

***Paenibacillus glycinicola* sp. nov., *Paenibacillus trititicola* sp. nov., *Pseudescherichia oryzae* sp. nov., *Pseudomonas gossypii* sp. nov., and *Tardiphaga zae* sp. nov., five novel bacterial species from agricultural crops**

Sarah Seaton¹, Emily Auker¹, Allison Connolly¹, Mekdem Tesfaye¹, Patrik Inderbitzin¹, Stephen J. Barnett², Christopher M. M. Franco³, Martha E. Trujillo⁴

¹Indigo Ag, Inc., 500 Rutherford Avenue, Boston, MA 02129, United States

²South Australian Research and Development Institute, Adelaide, Australia

³Medical Biotechnology, College of Medicine and Public Health, Flinders University, Adelaide, Australia

⁴Departamento de Microbiología y Genética, Campus Miguel de Unamuno, University of Salamanca, Salamanca, Spain

ABSTRACT

Five novel species from healthy plants and other substrates are described. These are *Paenibacillus glycinicola* sp. nov. from *Glycine max* in Iowa, United States; *Paenibacillus trititicola* sp. nov. from *Triticum aestivum* in Australia, *Allium cepa* in Texas, garden soil in Wisconsin, cow's milk in New York state, *Panicum virgatum* in Michigan, sweet sorghum in Nebraska, and *Zea mays* in Ohio; *Pseudescherichia oryzae* sp. nov. from *Oryza sativa* in Arkansas; *Pseudomonas gossypii* sp. nov. from a healthy *Bos taurus* in Kansas, *Capsicum annuum* from Arizona, *Gossypium hirsutum* in New Mexico, Oklahoma, and Texas, and *Zea mays* from Arkansas; and *Tardiphaga zae* sp. nov. from *Zea mays* in Iowa. No pathogenic strains are known for any of the novel species.

INTRODUCTION

Agricultural crops have diverse microbiomes (Favela et al. 2022; Kawasaki et al. 2021; Kuźniar et al. 2020; Longley et al. 2020; Wei et al. 2019) that differ between plant species and compartments (Xiong et al. 2021), and play crucial roles for plant health (Berendsen et al. 2012). A better understanding of the interactions between microbiomes, plants, and the environment, is

essential for microbiome manipulation to the benefit of increased sustainability and climate change resilience in agroecosystems (Mariotte et al. 2018; Sergaki et al. 2018).

Here we describe five plant-associated strains, representing five novel species in the genera *Paenibacillus*, *Pseudescherichia*, *Pseudomonas*, and *Tardiphaga*. *Paenibacillus glycinicola* strain SS118 was isolated from healthy field-grown *Glycine max* in Iowa, United States; *Paenibacillus trititicola* strain PII19 was isolated from the roots of *Triticum aestivum* grown in field soil heated to 80°C from South Australia, Australia; *Pseudescherichia oryzae* strain SS120 was isolated from field-grown *Oryza sativa* in Arkansas, United States; *Pseudomonas gossypii* strain PII21 was isolated from healthy field-grown *Gossypium hirsutum* in New Mexico, United States; and *Tardiphaga zae* strain SS122 was isolated from healthy field-grown *Zea mays* in Iowa, United States. We provide phenotypic and phylogenomic details describing these five species, and report geographic distribution and host range based on comparison to whole genome assemblies from GenBank of conspecific, cultured strains.

METHODS

Isolation.

Strain SS118 was isolated from a stem of healthy field-grown *Glycine max* in Iowa, United States; strain PII19 was isolated from the roots of *Triticum aestivum* grown in field soil heated to 80°C, the soil was collected in South Australia, Australia; strain SS120 was isolated from a field-grown *Oryza sativa* seedling in Arkansas, United States; strain PII21 was isolated from a cotton bowl of a healthy field-grown *Gossypium hirsutum* in New Mexico, United States; and strain SS122 was isolated from a root of healthy field-grown *Zea mays* in Iowa, United States. For strains PII21, SS118, SS120, and SS122, plant tissue was washed with a mild detergent to remove particulates, surface-sterilized with bleach (1% v/v sodium hypochlorite) and ethanol (70% v/v), and homogenized. Serial dilutions of tissue homogenate were plated on a panel of media types for endophyte cultivation. For strain PII19, the roots were washed well, heated to 80°C for 30 min, and five-millimeter segments were placed onto Baker's Yeast agar (BYA1: 4g/L baker's yeast, 10 g/L agar, 100 mg/L cycloheximide) incubated at 25°C. All strains were streaked to purity and stored in glycerol (20% v/v) at -80°C until subjected to further testing.

