase¹⁵ and have been investigated as biofertilizers¹⁶. While these functions underscore their potential, the capacity for efficient plant biomass degradation among *Dyella* species remains less explored.

In the current study, a Gram-negative bacterium, designated strain KULCS107^T, was isolated from soil associated with decomposing leaf litter during a screening for lignocellulolytic microorganisms. This isolate exhibited the largest combined hydrolysis zones for both cellulase and xylanase activities, indicating its strong potential for lignocellulose degradation and warranting further investigation. Accordingly, the aims of the current study were to determine the taxonomic position of strain KULCS107^T through a polyphasic approach and to perform whole-genome sequencing and analysis, with a particular emphasis on identifying its carbohydrate-active enzymes (CAZymes), to assess both its novelty within the genus *Dyella* and its potential for biotechnological applications.

Materials and methods

Bacterial isolation and screening

A soil sample associated with leaf compost was collected from a residential garden in Kamphaeng Saen, Nakhon Pathom, Thailand (13°57'34.8" N, 100°00'07.9" E) in June 2023. The sample was enriched in carboxymethyl cellulose (CMC) broth (5 g/L CMC, 0.5 g/L MgSO₄·7 H₂O, 0.5 g/L yeast extract, 1 g/L KH₂PO₄, 1 g/L (NH₄)₂SO₄, and 1.0 g/L KCl) at 37 °C with shaking at 150 rpm for 1–2 days. Then, the enriched culture was serially diluted, spread onto CMC agar, followed by incubation at 37 °C for 1–2 days to obtain bacterial colonies. Following purification, the isolates were screened for cellulolytic and xylanolytic on CMC and xylan agar, respectively, using the Gram's iodine overlay method¹⁷. A clear hydrolysis zone, or halo, around a colony indicated positive activity. Pure cultures were preserved in 20% (w/v) glycerol at -80 °C.

16S rRNA gene and phylogenetic analysis

Genomic DNA was extracted using a MagAttract HMW DNA kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. The 16S rRNA gene was amplified using PCR based on the universal bacterial primers 27F and 1492R¹⁸, with the amplicon being purified and sequenced by Bionics Inc. (Seoul, Republic of Korea). The partial 16S rRNA sequence (1,459 bp) was deposited in GenBank under accession number PV635141. For the phylogenetic analysis, the full-length 16S rRNA gene sequence of strain KULCS107^T (1,545 bp) was extracted from the whole-genome sequence using the ContEst16S tool¹⁹. The sequence was initially compared against the EzBioCloud 16S rRNA gene database²⁰ to determine the most closely related taxa. For the phylogenetic tree construction, complete 16S rRNA gene of the type strains for all 36 recognized *Dyella* species and all 7 type strains of the genus *Frateuria* were retrieved from their respective genome assemblies available in the NCBI database. Multiple sequence alignment was performed using MUSCLE²¹. The phylogenetic trees were constructed using the MEGA11²² software based on the maximum-likelihood²³, neighbor-joining²⁴, and minimum-evolution²⁵ methods. Evolution distances were calculated using a Kimura two-parameter model²⁶ and bootstrap values were calculated based on 1,000 replications.

Whole-genome sequencing and analysis

Genomic DNA was sequenced using a hybrid approach combining PacBio Sequel II long-reads and Illumina HiSeq short-reads (Macrogen, Seoul, Republic of Korea). A *de novo* hybrid assembly was generated by first assembling the long reads into a consensus sequence using Tricycler v.0.5.0²⁷. This consensus was then polished with the Illumina short reads using Polypolish v.0.5.0²⁸. The quality and contiguity of the final genome assembly were evaluated using QUAST²⁹, while genome completeness and contamination were assessed using CheckM v1.2.3³⁰. The complete genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP)³¹ and was deposited in GenBank under accession number CP192735.

The functional potential of the genome was characterized using multiple tools. A comprehensive functional annotation was performed using the Rapid Annotation using Subsystem Technology (RAST) server³². Genes were classified further into Clusters of Orthologous Groups (COG) using eggNOG-mapper³³, while the CAZymes were predicted with the dbCAN3 meta v2.2.12³⁴. A circular map of the genome was generated using Proksee³⁵.

Phylogenomics and comparative genomics

Phylogenomic analysis was performed using the Type Strain Genome Server (TYGS)³⁶. A phylogenomic tree of strain KULCS107^T and all 36 type strains of the genus *Dyella* and all 7 type strains of the genus *Frateuria*

was constructed via the Bacterial and Viral Bioinformatics Resource Center (BV-BRC)³⁷. This tree was based on a concatenated alignment of 1,000 conserved single-copy orthologous automatically selected by the BV-BRC Phylogenetic Analysis tool, with branch support values calculated using 100 rapid bootstrap replicates in RAxML³⁸. A heatmap and dendrogram of pairwise ANI values between strain KULCS107^T and all type strains of the genera *Dyella* and *Frateriella* were constructed using the Integrated Prokaryotes Genome and Pan-genome Analysis (IPGA) web server³⁹. The average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) of related type strains (*Dyella ginsengisoli* Gsoil 3046^T, *Dyella thiooxydans* ATSB10^T, *Dyella aluminiiresistens* A6^T, *Dyella soli* JS12-10^T, and *Dyella lutea* Sa^T) were calculated using the OrthoANIu algorithm⁴⁰ and the Genome-to-Genome Distance Calculator (GGDC 3.0, formula 2)⁴¹, respectively. Orthologous gene clusters between strain KULCS107^T and its closest relatives (*D. ginsengisoli* Gsoil 3046^T, *D. thiooxydans* ATSB10^T, and *D. lutea* Sa^T) were identified using the OrthoVenn3 server⁴².

Morphological and physiological characterization

Colony morphology was examined on R2A agar after incubation at 35 °C for 2 days. The Gram reaction was determined using the non-staining KOH method⁴³ and a standard staining protocol⁴⁴. Cell morphology and Gram reaction were observed under a light microscope (CX31, Olympus; Tokyo, Japan) at ×1000 magnification. Motility was assessed using the hanging drop technique. For determination of cell size and flagellation, cells were negatively stained with 1% (w/v) uranyl acetate and examined using transmission electron microscopy (HT7700, Hitachi; Tokyo, Japan) at 80 kV. Catalase activity was tested based on bubble formation in 3% (v/v) hydrogen peroxide, while oxidase activity was determined using 1% (w/v) tetramethyl-*p*-phenylenediamine. The temperature range for growth was determined on R2A agar incubated at 10, 15, 20, 25, 30, 35, 37, 40, and 45 °C for 2 days. The pH range for growth (pH 4.0–10.0, at 0.5-unit intervals) was evaluated in R2A broth at 35 °C for 5 days using 50 mM buffer systems: sodium citrate (pH 4.0–5.5), phosphate (pH 6.0–8.0), and sodium borate (pH 8.5–10.0). Salt tolerance was tested in R2A broth supplemented with NaCl (0–5% w/v, at 0.5% intervals) under the same conditions. Optimal temperature for growth was determined based on the best growth observed on R2A agar for 2 days. The optima for pH and NaCl concentration were determined from the highest optical density in R2A broth, with all incubations lasting for 5 days.

Growth on different media, including R2A agar, Luria-Bertani (LB) agar, nutrient agar (NA), and MacConkey agar was tested and observed after 2 days incubation at 35 °C. Anaerobic growth was examined on R2A agar incubated in a GasPak anaerobic jar (BBL GasPak™, Becton, Dickinson; Sparks, MD, USA) at 35 °C for 14 days. Enzymatic and biochemical characteristics were determined using API ZYM and API 20NE test kits (bioMérieux, Marcy-l'Étoile, France) according to the manufacturer's instructions. Positive and negative results were determined based on distinct color changes or growth as specified in the reading table, following the manufacturer's interpretation guidelines, whereas a weakly positive result (w) referred to an ambiguous or faint color reaction, or weak growth, that did not fully correspond to the defined positive or negative descriptions. All tests were carried out in triplicate.

Chemotaxonomic characterization

Cells grown in R2A broth at 35 °C for 3 days were harvested for chemotaxonomic analyses of respiratory quinones, fatty acid profiles, and polar lipid compositions. The respiratory quinones were extracted according to the method of Yamamoto⁴⁵, purified using thin-layer chromatography (silica gel; diethyl ether: hexane, 3:1, v/v), and analyzed using high-performance liquid chromatography (HPLC) as described by Komagata and Suzuki⁴⁶. Cellular fatty acids were analyzed using the Sherlock Microbial Identification System⁴⁷ with identification via the TSBA6 database (version 6.2B). Total lipids were extracted from freeze-dried cells using a modification of the method by Folch et al.⁴⁸, and the total lipid content was determined gravimetrically, expressed as milligrams per gram of dry cell weight (mg/g dry cell weight). Polar lipid classes within this extract were then quantified using HPLC with a charged aerosol detector (Corona, ESA Biosciences; Chelmsford, MA, USA) following Khoomrung et al.⁴⁹. Individual lipids were identified by comparing retention times with authentic standards (Sigma-Aldrich, St. Louis, Mo, USA) and quantified using calibration curves ($R^2 > 0.95$). For one-dimensional HP-TLC confirmation, lipid extracts in chloroform: methanol (2:1, v/v) were applied to a silica gel plate (60 F₂₅₄, CAMAG; Muttentz, Switzerland). The plate was developed in a chloroform-methanol-glacial acetic acid (65:25:10, v/v/v) solvent system⁵⁰. Separated lipid classes were visualized with iodine vapor⁵¹ and identified by comparing R_f values with authentic standards. Additionally, polar lipids were extracted and identified using two-dimensional TLC according to the method of Minnikin et al.⁵².