Motility. The strains were tested for flagellar-dependent swimming and swarming motility on R2A plates solidified with 0.3% and 0.6% agar, respectively. Three independent colonies were inoculated onto R2A broth and grown for 36 hr at 24°C. Broth cultures were normalized to an OD600 of 0.1, and 1.5 µl of culture was spotted directly onto the surface of the motility agar. The diameter of colony expansion was measured for 5 days.

Carbon source utilization. Substrate utilization was assessed using Biolog GenIII Microplates (Catalogue No. 1030) (Biolog Inc., Hayward, CA). Each bacterium was inoculated in duplicate plates using Protocol A, described by the manufacturer, except that plates were incubated at 30°C for 24-48 hours, and *Tardiphaga zeeae* strain SS122 required 8 days of incubation. Respiration leading to reduction of the tetrazolium indicator was measured by absorbance at 590 nm.

Biochemical analyses. Catalase activity was evaluated by immediate effervescence after applying 3% (v/v) hydrogen peroxide solution via the tube method; a positive reaction was indicated by the production of bubbles. *Staphylococcus aureus* NCIMB 12702 and *Streptococcus pyogenes* ATCC 19615 were used as positive and negative controls, respectively. Oxidase activity was evaluated via the oxidation of Kovács oxidase reagent, 1% (w/v) tetra-methyl-p-phenylenediamine dihydrochloride in water, via the filter-paper spot method. A positive reaction was indicated when the color of the microorganism changed to dark purple. *Pseudomonas aeruginosa* NCIMB 12469 and *Escherichia coli* ATCC 25922 were used as positive and negative controls, respectively. Urease activity was evaluated via the hydrolysis of urea in Christensen's Urea Agar, using phenol red as a pH indicator. *Proteus hauseri* ATCC 13315 and *Escherichia coli* ATCC 25922 were used as positive and negative controls, respectively. Gram staining was performed using standard protocols.

Phylogenetic and genomic analyses. DNA was extracted from pure cultures using the Omega Mag-Bind Universal Pathogen Kit according to manufacturer's protocol with a final elution volume of 60µl (Omega Biotek Inc., Norcross, GA). DNA samples were quantified using a Qubit fluorometer (ThermoFisher Scientific, Waltham, MA) and normalized to 100 ng. DNA was prepped using a Nextera DNA Flex Library Prep kit according to the manufacturer's instructions (Illumina Inc., San Diego, CA). DNA libraries were quantified via qPCR using a KAPA Library Quantification kit (Roche Sequencing and Life Science, Wilmington, MA) and combined in equimolar concentrations into one 24-sample pool. Libraries were sequenced on a

MiSeq using pair-end reads (2x200bp). Reads were trimmed of adapters and low-quality bases using Cutadapt (version 1.9.1) and assembled into contigs using MEGAHIT (version 1.1.2) (Mitter et al. 2013). Reads were mapped to contigs using Bowtie2 (version 2.3.4) (Langmead and Salzberg 2012), and contigs were assembled into scaffolds using BESST (2.2.8) (Sahlin et al. 2014). Phylogenomic trees were generated using GToTree (version 1.2.1) (Lee 2019).

16S rRNA gene sequences were extracted from genome assemblies using barrnap (Seemann 2019), and 16S rRNA gene phylogenetic analyses were performed using FastTree (Price et al. 2010) using a General Time Reversible substitution model. Taxon sampling for each species is described in the respective phylogenetic tree figure legend.

Average nucleotide identity (ANI) analyses between genome assemblies were performed using fastANI (Jain et al. 2018).

Geographic distribution and host range of novel species were inferred by ANI to assemblies from unidentified species from GenBank (Ciufo et al. 2018) and the Indigo internal collection. An ANI threshold of $\geq 95\%$ indicated conspecificity (Chun et al. 2018; Richter and Rosselló-Móra 2009).