Crude enzyme production and enzymatic assay

The strain KULCS107^T was cultured in 50 mL of basal medium¹⁷ supplemented with either 10.0 g/L CMC or 10.0 g/L xylan (pH 7.0). Cultures were incubated at 35 °C for 2 days with shaking at 150 rpm. After incubation, the culture was centrifuged (10,000 rpm, 15 min, 4 °C), and the cell-free supernatant was collected as the crude enzyme source for activity assays.

Cellulase (CMCase) and xylanase activities were quantified by measuring the release of reducing sugars using the 3,5-dinitrosalicylic acid (DNS) method. The CMCase activity assay, based on Namnuch et al.⁵³, consisted of 0.5 mL of diluted crude enzyme and 0.5 mL of 1% (w/v) CMC in 50 mM sodium citrate buffer (pH 4.8), incubated at 50 °C for 30 min. The commercial cellulase, Celluclast® 1.5 L, (Sigma-Aldrich, St. Louis, MO, USA) was used for positive control. The xylanase activity assay, modified from Bailey et al.⁵⁴, contained 0.2 mL of crude enzyme and 1.8 mL of 1% (w/v) xylan in 50 mM Tris-HCl buffer (pH 7.0), incubated at 50 °C for 5 min. The commercial xylanase, X2753, (Sigma-Aldrich, St. Louis, MO, USA) was used for positive control. For both assays, the reactions were terminated by adding 3.0 mL of DNS reagent, followed by boiling for 5 min. After cooling, the absorbance was measured at 540 nm. The amount of released reducing sugar was quantified using

a glucose standard curve for CMCase and a xylose standard curve for xylanase. Negative controls for each assay consisted of enzyme without substrate and substrate without enzyme. All enzymatic assays were performed in triplicates. One unit (U) of activity was defined as the amount of enzyme required to release 1 μmol of the respective monosaccharide (glucose or xylose) per minute under the specified assay conditions.

Results and discussion

Isolation and screening of a novel bacterial strain

The screening for lignocellulolytic microorganisms from soil associated with decomposing leaf litter, termite mound soil, and cow dung yielded 92 isolates capable of utilizing CMC. Subsequent plate assays identified 23 isolates with cellulolytic and 16 isolates with xylanolytic activity, based on the formation of hydrolysis zones on CMC and xylan agar, respectively. Of these 39 isolates, only 7 demonstrated both activities. The isolate designated KULCS107^T, which exhibited the highest combined cellulase and xylanase activity as indicated by the largest hydrolysis zones (Fig. S1), was selected for further study.

16S rRNA gene and phylogenetic analysis

The nearly complete 16S rRNA gene sequence of strain KULCS107^T was 1,459 bp in length and was deposited in GenBank under accession number PV635141. A comparative analysis using the EzBioCloud database indicated that strain KULCS107^T shared the highest 16S rRNA gene sequence identity with *Dyella ginsengisoli* Gsoil 3046^T (99.0%). Sequence identities with other related type strains were 98.6% to *Dyella thiooxydans* ATSB10^T, 98.1% to *Dyella aluminiresistens* A6^T, 98.0% to *Dyella soli* JS12-10^T, and 97.8% to *Dyella lutea* Sa^T (Table 1).

Phylogenetic analyses based on complete 16S rRNA gene sequences of all 36 recognized *Dyella* type strains and all 7 type strains of the genus *Frateriuria*, using the maximum-likelihood (Fig. 1), neighbor-joining (Fig. S2), and minimum-evolution (Fig. S3) algorithms, consistently positioned strain KULCS107^T within the genus *Dyella*. In the resulting trees, the strain formed a distinct phylogenetic lineage with its closest relative, *D. ginsengisoli* Gsoil 3046^T. Since the 16S rRNA gene sequence identity between these 2 strains exceeds the 98.65% threshold commonly used for species delineation⁵⁵, further genome-based analyses were performed to clarify the taxonomic status of strain KULCS107^T.

Genomic features, phylogenomic and genomic analysis

A hybrid genome assembly using both PacBio long reads (328x coverage) and Illumina short reads (901x coverage) resulted in a complete, circular chromosome for strain KULCS107^T, with no plasmid detected. Assessment using the CheckM v1.2.3 confirmed the high quality of the assembly, indicating 99.62% completeness with 0.57% contamination. A summary of the key genomic features and a comparison with closely related species are presented in Table 2. The genome has a size of 3,978,691 bp, a G + C content of 67.8%, and contains 3,549 coding sequences (CDS), 6 rRNA genes (2 5S, 2 16S, 2 23S) and 48 tRNA genes. The two 16S rRNA gene copies were identical and showed 100% identity with the amplified 16S rRNA gene sequence. Notably, as shown in Table 2, the genome of KULCS107^T is smaller and contains fewer CDS than its relatives, with the exception of *D. aluminiresistens* A6^T.

Genome-based taxonomic analysis using the Type Strain Genome Server (TYGS) placed strain KULCS107^T within the genus *Dyella*, forming a distinct phylogenomic lineage most closely related to *D. ginsengisoli* Gsoil 3046^T (Fig. S4). This relationship was further quantified using pairwise genomic comparisons, which yielded ANI and dDDH values of 88.9% and 36.2%, respectively, when compared with *D. ginsengisoli* Gsoil 3046^T (Table 1). Both values fall well below the established species delineation thresholds of 95% for ANI and 70% for dDDH⁵⁶, confirming that strain KULCS107^T represents a novel species within the genus *Dyella*.

The phylogenetic position of strain KULCS107^T was strongly supported by multiple analyses. A phylogenomic tree constructed from 1,000 conserved single-copy orthologous genes showed that strain KULCS107^T forms a distinct and well-supported clade with *D. ginsengisoli* Gsoil 3046^T, *D. thiooxydans* ATSB10^T, and *D. lutea* Sa^T, with 100% bootstrap support (Fig. 2). Furthermore, this placement was confirmed with a pairwise ANI analysis (Fig. 3). Comparative genomic analysis of KULCS107^T and its 3 closest relatives revealed a core genome of 2,682 orthologous gene clusters and 3 unique gene clusters in strain KULCS107^T. Strain KULCS107^T shared the highest number of orthologs with *D. ginsengisoli* Gsoil 3046^T (2,943), further supporting their close phylogenetic proximity (Fig. 4). A circular genome map comparing KULCS107^T with *D. ginsengisoli* Gsoil 3046^T (Fig. 5) also supported this close relationship, revealing a high degree of genomic synteny and BLAST-based sequence identity. Based on this polyphasic evidence, the name *Dyella thailandensis* sp. nov. is proposed, with KULCS107^T designated as the type strain.

Species	Strain	16 S rRNA gene identity (%)	ANI value (%)	dDDH value (%)
<i>D. ginsengisoli</i>	Gsoil 3046 ^T	99.0	88.9	36.2
<i>D. lutea</i>	Sa ^T	97.8	84.3	27.4
<i>D. thiooxydans</i>	ATSB10 ^T	98.6	84.1	26.9
<i>D. soli</i>	JS12-10 ^T	98.0	77.9	21.5
<i>D. aluminiresistens</i>	A6 ^T	98.1	77.7	21.5

Table 1. Comparison of 16S rRNA gene sequence identity, ANI, and dDDH values between strain KULCS107^T and its closest relatives.

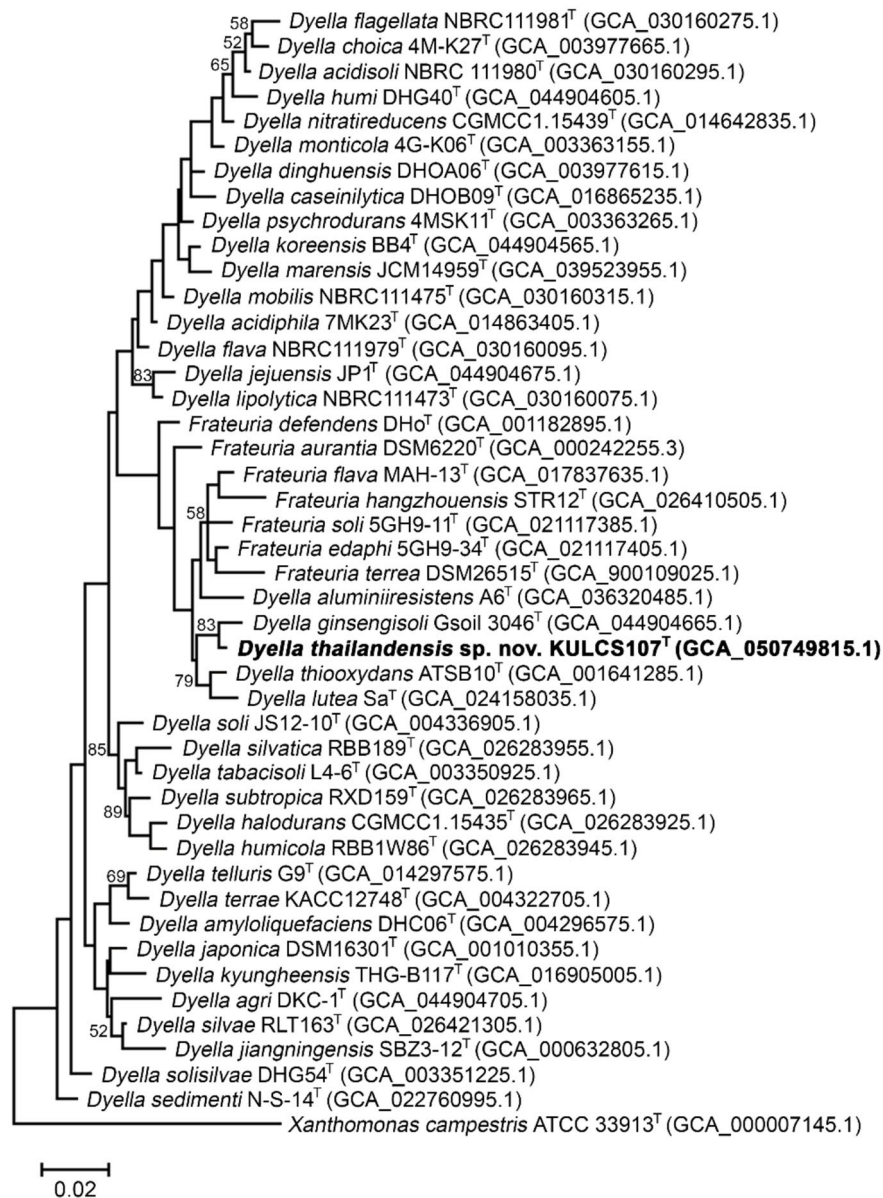


Fig. 1. Maximum-likelihood phylogenetic tree based on 16S rRNA gene sequences illustrating position of strain KULCS107^T relative to all closely related species of the genera *Dyella* and *Frateuria*, with *Xanthomonas campestris* ATCC 33913^T as the outgroup. Bootstrap values (based on 1,000 replications) greater than 50% are shown at the nodes. The analysis was based on an alignment of 1,545 nucleotide positions. GenBank accession numbers are in parentheses. Bar, 0.02 substitutions per nucleotide site.