RESULTS

Phylogenetic and genomic analyses

Paenibacillus glycinicola sp. nov. strain SS118

The strain SS118 16S rRNA gene sequence OP556612 shared 99.6% identity with the 16S of *Paenibacillus azotifigens* NF2-4-5^T. A phylogenetic tree based on 16S using FastTree (Price et al. 2010) confirmed the affiliation of strain SS118 to the genus *Paenibacillus*. SS118 was most closely related to *Paenibacillus helianthi* P26E^T and *Paenibacillus sonchi* X19- 5^T with 82% support (Figure 1). The top average nucleotide identity (ANI) was to *Paenibacillus riograndensis* SBR5^T with 83.4%, well below the threshold for species demarcation (Chun et al. 2018; Richter and Rosselló-Móra 2009). This showed that strain SS118 represents a new genomic species of *Paenibacillus*.

Paenibacillus triticicola sp. nov. strain PII19

The strain PI119 16S rRNA gene sequence OP556609 shared 99.9% identity with the 16S of *Paenibacillus peoriae* NBRC 15541^T. A phylogenetic tree based on 16S rRNA gene sequences using FastTree (Price et al. 2010) confirmed the affiliation of strain PI119 to the genus *Paenibacillus*. PI119 was most closely related to *P. peoriae* and *Paenibacillus kribbensis* AM49^T with 76% support (Figure 2). The top average nucleotide identity (ANI) was to *Paenibacillus polymyxa* ATCC 842^T with 89.8%, well below the threshold for species demarcation of 95% ANI (Chun et al. 2018; Richter and Rosselló-Móra 2009). This showed that strain PI119 represents a new genomic species of *Paenibacillus*.

Pseudescherichia oryzae sp. nov. strain SS120

The strain SS120 16S rRNA gene sequence OP556613 shared 98.3% identity with the 16S of *Klebsiella aerogenes* ATCC 13048^T. Due to this low top 16S identity, and the fact that the 16S is not a reliable phylogenetic marker in the *Enterobacteriaceae* (Dauga 2002), we performed a phylogenomic analysis using GToTree (Lee 2019) with all 261 ex-type assemblies of the *Enterobacteriaceae* in GenBank. Strain SS120 was most closely related to *Pseudescherichia vulneris* NCTC12130^T with 100% support (Figure 3). The top average nucleotide identity (ANI) was to *P. vulneris* with 82.8%, well below the threshold for species demarcation (Chun et al. 2018; Richter and Rosselló-Móra 2009). This showed that strain SS120 represents a new genomic species of *Pseudescherichia*.

Pseudomonas gossypii sp. nov. strain PI121

The strain PI121 16S rRNA gene sequence OP556610 shared 99.3% identity with the 16S of *Pseudomonas oryzae* LMG 7040^T. A phylogenetic tree using FastTree (Price et al. 2010) confirmed the affiliation of strain PI121 to the genus *Pseudomonas*. PI121 was most closely related to *P. oryzae* and *Pseudomonas oleovorans* IAM 1508^T with 100% support (Figure 4). The top average nucleotide identity (ANI) was to *P. oryzae* with 91.2%, well below the threshold for species demarcation (Chun et al. 2018; Richter and Rosselló-Móra 2009), and showing that strain PI121 represents a new genomic species of *Pseudomonas*.

Tardiphaga zeae sp. nov. strain SS122

The strain SS122 16S rRNA gene sequence OP556611 shared 99.5% identity with the 16S rRNA gene sequence of *Tardiphaga robiniae* R-45977^T. A phylogenetic tree using FastTree (Price et al. 2010) confirmed the affiliation of strain SS122 to the genus *Tardiphaga*. SS122 was most closely related to *T. robiniae* with 100% support (Figure 5). The top average nucleotide identity (ANI) was to *T. robiniae* with 91.8%, well below the threshold for species demarcation (Chun et al. 2018; Richter and Rosselló-Móra 2009), and showing that strain PI121 represents a new genomic species of *Tardiphaga*.

Geographic distribution and host range

Geographic distribution and host range of the novel species was inferred by comparison to congeneric genome assemblies from GenBank and the Indigo internal collection. Distribution maps of the individual species for the United States are included (Figure 6, Figure 7, Figure 8, Figure 9, Figure 10). A summary of the known substrates is provided in Table 1.

Paenibacillus glycinicola sp. nov. strain SS118

There were no hits of the strain SS118 genome assembly against GenBank and Indigo genomes at an ANI \geq 95% threshold. The known geographic distribution of *Paenibacillus glycinicola* sp. nov. in the United States is illustrated in Figure 6.