The functional capabilities of strain KULCS107^T and its closely related species, *D. ginsengisoli* Gsoil 3046^T, *D. lutea* Sa^T, and *D. thiooxydans* ATSB10^T, were annotated and classified using the RAST subsystem (Table S1) and COG databases (Table S2). Overall, the total number of genes assigned to functional categories was slightly lower in KULCS107^T, consistent with its smaller genome size. RAST identified 1,341 features assigned to subsystems, and COG classified 3,431 genes, both fewer than in the related genomes. All four *Dyella* strains exhibited a well-conserved core metabolic framework, with high proportions of genes associated with essential processes such as protein metabolism, amino acid and derivatives and carbohydrate (RAST), as well as transcription, amino acid transport and metabolism, and cell wall, membrane, envelope biogenesis (COG). Notably, the RAST annotation revealed marked variation in the motility and chemotaxis subsystem. Strain KULCS107^T encoded 62 genes in this category, a number comparable to *D. lutea* Sa^T (63) but nearly threefold greater than those in *D. ginsengisoli* Gsoil 3046^T (23) and *D. thiooxydans* ATSB10^T (18). This suggests that KULCS107^T may have a more complex chemotactic and motility system, potentially enhancing its ability to sense and respond to environmental gradients. While sharing a core metabolic framework, the genomic analysis of KULCS107^T reveals a distinct functional profile that, combined with its phylogenetic and genomic distinctiveness, supports its recognition as a novel species within the genus *Dyella*.

Genomic feature	1	2	3	4	5	6
Genome size (Mb)	3.98	4.26	4.40	4.23	8.19	3.72
Contig	1	19	8	1	19	1
GC content (%)	67.8	67.5	67.9	67.9	65.0	65.0
CDS	3,549	3,822	3,987	3,782	7,218	3,239
rRNA genes	6	5	6	6	13	6
tRNA genes	48	46	50	49	96	50

Table 2. Comparison of genomic features between strain KULCS107^T and its closest relatives. Strains: 1, KULCS107^T; 2, *D. ginsengisoli* Gsoil 3046^T; 3, *D. lutea* Sa^T; 4, *D. thiooxydans* ATSB10^T; 5, *D. soli* JS12-10^T; 6, *D. aluminiiresistens* A6^T. Data for strains 2–6 were retrieved from GenBank.

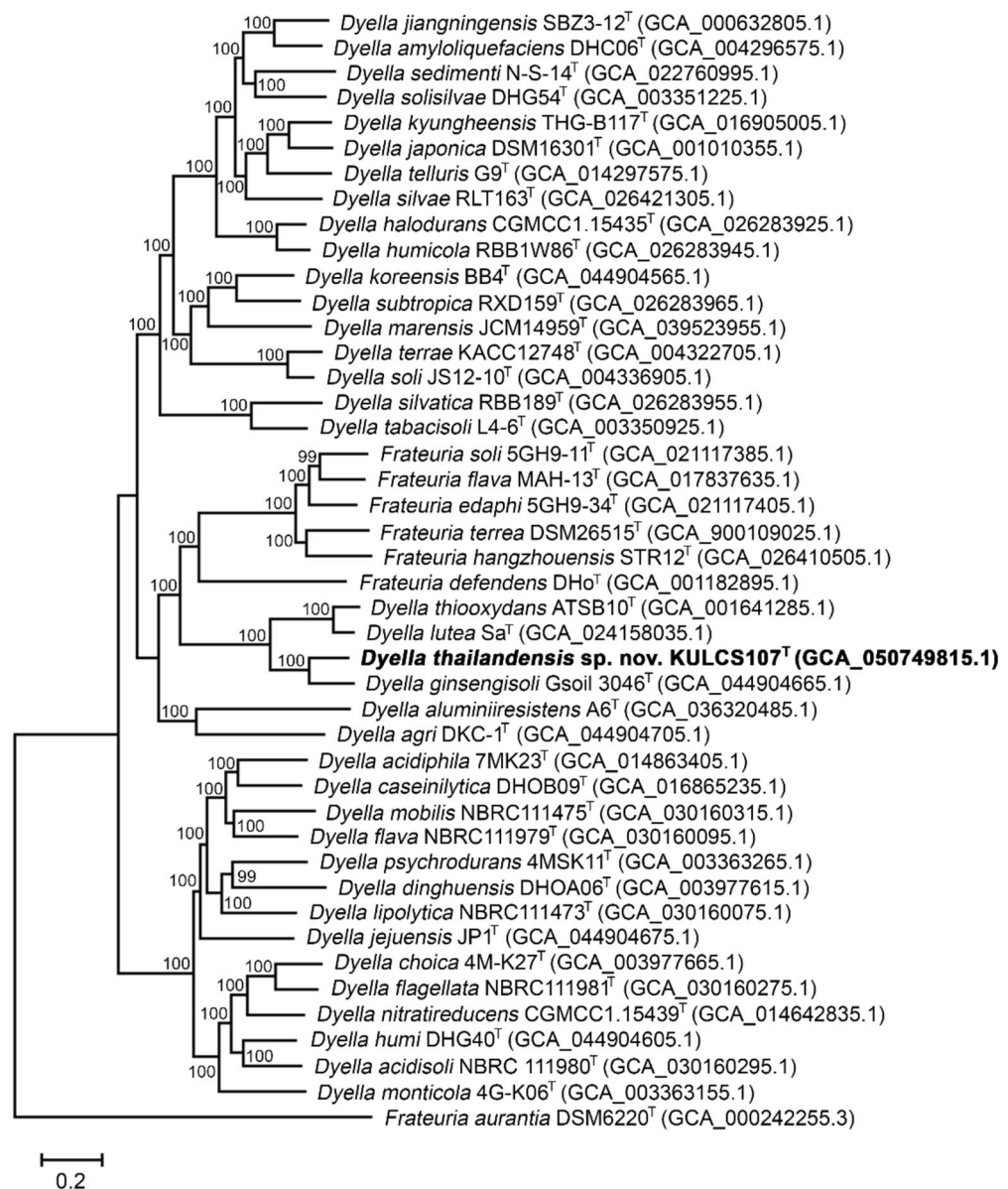


Fig. 2. Phylogenomic tree constructed using 1,000 conserved single-copy orthologous genes, illustrating the phylogenetic relationship of strain KULCS107^T with all type strains of the genera *Dyella* and *Frateuria*. Bootstrap support values based on 100 replications are shown at the nodes. GenBank accession numbers are given in parentheses. Bar, 0.2 substitutions per nucleotide site.

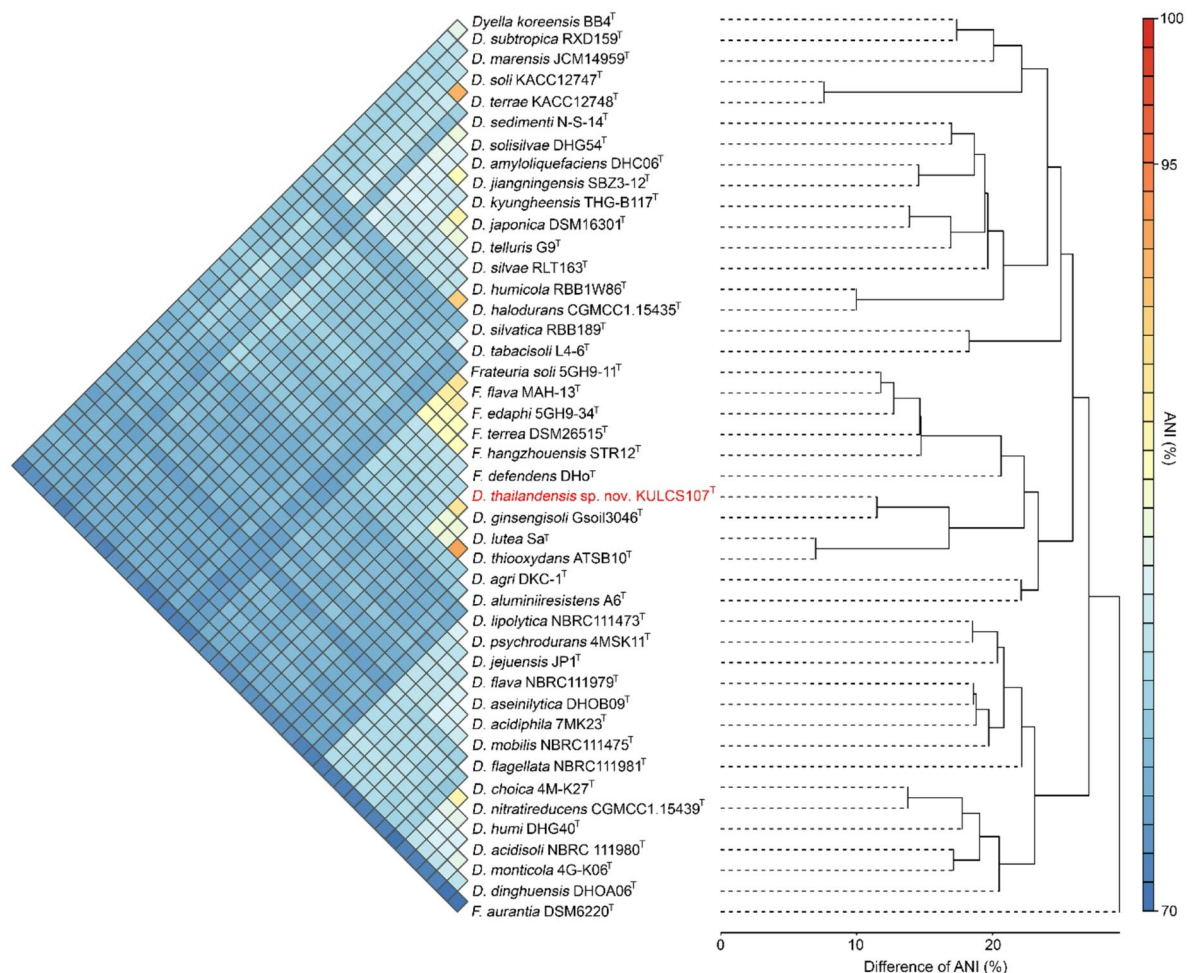


Fig. 3. Heatmap and dendrogram of pairwise ANI values between strain KULCS107^T and all type strains of the genera *Dyella* and *Frateruia*. Color gradient represents ANI values from 70% (blue) to 100% (red).

CAZymes are fundamental to the biovalorization of plant biomass, facilitating the degradation of complex lignocellulosic polymers such as cellulose, hemicellulose, and pectin into soluble sugars for biotechnological applications^{57,58}. The genome of strain KULCS107^T encodes 179 putative CAZymes, a repertoire dominated by glycoside hydrolases (GHs; $n = 68$) and glycosyltransferases (GTs; $n = 74$), along with carbohydrate-binding modules (CBMs; $n = 18$), carbohydrate esterases (CEs; $n = 11$), auxiliary activities (AAs; $n = 7$), and a single polysaccharide lyase (PL) (Fig. 6a and Table S3). This repertoire is consistent with its isolation from decomposing plant material and indicates a substantial lignocellulolytic potential. Although the total number of CAZyme genes in strain KULCS107^T is slightly lower than in the related species *D. ginsengisoli* Gsoil 3046^T (190), *D. lutea* Sa^T (215), and *D. thiooxydans* ATSB10^T (218) (Fig. 6a), it remains notably higher than in several *Dyella* species isolated from the same ecological niche (subtropical forest soils), including *D. humicola* RBB1W86^T (119), *D. subtropica* RXD159^T (100), *D. silvatica* RBB189^T (134), and *D. silvae* RLT163^T (96)⁶. In addition, *Dyella jiangningensis* FCAV SCS01, recovered from the metagenome of a lignocellulose-degrading microbial consortium, contained only 36 GHs in its genome⁵⁹. Other studies have emphasized the importance of such repertoires in other lignocellulose-degrading bacteria. For example, *Halosquirtibacter xylanolyticus* sp. nov. DS1-an-2312^T, capable of hydrolyzing xylan into xylotriose and xylotetraose, harbors 128 GHs, 37 GTs, 27 PLs, 35 CEs, and 22 CBMs⁶⁰. *Bacillus amyloliquefaciens*, has broad CAZyme repertoires to ferment lignocellulosic biomass, such as chrysanthemum stems, where GT and PL families are key contributors⁵⁸. Taken together, the robust CAZyme repertoire targeting lignocellulosic substrates of strain KULCS107^T (Fig. 6b) positions it as a promising candidate for the deconstruction of complex plant biomass. These findings not only reinforce its ecological role in natural biomass turnover but also highlight its potential value in biomass valorization and industrial biotechnology.

Physiological characterization

Colonies of strain KULCS107^T were yellow-pigmented, circular, transparent, and smooth with entire margins after 2 days of incubation on R2A agar at 35 °C. Cells were Gram-negative, aerobic, rod-shaped (0.5–0.8 × 1.5–2.0 μm), and motile, with a single polar flagellum (Fig. 7). The strain was positive for both catalase and oxidase

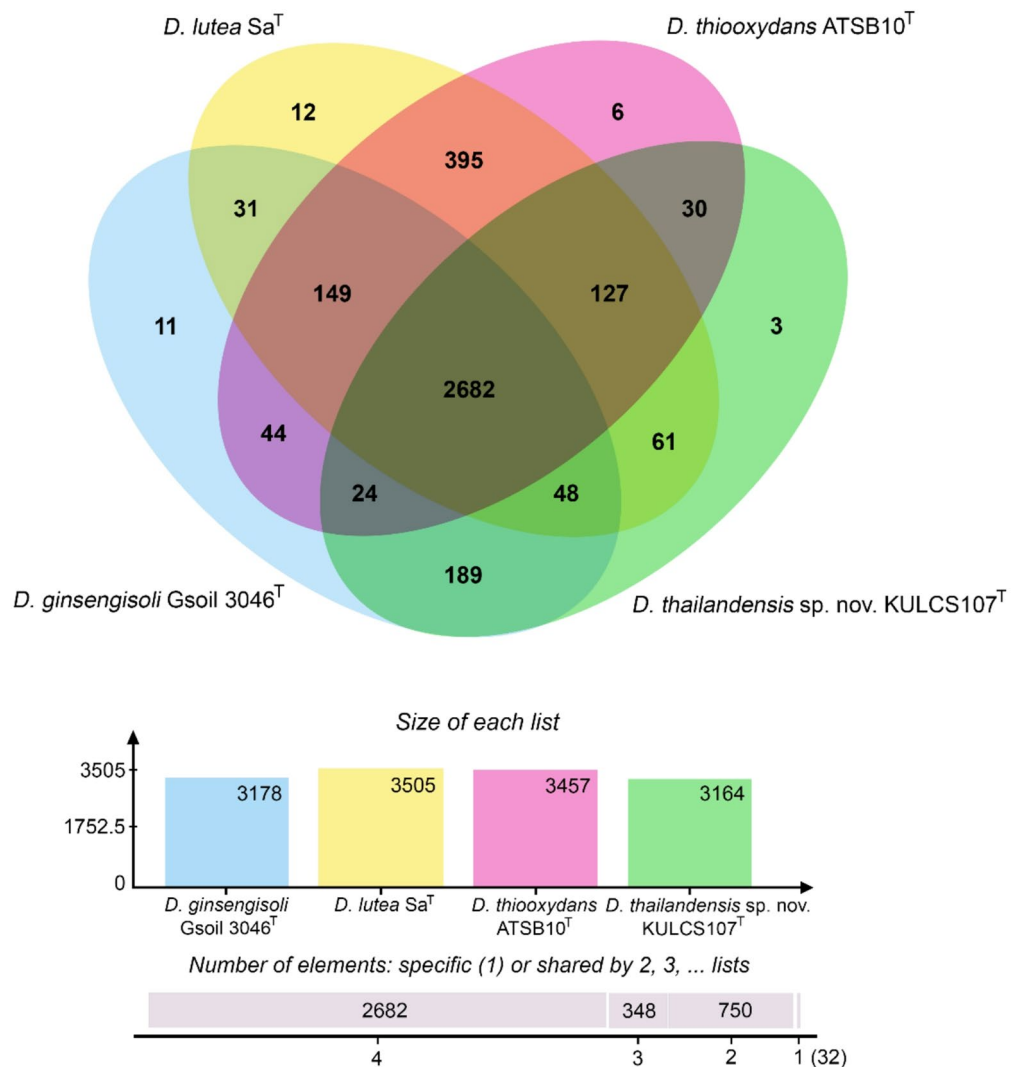


Fig. 4. Pangenome analysis of strain KULCS107^T and its closest relatives. Venn diagram illustrates distribution of shared and unique orthologous gene clusters among strain KULCS107^T (green) and the type strains of *D. ginsengisoli* Gsoil 3046^T (blue), *D. lutea* SA^T (yellow), and *D. thiooxydans* ATSB10^T (pink). Bar graph indicates total number of proteins for each genome.

activities. Growth occurred at temperatures in the range 25–40 °C and within a pH range of 4.5–10.0. The strain tolerated NaCl concentrations up to 4.5% (w/v). A summary of detailed phenotypic characteristics is provided in Table 3. Collectively, these properties distinguish strain KULCS107^T from related *Dyella* species.

Chemotaxonomic characterization

The cellular fatty acid composition of strain KULCS107^T is shown in Table 4. The overall fatty acid profile of strain KULCS107^T was consistent with its classification within the genus *Dyella*, being dominated by branched-chain fatty acids. The major fatty acids of strain KULCS107^T were anteiso-C_{15:0} (18.7%), iso-C_{15:0} (13.8%), and Summed Feature 9 (13.3%). Compared to its closest relative, *D. ginsengisoli* Gsoil 3046^T, strain KULCS107^T contained considerably higher proportions of anteiso-C_{15:0} (18.7% vs. 15.4%) and iso-C_{15:0} (13.8% vs. 6.8%), but lower proportions of iso-C_{16:0} (8.1% vs. 15.4%) and Summed Feature 9 (13.3% vs. 16.7%). Furthermore, strain KULCS107^T produced a small amount of cyclo C_{17:0} (0.7%), which was not detected in *D. ginsengisoli* Gsoil 3046^T. These differences provide strong chemotaxonomic support for classifying strain KULCS107^T as a novel species. The major respiratory quinone in strain KULCS107^T was ubiquinone-8, which is consistent with other members of the genus *Dyella*^{7,61,62}. The major polar lipids were phosphatidylethanolamine (PE), Phosphatidylglycerol (PG), and diphosphatidylglycerol (DPG) (Fig. S5), a profile also characteristic of the genus *Dyella*^{6,7,12}. The unidentified aminolipid and unidentified aminophospholipids were also present. Notably, phosphatidylserine (PS) was detected in minor amounts (Fig. S6, Table S4), which, to our knowledge, is the first report of this lipid in the genus.