Paenibacillus tritivicola sp. nov. strain PI119

Hits to strain PI119 included the following GenBank assemblies: Assembly GCF_009827695.1 from the rhizosphere of *Allium cepa* in Texas (Garcia et al. 2020) (ANI: 95.5%; query coverage: 63.7%), assembly GCF_022537985.1 from garden soil in Wisconsin, (ANI: 95.5%, query coverage: 87.0%), five assemblies from milk in New York state (Trmčić et al. 2015) (GCF_001955935.1, GCF_001956115.1, GCF_001956155.1, GCF_001956225.1, GCF_001956235.1; ANI range: 95.5 – 95.9%, query coverage range: 86.7 – 88.3%), assembly GCF_017833155.1 from the rhizosphere of *Panicum virgatum* in Michigan (ANI: 96.0%, query coverage: 87.8%), and assembly GCF_017833915.1 from the roots of sweet sorghum (ANI: 95.1%, query coverage: 86.0%) from Nebraska. From the Indigo internal collection, there was a hit to the assembly of strain P123 from *Zea mays* from Ohio (Figure 7).

Pseudoscherichia oryzae sp. nov. strain SS120

There were no hits of the strain SS120 genome assembly against GenBank and Indigo genomes at an ANI \geq 95% threshold. The known geographic distribution of *Pseudoscherichia oryzae* sp. nov. in the United States is illustrated in Figure 8.

Pseudomonas gossypii sp. nov. strain PI121

There was one hit among GenBank assemblies to strain PI121, which was GCF_001518815.1 from the nasopharynx of a healthy calf (*Bos taurus*) in Kansas. Among the Indigo collection, there were hits against strain SS124 from a leaf of *Capsicum annum* in Arizona (ANI: 98.2%, query coverage: 91.2%), strain PI125 from a leaf of *Gossypium hirsutum* in Oklahoma (ANI: 98.0%, query coverage: 94.3%), strain SS126 from a seedling of *Gossypium hirsutum* in Texas (ANI: 98.1%, query coverage: 91.2%), and strain PI127 from a seedling of *Zea mays* from Arkansas (ANI: 98.0%, query coverage: 90.9%) (Figure 9).

Tardiphaga zeae sp. nov. strain SS122

There were no hits of the strain SS122 genome assembly against GenBank and Indigo genomes at an ANI \geq 95% threshold. The known geographic distribution of *Tardiphaga zeae* sp. nov. in the United States is illustrated in Figure 10.

Morphology, physiology and biochemical characteristics

Paenibacillus glycinicola sp. nov. strain SS118, and *P. tritricicola* sp. nov. strain PI119

The two *Paenibacillus* strains, SS118 and PI119 had similar morphological characteristics but important differences were found. On R2A agar, colonies of SS118 were irregular and flat, with undulate margins, and were non-pigmented. The same colony characteristics were recorded for PI119 on Tryptic Soy agar (TSA), as this strain did not grow on R2A. Concerning Gram staining, SS118 was variable, while PI119 yielded a positive reaction. Cells of both strains are rod-shaped but differ in size, with SS118 measuring 2 – 4 μ m long and those of PI119 only 1 – 2 μ m (Figure 11A, Figure 11B). Motility was recorded for strain PI119

and showed growth at 30 and 37°C, while strain SS118 is non-motile and only grew at 30°C. Neither strain grew at 42°C. The new *Paenibacillus* strains are aerobic with a respiratory metabolism. Catalase production was detected in PI119, but neither strain produced oxidase nor urease. Physiological characterization using the Biolog GenIII system showed that the strains could use various common compounds as carbon and energy sources: D-maltose, D-cellobiose, lactose, melibiose, D-salicin, α -D-glucose, D-fructose, D-galactose and sodium lactate (1%). Differences were found, as only strain PI119 metabolized D-trehalose, sucrose, D-turanose, stachyose, D-raffinose, β -Methyl-D-glucoside, D-sorbitol D-mannitol and glycerol. Other results are given in the corresponding species descriptions. Concerning NaCl tolerance, both strains grew at concentrations up to 8%. In addition, the new bacteria were sensitive to the antibiotics troleandomycin, lincomycin and rifamycin SV, however, PI119 showed resistance to minocycline.

Distinguishing phenotypic characteristics between *Paenibacillus glycinicola* sp. nov. and the closest phylogenetic type strains *P. helianthi* P26E^T and *P. sonchi* X19-5^T is summarized in Table 2. Differences between *P. triticicola* sp. nov. and *P. peoriae* NBRC 1554^T and *P. kribbensis* AM49^T are given in Table 3.

Pseudodescherichia oryzae sp. nov. strain SS120

The cells of strain SS120 are non-motile and stain Gram-negative. On TSA agar, circular, non-pigmented, and umbonate colonies with undulate margins are observed. Cells are small rod-shaped and measure 0.5 – 1µm long (Figure 11C). Strain SS120 can grow 30 – 42° and at pH 5 and 6, and in the presence of up to 4% NaCl. There is a weak reaction for catalase production but not for oxidase or urease. Under the Biolog GenIII system to assess carbohydrate utilization, the strain weakly catabolized D-maltose, D-trehalose, D-cellobiose and gentobiose. Other sugars used as carbon sources include glycerol, D-glucose-6-phosphate and D-fructose-6-phosphate. Other results are given in the species description. There is resistance to rifamycin SV and lincomycin, but not to minocycline. SS120 is also capable of growing in the presence of the chemicals tetrazolium blue and tetrazolium violet, but only shows weak growth against lithium chloride.

Differential phenotypic characteristics between strain SS120 and the only representative of the genus *Pseudodescherichia*, *P. vulneris* NBRC 102420^T are presented in Table 4.

Pseudomonas gossypii sp. nov. strain PI121

Strain PI121 stained Gram-negative. On TSA, flat, yellow-pigmented colonies with entire margins are observed. Cells are rod-shaped and motile, measuring approximately 0.5 – 1 µm long (Figure 11D). Growth was recorded at 30 and 37°C but not at 42°C. The strain also grew at pH 5 and 6 and tolerated NaCl concentrations up to 4%. Strain PI121 had a weak catalase reaction. On the contrary, it did not produce oxidase or urease. Being aerobic, the strain has a respiratory metabolism. Based on the Biolog GenIII system, strain PI121 can use a diverse range of compounds as carbon and energy sources. These include D-trehalose, sucrose, α-D-glucose, D-mannose, D-fructose, D-galactose, D-sorbitol, D-mannitol, D-arabitol, myo-inositol and glycerol in addition to several organic acids. Other results are given in the species description. The strain shows resistance to rifamycin SV and lincomycin but is susceptible to minocycline, and nalidixic acid. The strain grows in the presence of several chemical inhibitors, which include tetrazolium violet, tetrazolium blue, lithium chloride, potassium tellurite, and Tween 40.

Distinguishing characteristics between the new strain and the closest phylogenetic type strains of the species, *P. luteola* NBRC 103146, *P. oryzihabitans* LMG 7040 and *P. oleovorans* IAM 1508, are given in Table 5.

Tardiphaga zeae sp. nov. strain SS122

Cells of strain SS122 stain Gram-negative. On R2A agar colonies are non-pigmented, circular, flat with entire margins, and shows very slow growth similar to *Tardiphaga robiniae* LMG 26467^T, the only representative of the genus. Cells are rod-shaped, motile and small, measuring less than 2 µm. The strain can grow at 30°C but not at 37 or 42°C, and at pH 6 but not pH 5. Carbon source characterization with the Biolog Gen III system showed that the strain can use fusidic acid, glucuronamide, quinic acid, α-keto-glutaric acid, D-malic acid, L-malic acid, and acetic acid. The strain shows resistance to aztreonam, nalidixic acid, lincomycin, minocycline and rifamycin SV.

Several phenotypic differences between strain SS122 and *Tardiphaga robiniae* LMG 26467^T are given in Table 6.

DESCRIPTION OF *PAENIBACILLUS GLYCINICOLA* SP. NOV.

Paenibacillus glycinicola (gly.ci.ni'co.la. N.L. fem. n. *Glycine*, the genus of the soybean; L. masc./fem. n. suff. -cola, inhabitant dweller; from L. masc./fem. n. *incola*, dweller; N.L.).