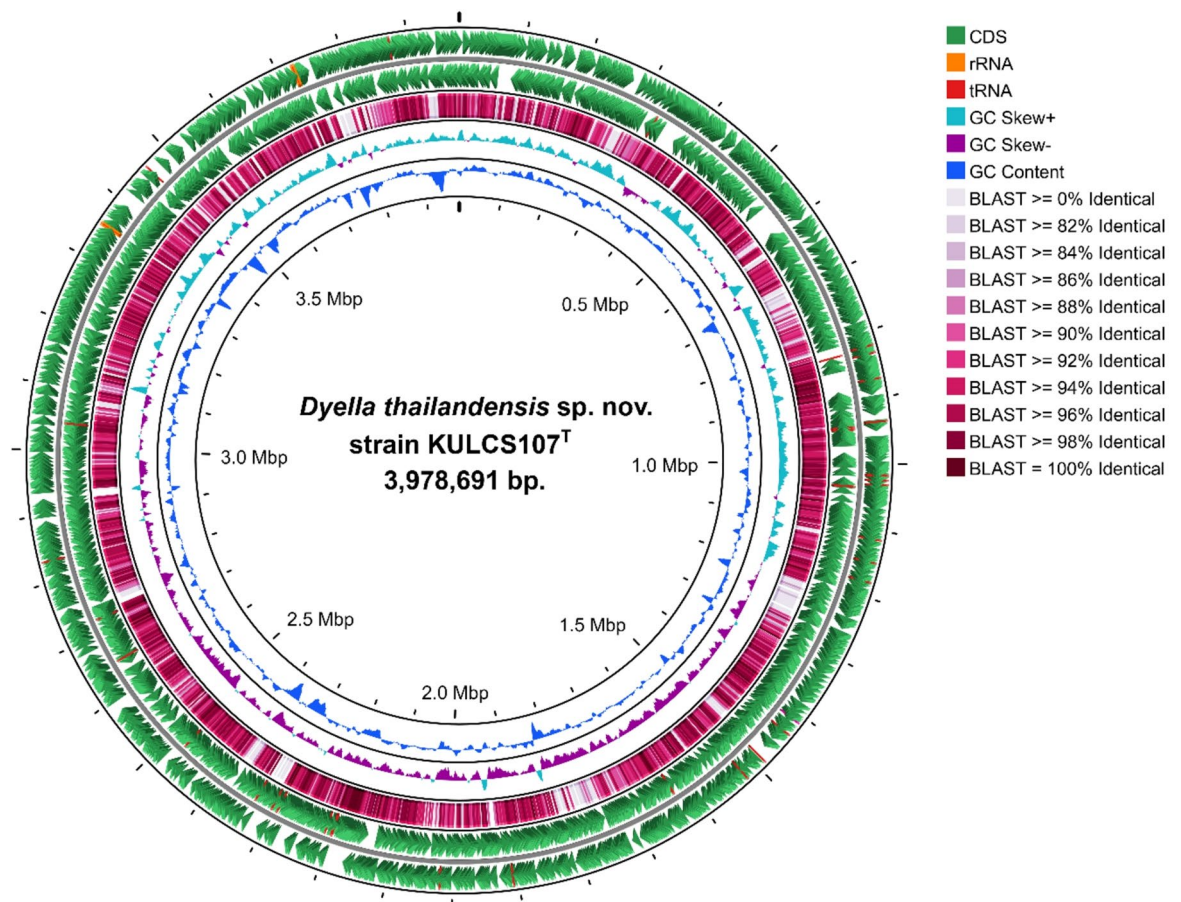


Fig. 5. Circular genome map of strain KULCS107^T. From outer to inner rings: protein-coding genes on forward and reverse strands (green); BLAST comparison of the CDS from KULCS107^T genome with *D. ginsengisoli* Gsoil 3046^T (color gradient from pale pink to dark red indicating level of sequence identity); GC skew (positive in cyan, negative in purple); and GC content (blue).

Lignocellulolytic enzyme activity of strain KULCS107^T

The strain KULCS107^T demonstrated potent lignocellulolytic activity, with xylan induction yielding maximum xylanase and cellulase activities of 3,600 mU/mL and 180 mU/mL, respectively, at 24 h of cultivation (Fig. 8). A subsequent decline in both enzyme activities at 48 h likely resulted from nutrient depletion, catabolite repression, and enzyme degradation and deactivation^{17,63}. The enzymatic profile of KULCS107^T was substantiated by the genomic analysis, which revealed a corresponding repertoire of CAZymes providing the genetic basis for this function (Fig. 6). The lignocellulolytic activity of KULCS107^T distinguishes it from its closest phylogenetic neighbors. In comparison, *D. lutea* Sa^T and *Dyella thiooxydans* ATSB10^T have no reported xylanase activity and were negative for CMC hydrolysis^{61,64}, while *D. ginsengisoli* Gsoil 3046^T was inactive on both substrates⁸. Interestingly, these related species harbor a greater number of predicted CAZyme genes than KULCS107^T (Fig. 6). This discrepancy suggests that the magnitude of lignocellulolytic function in *Dyella* species is not solely determined by CAZyme gene abundance. Instead, it highlights the importance of gene expression, regulation, or differences in enzyme-specific activity. Other *Dyella* strains show variable or condition-dependent activity. For instance, *Dyella* sp. BM6 formed a hydrolysis zone on CMC agar, but its CMCase activity was not measurable in a liquid basal medium⁶⁵. While *Dyella* sp. SSA-1562^T exhibited high CMCase (6.52 U/mL) and xylanase (~4 U/mL) activities, those results were obtained via solid-state fermentation on sawdust⁶⁶, limiting a direct comparison with this study.

Notably, the enzyme yields for KULCS107^T were achieved under unoptimized batch-culture conditions. This indicates that there is considerable potential for yield enhancement through the optimization of cultivation parameters, such as medium composition, pH, and temperature. Collectively, these findings underscore the potential of KULCS107^T for biotechnological applications in plant biomass deconstruction. Future studies evaluating the strain's performance on raw, unprocessed lignocellulosic biomass are essential to fully assess its bioconversion capabilities.

Description of *Dyella thailandensis* sp. nov.

Dyella thailandensis (thai.land.en'sis. N.L. fem. adj. *thailandensis* pertaining to Thailand, where the type strain was isolated). Cells are Gram-negative, aerobic, and rod-shaped (0.5–0.8 × 1.5–2.0 μm), and are motile with a

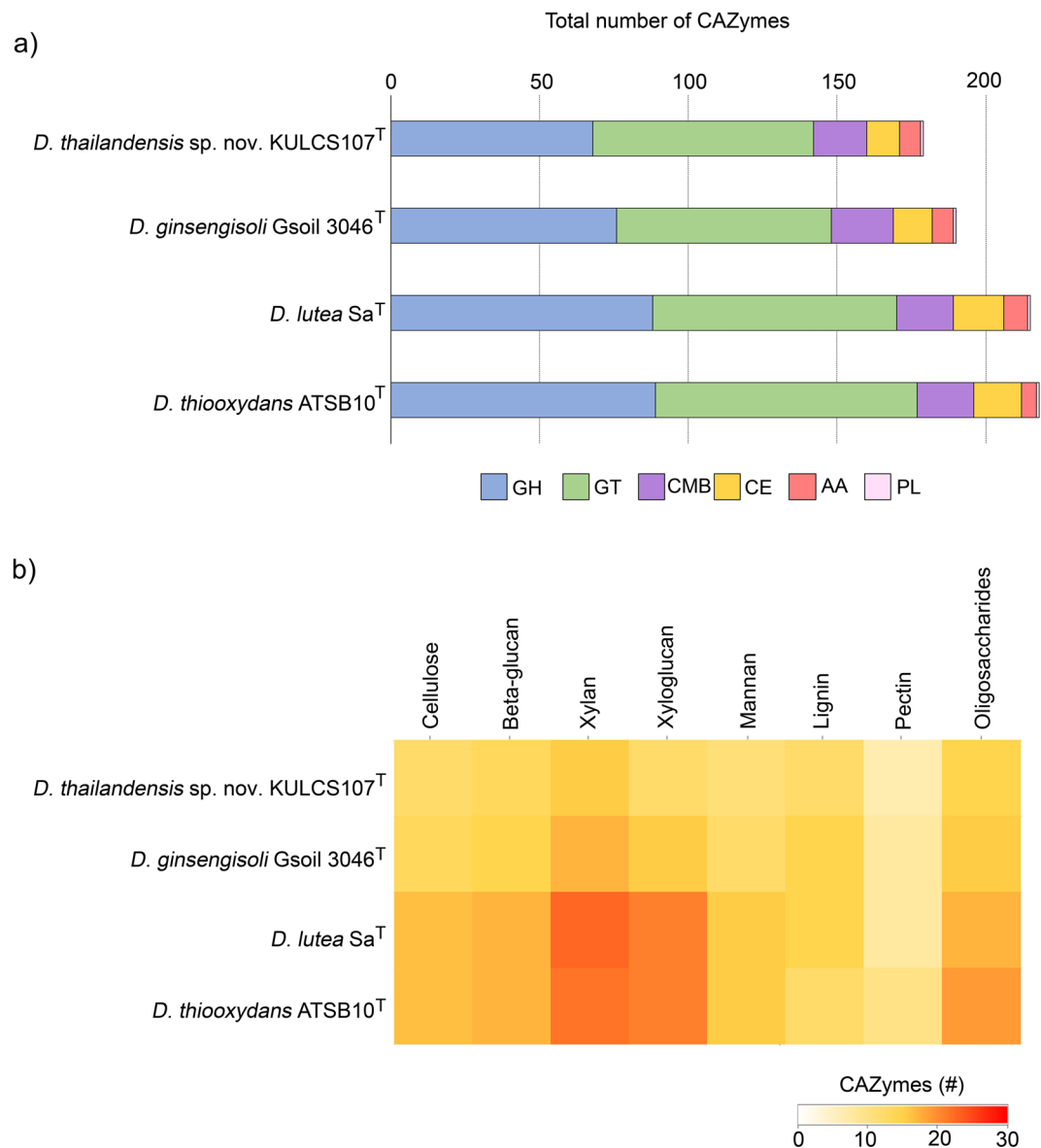


Fig. 6. Comparative analysis of CAZyme genes in strain KULCS107^T and the type strains *D. ginsengisoli* Gsoil 3046^T, *D. lutea* SA^T, and *D. thiooxydans* ATSB10^T. **(a)** Total number of CAZyme genes in each species. **(b)** Heatmap of CAZyme genes abundance involved in lignocellulose degradation (color gradient from light yellow to dark orange indicating gene abundance). Abbreviations: GH, glycoside hydrolases; GT, glycosyltransferases; CBM, carbohydrate-binding module; CE, carbohydrate esterase; AA, auxiliary activities; PL, polysaccharide lyase.

single polar flagellum. Growth was observed on R2A agar at 25–40 °C (optimum, 35 °C). Growth occurred at pH 4.5–10.0 (optimum, pH 8), and in the presence of up to 4.5% (w/v) NaCl in R2A broth. The strain grows well on R2A and NA, and weakly on LB agar, but not on MacConkey agar. Catalase and oxidase are positive. In API 20NE tests, strain KULCS107^T hydrolyzes aesculin and shows weak gelatinase activity. Negative results are obtained for nitrate reduction, indole production, glucose acidification, arginine dihydrolase, urease, and β-galactosidase activities. The strain assimilates D-glucose, N-acetyl-glucosamine, and D-maltose, but does not utilize L-arabinose, D-mannitol, D-mannose, potassium gluconate, caprate, adipate, malate, citrate, and phenylacetate. In API ZYM tests, the strain KULCS107^T exhibits positive reactions for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-glucosidase, β-glucosidase, and N-acetyl-β-glucosaminidase. Weak enzymatic activities are observed for lipase (C14), cystine arylamidase, and α-chymotrypsin. Negative enzymatic activities include trypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-mannosidase, and α-fucosidase. The major fatty acids are anteiso-C_{15:0} (18.7%), iso-C_{15:0} (13.8%), and Summed Feature 9 (13.3%). The major respiratory quinone is ubiquinone-8. The major polar lipids of strain KULCS107^T are phosphatidylethanolamine,

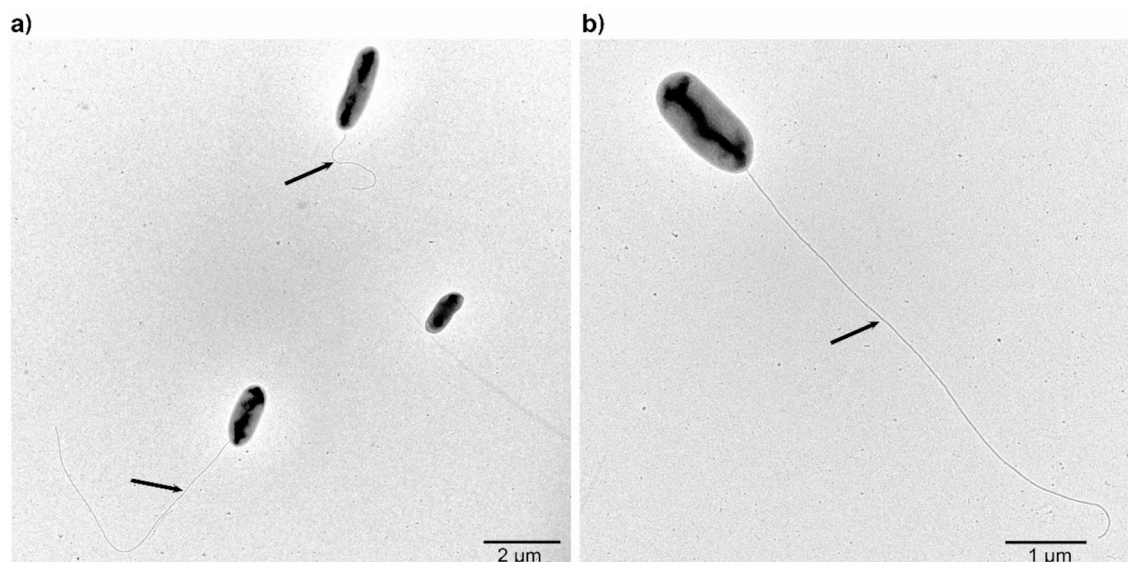


Fig. 7. Transmission electron micrograph of cells of strain KULCS107^T. (a) Cell at ×2,000 magnification. (b) Cell at ×4,000 magnification, where arrow indicates single polar flagellum.

Characteristic	1	2	3	4	5	6
Catalase/oxidase	+/+	+/+	+/+	-/+	+/+	-/-
Temperature range (°C)	25–40	20–40	20–40	25–40	20–35	15–37
pH range	4.5–10.0	4.5–8.0	5.0–7.5	5.0–8.0	5.0–10.0	4.5–8.0
NaCl tolerance (%)	0–4.5	0–5.0	0–4.5	0–5.0	0–2.0	0–0.5
Biochemical characteristic (API 20NE)						
Nitrate reduction	-	-	-	-	-	+
Aesculin hydrolysis	+	+	+	+	-	+
Gelatinase	w	-	-	-	-	-
β-Galactosidase	-	-	-	-	-	+
Assimilation						
D-Mannose	-	-	-	w	-	-
N-Acetyl-glucosamine	+	+	w	+	w	-
D-Maltose	+	+	+	+	w	-
Enzyme activity (API ZYM)						
Alkaline phosphatase	+	+	w	+	w	-
Lipase (C14)	w	w	w	w	-	-
Cystine arylamidase	w	+	w	+	-	-
α-Chymotrypsin	w	w	w	+	-	-
Naphthol-AS-BI-phosphohydrolase	+	+	w	+	w	+
β-Galactosidase	-	-	-	-	-	+
α-Glucosidase	+	+	+	+	w	-
β-Glucosidase	+	w	+	+	-	+
N-Acetyl-β-glucosaminidase	+	+	w	+	w	-

Table 3. Physiological characteristics of strain KULCS107^T compared with related *Dyella* species. Strains: 1, KULCS107^T; 2, *D. ginsengisoli* Gsoil 3046^T; 3, *D. lutea* Sa^T; 4, *D. thiooxydans* ATSB10^T; 5, *D. soli* JS12-10^T; 6, *D. aluminiiresistens* A6^T. Data for strains 1–5 obtained in the current study; data for strain 6 from Li et al.⁷. +, positive; w, weakly positive; –, negative.

phosphatidylglycerol, and diphosphatidylglycerol. Phosphatidylserine is present as a minor polar lipid. The G + C content of the genomic DNA of the type strain is 67.8% based on the complete genome sequence.

Fatty acid	1	2	3	4	5	6
Iso-C _{11:0}	6.1	4.1	3.9	3.8	3.5	3.2
Iso-C _{11:0} 3-OH	6.4	4.5	4.6	5.0	4.3	4.3
Iso-C _{13:0} 3-OH	4.1	2.3	2.4	2.5	2.4	2.9
Iso-C _{14:0}	1.1	-	-	1.0	-	-
Iso-C _{15:0}	13.8	6.8	13.5	14.3	14.9	7.1
Anteiso-C _{15:0}	18.7	15.4	11.41	8.7	5.8	4.5
C _{16:0}	5.5	8.5	4.1	2.7	3.3	15.3
Iso-C _{16:0}	8.1	15.4	14.4	18.5	15.1	6.6
Iso-C _{17:0}	10.0	10.2	16.3	13.8	18.3	23.6
Anteiso C _{17:0}	1.7	3.2	2.3	1.7	3.5	6.4
Cyclo _{17:0}	0.7	-	-	-	-	3.3
Summed Feature 3	5.3	7.4	3.8	4.1	5.4	1.2
Summed Feature 9	13.3	16.7	17.9	18.0	16.6	14.3

Table 4. Cellular fatty acid composition (%) of strain KULCS107^T compared to related *Dyella* species. Major components ($\geq 10.0\%$) are shown in bold. The symbol “-” indicates fatty acids not detected or $< 1\%$ of the total fatty acid content, except cyclo 17:0, which is reported for diagnostic significance. Strains: 1, KULCS107^T; 2, *D. ginsengisoli* Gsoil 3046^T; 3, *D. lutea* Sa^T; 4, *D. thiooxydans* ATSB10^T; 5, *D. soli* JS12-10^T; 6, *D. aluminii*resistens A6^T. summed feature 3 comprises C16:1 ω 7c and/or C16:1 ω 6c; summed feature 9 comprises iso-C17:1 ω 9c. Data for strains 1–5 were obtained in this study; data for strain 6 were from Li et al.⁷