Gram variable, rod-shaped, nonmotile bacterium (2-4 µm long). Irregular, non-pigmented colonies on R2A agar. Does not grow on TSA. Growth observed at 30°C but not at 37 or 42 °C. Grows at pH 6 but not at 5 and on 8% (w/v) NaCl. Neither catalase, oxidase nor urease are produced. Weakly uses gentiobiose as a carbon source, but not N-acetyl-D-glucosamine, N-acetyl-β-D-mannosamine, N-Acetyl-D-galactosamine, N-Acetyl-neuraminic acid, D-mannose, 3-methyl-glucose, DL-fucose, L-rhamnose or inosine. Most amino acids are not used as carbon sources. Growth is not inhibited by potassium tellurite.

The type strain SS118 was isolated from a stem of a healthy field-grown plant of *Glycine max* in Iowa.

The G+C content is 49.4%.

DESCRIPTION OF *PAENIBACILLUS TRITICICOLA* SP. NOV.

Paenibacillus trititicola (tri.ti.ci.co'la. L. neut. n. *Triticum*, wheat; L. masc./fem. n. suff. -cola, inhabitant dweller; from L. masc./fem. n. *incola*, dweller; N.L. masc./fem. n. *trititicola*, wheat dweller).

Gram stain positive, rod-shaped, motile bacterium (1 – 2 µm long). Irregular, non-pigmented colonies with undulate margin on TSA agar. Growth is observed at 30 and 37°C but not at 42°C. Grows at pH 6 but only weakly at pH 5 and in 4% (w/v) NaCl. Catalase is produced but not oxidase or urease. The following compounds are used as carbon sources: dextrin, D-trehalose, sucrose, D-turanose, stachyose, raffinose, β-methyl-D-glucoside, D-mannose, D-sorbitol, D-mannitol, glycerol, pectin, and D-gluconic acid. Growth not inhibited by tetrazolium violet and lithium chloride.

The type strain PI119 was isolated from the roots of *Triticum aestivum* collected in South Australia.

The G+C content is 45.6%.

DESCRIPTION OF *PSEUDESCHERICHIA ORYZAE* SP. NOV.

Pseudescherichia oryzae (o.ry'zæ. L. gen. fem. n. *oryzae*, of rice).

Gram stain-negative, rod-shaped (0.5-1µm long), non-motile bacterium. Circular, non-pigmented, umbonate colonies with undulate margin on TSA agar. Growth observed at 30°C but not at 37 or 42°C. Grows at pH 6 but not at pH 5 and in 4% (w/v) NaCl. Catalase is weakly produced but not oxidase or urease. Dextrin, D-maltose, D-trehalose, D-cellobiose, gentiobiose, α-D-lactose, β-methyl-D-glucoside, D-salicin, D-galactose, inosine, 1% sodium lactate, fusidic acid, D-sorbitol, D-arabitol, glycerol, D-glucose-6-PO₄ and D-fructose-6-PO₄. Growth is not inhibited by tetrazolium violet, tetrazolium blue.

The type strain SS120 was isolated from a field-grown *Oryza sativa* seedling in Arkansas.

The G+C content is 55.9%.

DESCRIPTION OF *PSEUDOMONAS GOSSYPHII* SP. NOV.

Pseudomonas gossypii (gos.sy'pi.i. N.L. gen. neut. n. *gossypii*, of the cotton genus *Gossypium*).

Gram stain-negative, rod-shaped (0.5 – 1µm long), motile bacterium. Circular, yellow-pigmented, colonies on TSA agar. Growth observed at 30 and 37°C but not at 42 °C. Grows at pH 5 and 6 and in 4% (w/v) NaCl. Catalase is weakly produced but not oxidase or urease. In addition to the substrates above, the following compounds are used as carbon sources: 1% sodium lactate, fusidic acid, inosine, D-sorbitol, D-arabitol, glycerol, D-glucose-6-PO₄, D-

fructose-6-PO₄, DL-galacturonic acid, D-gluconic acid, D-glucuronic acid, mucic acid, D-saccharic acid, lactic acid and citric acid. Grows in the presence of tetrazolium violet and tetrazolium blue, niaproof 4, and weakly in potassium tellurite.

The type strain PI121 was isolated from a cotton bowl of a healthy field-grown *Gossypium hirsutum* in New Mexico.

The G+C content is 65.1%.

DESCRIPTION OF *TARDIPHAGA ZEAE* SP. NOV.

Tardiphaga zeae (ze'ae. L. gen. fem. n. *zeae*, of *Zea mays*).