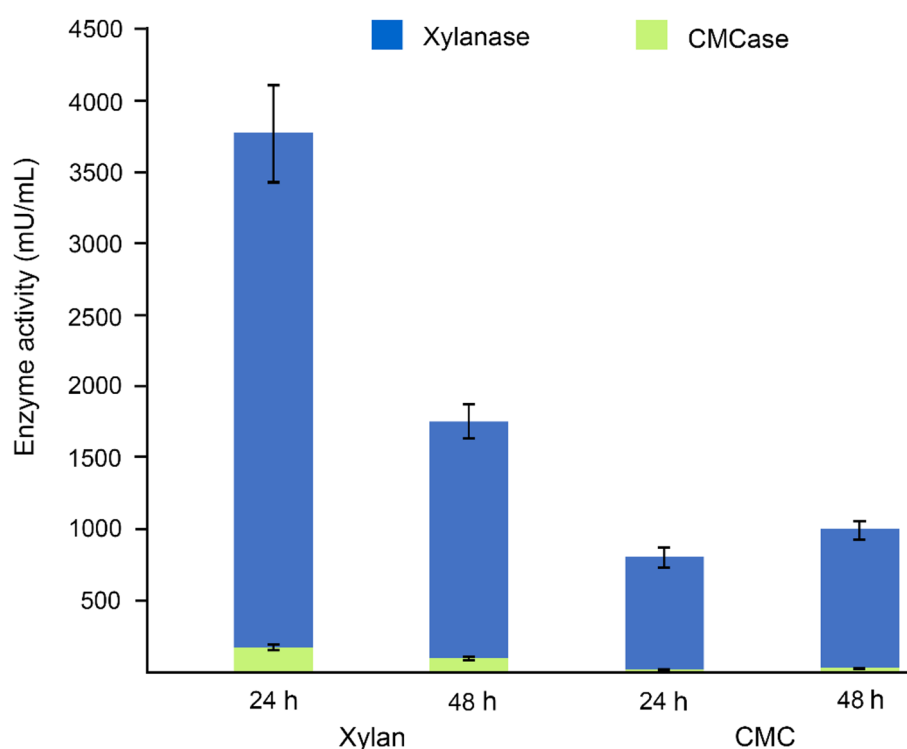


Fig. 8. Xylanase and CMCase activities in supernatant of strain KULCS107^T after 24 and 48 h of cultivation in broth containing either xylan or CMC as sole carbon source.

Conclusion

This study successfully characterized a novel, lignocellulolytic bacterium, strain KULCS107^T, isolated from soil associated with leaf compost in Thailand. A polyphasic taxonomic approach, combining phylogenetic analysis of the 16S rRNA gene with genome-wide comparisons using ANI and dDDH values, demonstrated that this strain represents a novel species within the genus *Dyella*. Accordingly, the name *Dyella thailandensis* sp. nov. is proposed. The complete genome analysis revealed a relatively small genome size (3.98 Mb) with a high G + C content and an abundance of genes encoding CAZyme, consistent with its demonstrated cellulase and xylanase activities. This discovery not only contributes to the known microbial diversity but also identifies *D. thailandensis*

KULCS107^T as a promising candidate for biotechnological applications, particularly in the biorefinery sector for the efficient conversion of lignocellulosic biomass into biofuels and other high-value biochemicals. Future studies should focus on the detailed characterization of these specific enzymes to fully harness their industrial potential.

Data availability

The GenBank accession numbers for the 16S rRNA gene and whole-genome sequences are PV635141 and CP192735, respectively. The SRA accession number for the PacBio (long-read) sequencing data is SRR35922960, and the Illumina (short-read) sequencing data is SRR35922961.

Received: 6 October 2025; Accepted: 22 December 2025

Published online: 09 January 2026

References

- Sharma, S. & Kumar, S. K. Exploration of thermophilic lignocellulolytic enzymes producing bacterial strains from hot springs of Western Himalayan range. *Indian J. Microbiol.* <https://doi.org/10.1007/s12088-024-01385-5> (2024).
- Wang, K., Shen, Z., Wang, X., Li, Z. & Cheng, S. Advances in enhancing the enzymatic saccharification process of lignocellulosic biomasses for bioethanol production. *Biomass Bioenergy*. **191**, 107450. <https://doi.org/10.1016/j.biombioe.2024.107450> (2024).
- Ren, F. et al. Fungal systems for lignocellulose deconstruction: from enzymatic mechanisms to hydrolysis optimization. *GCB Bioenergy*. **16**, e13130. <https://doi.org/10.1111/gcbb.13130> (2024).
- Xie, C. H. & Yokota, A. *Dyella japonica* gen. nov., sp. nov., a gamma-proteobacterium isolated from soil. *Int J Syst. Evol. Microbiol.* **55**, 753–756. <https://doi.org/10.1099/ijs.0.63377-0> (2005).
- Chen, M. H., Lv, Y. Y., Wang, J., Tang, L. & Qiu, L. H. *Dyella Humi* sp. nov., isolated from forest soil. *Int J Syst. Evol. Microbiol.* **66**, 4372–4376. <https://doi.org/10.1099/ijsem.0.001358> (2016).
- Feng, G. D. et al. *Dyella humicola* sp. nov., *Dyella subtropica* sp. nov., *Dyella silvatica* sp. nov. and *Dyella silvae* sp. nov., isolated from subtropical forest soil. *Int. J. Syst. Evol. Microbiol.* **73**, 005878. <https://doi.org/10.1099/ijsem.0.005878> (2023).
- Li, Y. et al. *Dyella aluminirensistens* sp. nov., a Al³⁺-tolerant bacterium with the ability to inhibit *Fusarium oxysporum* f. sp. melonis isolated from rhizosphere soil of muskmelon. *Int. J. Syst. Evol. Microbiol.* **75**, 006611. <https://doi.org/10.1099/ijsem.0.006611> (2025).
- Jung, H. M. et al. *Dyella ginsengisoli* sp. nov., isolated from soil of a ginseng field in South Korea. *Int J Syst. Evol. Microbiol.* **59**, 460–465. <https://doi.org/10.1099/ijs.0.64514-0> (2009).
- Chaudhary, D. K. & Kim, J. *Dyella Agri* sp. nov., isolated from reclaimed grassland soil. *Int J Syst. Evol. Microbiol.* **67**, 4246–4252. <https://doi.org/10.1099/ijsem.0.002291> (2017).
- Lee, D. W. & Lee, S. D. *Dyella marensis* sp. nov., isolated from cliff soil. *Int J Syst. Evol. Microbiol.* **59**, 1397–1400. <https://doi.org/10.1099/ijs.0.000968-0> (2009).
- Zhao, F. et al. *Dyella jiangningensis* sp. nov., a γ-proteobacterium isolated from the surface of potassium-bearing rock. *Int J Syst. Evol. Microbiol.* **63**, 3154–3157. <https://doi.org/10.1099/ijs.0.048470-0> (2013).
- Gao, L. et al. *Dyella sedimenti* sp. nov., isolated from the sediment of a winery. *Curr. Microbiol.* **79**, 316. <https://doi.org/10.1007/s00284-022-03007-z> (2022).
- Li, A., Qu, Y., Zhou, J. & Gou, M. Isolation and characteristics of a novel biphenyl-degrading bacterial strain, *Dyella ginsengisoli* LA-4. *J. Environ. Sci.* **21**, 211–217. [https://doi.org/10.1016/S1001-0742\(08\)62253-6](https://doi.org/10.1016/S1001-0742(08)62253-6) (2009).
- Wang, S., Yin, Y. & Wang, J. Microbial degradation of trichosan by a novel strain of *Dyella* Sp. *Appl. Microbiol. Biotechnol.* **102**, 1997–2006. <https://doi.org/10.1007/s00253-018-8740-z> (2018).
- Silano, V. et al. Safety evaluation of the food enzyme isoamylase from a *Dyella* sp. strain. *EFSA J.* **18**, e06250. <https://doi.org/10.2903/j.efsa.2020.6250> (2020).
- Sembiring, M. & Sabrina, T. Effectiveness of *Dyella japonica* and *Enterobacter cloacae* as biofertilizers to increase maize (*Zea mays*) production on andisol soil. *Biodiversitas* **23**, 3338–3343. <https://doi.org/10.13057/biodiv/d230704> (2022).
- Ketsakhon, P., Thammasittirong, A. & Thammasittirong, S. N. R. Adding value to rice straw waste for high-level xylanase production using a new isolate of *Bacillus altitudinis* RS3025. *Folia Microbiol. (Praha)*. **68**, 87–99. <https://doi.org/10.1007/s12223-022-00998-x> (2023).
- Khamberk, S., Thammasittirong, S. N. R. & Thammasittirong, A. Valorization of sugarcane bagasse for co-production of poly (3-hydroxybutyrate) and bacteriocin using *Bacillus cereus* strain S356. *Polymers* **16** (2015). <https://doi.org/10.3390/polym16142015> (2024).
- Lee, I. et al. ContEst16S: an algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *Int J Syst. Evol. Microbiol.* **67**, 2053–2057. <https://doi.org/10.1099/ijsem.0.001872> (2017).
- Yoon, S. H. et al. Introducing ezbiocloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst. Evol. Microbiol.* **67**, 1613–1617. <https://doi.org/10.1099/ijsem.0.001755> (2017).
- Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* **5**, 113. <https://doi.org/10.1186/1471-2105-5-113> (2004).
- Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027. <https://doi.org/10.1093/molbev/msab120> (2021).
- Thorne, J. L., Kishino, H. & Felsenstein, J. An evolutionary model for maximum likelihood alignment of DNA sequences. *J. Mol. Evol.* **33**, 114–124. <https://doi.org/10.1007/bf02193625> (1991).
- Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454> (1987).
- Fitch, W. M. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Biol.* **20**, 406–416. <https://doi.org/10.1093/sysbio/20.4.406> (1971).
- Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120. <https://doi.org/10.1007/BF01731581> (1980).
- Wick, R. R. et al. Trycycler: consensus long-read assemblies for bacterial genomes. *Genome Biol.* **22**. <https://doi.org/10.1186/s13059-021-02483-z> (2021).
- Wick, R. R. & Holt, K. E. Polypolish: short-read Polishing of long-read bacterial genome assemblies. *PLoS Comput. Biol.* **18**, e1009802. <https://doi.org/10.1371/journal.pcbi.1009802> (2022).
- Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D. & Gurevich, A. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* **34**, i142–i150. <https://doi.org/10.1093/bioinformatics/bty266> (2018).
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055. <https://doi.org/10.7287/peerj.preprints.554v2> (2015).

31. Tatusova, T. et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **44**, 6614–6624. <https://doi.org/10.1093/nar/gkw569> (2016).
32. Aziz, R. K. et al. The RAST server: rapid annotations using subsystems technology. *BMC Genom.* **9**, 75. <https://doi.org/10.1186/1471-2164-9-75> (2008).
33. Cantalapiedra, C. P. et al. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829. <https://doi.org/10.1093/molbev/msab293> (2021).
34. Zheng, J. et al. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res.* **51**, W115–W121. <https://doi.org/10.1093/nar/gkad328> (2023).
35. Grant, J. R. et al. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res.* **51**, W484–W492. <https://doi.org/10.1093/nar/gkad326> (2023).
36. Meier-Kolthoff, J. P. & Göker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **10**, 2182. <https://doi.org/10.1038/s41467-019-10210-3> (2019).
37. Olson, R. D. et al. Introducing the bacterial and viral bioinformatics resource center (BV-BRC): a resource combining PATRIC, IRD and viPR. *Nucleic Acids Res.* **51**, D678–D689. <https://doi.org/10.1093/nar/gkac1003> (2022).
38. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033> (2014).
39. Liu, D. et al. IPGA: a handy integrated prokaryotes genome and pan-genome analysis web service. *iMeta* **1**, e55. <https://doi.org/10.1002/imt2.55> (2022).
40. Chalita, M. et al. EzBioCloud: a genome-driven database and platform for Microbiome identification and discovery. *Int. J. Syst. Evol. Microbiol.* **74**, 006421. <https://doi.org/10.1099/ijsem.0.006421> (2024).
41. Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L. & Göker, M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* **50**, D801–D807. <https://doi.org/10.1093/nar/gkab902> (2021).
42. Xu, L. et al. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* **47**, W52–W58. <https://doi.org/10.1093/nar/gkz333> (2019).
43. Buck, J. D. Nonstaining (KOH) method for determination of gram reactions of marine bacteria. *Appl. Environ. Microbiol.* **44**, 992–993. <https://doi.org/10.1128/aem.44.4.992-993.1982> (1982).
44. Smith, A. C. & Hussey, M. A. Gram stain protocols. *Am. Soc. Microbiol.* <https://asm.org/protocols/gram-stain-protocols> (2019).
45. Yamamoto, K., Murakami, R. & Takamura, Y. Isoprenoid quinone, cellular fatty acid composition and diaminopimelic acid isomers of newly classified thermophilic anaerobic gram-positive bacteria. *FEMS Microbiol. Lett.* **161**, 351–358. <https://doi.org/10.1111/j.1574-6968.1998.tb12968.x> (1998).
46. Komagata, K. & Suzuki, K. I. in *Methods in Microbiology*, vol. 19, 161–207 (eds Colwell, R. R. & Grigorova, R.) (Academic Press, 1988).
47. Sasser, M. *Identification of Bacteria by Gas Chromatography of Cellular Fatty Acids, MIDI Technical Note 101* (MIDI Inc, 1990).
48. Folch, J., Lees, M. & Stanley, G. H. A simple method for the isolation and purification of total lipides from animal tissues. *J. Biol. Chem.* **226**, 497–509. [https://doi.org/10.1016/S0021-9258\(18\)64849-5](https://doi.org/10.1016/S0021-9258(18)64849-5) (1957).
49. Khoomrung, S. et al. Rapid quantification of yeast lipid using microwave-assisted total lipid extraction and HPLC-CAD. *Anal. Chem.* **85**, 4912–4919. <https://doi.org/10.1021/ac3032405> (2013).
50. Sandoval-Calderón, M., Geiger, O., Guan, Z., Barona-Gómez, F. & Sohlenkamp, C. A eukaryote-like cardiolipin synthase is present in *Streptomyces coelicolor* and in most actinobacteria. *J. Biol. Chem.* **284**, 17383–17390. <https://doi.org/10.1074/jbc.M109.006072> (2009).
51. Šajbidor, J., Čertík, M. & Grego, J. Lipid analysis of baker's yeast. *J. Chromatogr. A*. **665**, 191–195. [https://doi.org/10.1016/0021-9673\(94\)87048-9](https://doi.org/10.1016/0021-9673(94)87048-9) (1994).
52. Minnikin, D. E. et al. An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *J. Microbiol. Methods*. **2**, 233–241. [https://doi.org/10.1016/0167-7012\(84\)90018-6](https://doi.org/10.1016/0167-7012(84)90018-6) (1984).
53. Namnuch, N., Thammasittirong, A. & Thammasittirong, S. N. R. Lignocellulose hydrolytic enzymes production by *Aspergillus flavus* KUB2 using submerged fermentation of sugarcane bagasse waste. *Mycology* **12**, 119–127. <https://doi.org/10.1080/21501203.2020.1806938> (2021).
54. Bailey, M. J., Biely, P. & Poutanen, K. Interlaboratory testing of methods for assay of xylanase activity. *J. Biotechnol.* **23**, 257–270. [https://doi.org/10.1016/0168-1656\(92\)90074-J](https://doi.org/10.1016/0168-1656(92)90074-J) (1992).
55. Kim, M., Oh, H. S., Park, S. C. & 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.* **64**, 346–351. <https://doi.org/10.1099/ijss.0.059774-0> (2014).
56. Goris, J. et al. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst. Evol. Microbiol.* **57**, 81–91. <https://doi.org/10.1099/ijss.0.64483-0> (2007).
57. Liew, K. J. et al. Thermophiles and carbohydrate-active enzymes (CAZymes) in biofilm microbial consortia that decompose lignocellulosic plant litters at high temperatures. *Sci. Rep.* **12**, 2850. <https://doi.org/10.1038/s41598-022-06943-9> (2022).
58. Ren, P. et al. Decoding critical cazyne genes and transcription factors for pathogen-suppressing lignocellulosic biomass valorization via fermentation. *Bioresour Technol.* **437**, 133102. <https://doi.org/10.1016/j.biortech.2025.133102> (2025).
59. Desiderato, J. G., Alvarenga, D. O., Constancio, M. T. L., Alves, L. M. C. & Varani, A. M. The genome sequence of *Dyella jiangningensis* FCAV SCS01 from a lignocellulose-decomposing microbial consortium metagenome reveals potential for biotechnological applications. *Genet. Mol. Biol.* **41**, 507–513. <https://doi.org/10.1590/1678-4685-GMB-2017-0155> (2018).
60. Nguyen, T. T. H., Vuong, T. Q., Han, H. L. & Kim, S. G. *Halosquirtibacter laminarini* gen. nov., sp. nov. and *Halosquirtibacter Xylanolyticus* sp. nov., marine anaerobic laminarin and xylan degraders in the phylum Bacteroidota. *Sci. Rep.* **14**, 24329. <https://doi.org/10.1038/s41598-024-74787-6> (2024).
61. Park, S. et al. *Runella salmonicolor* sp. nov. and *Dyella lutea* sp. nov., isolated from paddy field soil. *Int J Syst. Evol. Microbiol.* **73**, 005938. <https://doi.org/10.1099/ijsem.0.005938> (2023).
62. Weon, H. Y. et al. *Dyella soli* sp. nov. and *Dyella terrae* sp. nov., isolated from soil. *Int J Syst. Evol. Microbiol.* **59**, 1685–1690. <https://doi.org/10.1099/ijss.0.004838-0> (2009).
63. Ibrahim, A. M., Hamouda, R. A., El-Naggar, N. E. A. & Al-Shakankery, F. M. Bioprocess development for enhanced endoglucanase production by newly isolated bacteria, purification, characterization and *in-vitro* efficacy as anti-biofilm of *Pseudomonas aeruginosa*. *Sci. Rep.* **11**, 9754. <https://doi.org/10.1038/s41598-021-87901-9> (2021).
64. Anandham, R. et al. *Dyella thiooxydans* sp. nov., a facultatively chemolithotrophic, thiosulfate-oxidizing bacterium isolated from rhizosphere soil of sunflower (*Helianthus annuus* L.). *Int J Syst. Evol. Microbiol.* **61**, 392–398. <https://doi.org/10.1099/ijss.0.022012-0> (2011).
65. Liang, Y. L., Zhang, Z., Wu, M., Wu, Y. & Feng, J. X. Isolation, screening, and identification of cellulolytic bacteria from natural reserves in the subtropical region of China and optimization of cellulase production by *Paenibacillus Terrae* ME27-1. *Biomed. Res. Int.* **2014**, 512497. <https://doi.org/10.1155/2014/512497> (2014).
66. Samir Ali, S., Al-Tohamy, R., Sun, J., Wu, J. & Huizi, L. Screening and construction of a novel microbial consortium SSA-6 enriched from the gut symbionts of wood-feeding termite, *Coptotermes formosanus* and its biomass-based biorefineries. *Fuel* **236**, 1128–1145. <https://doi.org/10.1016/j.fuel.2018.08.117> (2019).

Author contributions

N. I, A. T., S. J., P. Y., S. N-R T.: Investigation, Formal analysis. A. T., S. N-R T.: Conceptualization, Supervision, Writing–original draft, Writing–review & editing. S. J., A. T., S. N-R T.: Methodology. A. T., S. J., P. Y., S. N-R T.: Review & editing.

Funding

This work was financially supported by Kasetsart University, Bangkok, Thailand through the Graduate School Fellowship Program and was partially supported by Microbiology Division, Department of Science and Bioinnovation, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Sean Campus, Thailand (Grant year 2025).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-33717-w>.

Correspondence and requests for materials should be addressed to S.N.-R.T.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2026