Gram stain-negative. Rod-shaped (<2µm long), motile, and slow growing bacterium. Circular, non-pigmented colonies on R2A, does not grow on TSA agar. Growth observed at 30°C but not at 37 or 42 °C. Grows at pH 6 but not at pH 5. Does not grow in the presence of NaCl. Produces oxidase but not urease. Catalase is weakly produced. Based on the Biolog GENIII system only a few compounds are used as carbon sources: D-serine, 1% sodium lactate, fusidic acid, D-gluconic acid, glucuronamide, quinic acid, methyl pyruvate, L-lactic acid, acetic acid, and formic acid. The following sugars do not serve as carbon sources: dextrin, D-maltose, D-trehalose, D-cellobiose, D-turanose, gentiobiose, sucrose, stachyose, D-raffinose, α-D-lactose, D-melibiose, D-salicin, α-D-glucose, D-mannose, D-fructose, D-galactose, D-fucose, L-rhamnose, inosine, D-sorbitol, D-mannitol, D-arabitol, myo-inositol and glycerol. Gelatin is not hydrolyzed.

The type strain SS122 was isolated from a root of healthy field-grown *Zea mays* in Iowa.

The G+C content is 61.4%.

ACKNOWLEDGEMENTS

We thank Dr. Aljosa Trmcic, Wiedmann Lab, Cornell University, for pointing us toward the Food Microbe Tracker. We also thank Professor Aharon Oren, The Hebrew University of Jerusalem, for help with nomenclature.

REFERENCES

- Ambrosini, A., Sant' Anna, F. H., Heinzmann, J., de Carvalho Fernandes, G., Bach, E., and Passaglia, L. M. P. 2018. *Paenibacillus helianthi* sp. nov., a nitrogen fixing species isolated from the rhizosphere of *Helianthus annuus* L. *Antonie Van Leeuwenhoek*. 111:2463–2471.
- Berendsen, R. L., Pieterse, C. M., and Bakker, P. A. 2012. The rhizosphere microbiome and plant health. *Trends Plant Sci.* 17:478–486.
- Brenner, D. J., McWhorter, A. C., Knutson, J. K. L., and Steigerwalt, A. G. 1982. *Escherichia vulneris*: a new species of Enterobacteriaceae associated with human wounds. *J. Clin. Microbiol.* 15:1133–1140.
- Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., da Costa, M. S., et al. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68:461–466.
- Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., et al. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int. J. Syst. Evol. Microbiol.* 68:2386–2392.
- Dauga, C. 2002. Evolution of the *gyrB* gene and the molecular phylogeny of Enterobacteriaceae: a model molecule for molecular systematic studies. *Int. J. Syst. Evol. Microbiol.* 52:531–547.
- De Meyer, S. E., Coorevits, A., and Willems, A. 2012. *Tardiphaga robiniae* gen. nov., sp. nov., a new genus in the family Bradyrhizobiaceae isolated from *Robinia pseudoacacia* in Flanders (Belgium). *Syst. Appl. Microbiol.* 35:205–214.
- Favela, A., Bohn, M., and Kent, A. 2022. N-cycling microbiome recruitment differences between modern and wild *Zea mays*. *Phytobiomes J.* :151–160.
- Garcia, A. M., Carlo, P. P., and Mendez, M. O. 2020. Draft Genome Sequence of *Paenibacillus* sp. Strain OT2-17, a Triclosan-Degrading Rhizobacterium. *Microbiol. Resour. Announc.* 9:e01596-19.
- Hauser, E., Kämpfer, P., and Busse, H.-J. 2004. *Pseudomonas psychrotolerans* sp. nov. *Int. J. Syst. Evol. Microbiol.* 54:1633–1637.
- Hong, Y.-Y., Ma, Y.-C., Zhou, Y.-G., Gao, F., Liu, H.-C., and Chen, S.-F. 2009. *Paenibacillus sonchisp.* nov., a nitrogen-fixing species isolated from the rhizosphere of *Sonchus oleraceus*. *Int. J. Syst. Evol. Microbiol.* 59:2656–2661.
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114.

- Kawasaki, A., Dennis, P. G., Forstner, C., Raghavendra, A. K., Richardson, A. E., Watt, M., et al. 2021. The microbiomes on the roots of wheat (*Triticum aestivum* L.) and rice (*Oryza sativa* L.) exhibit significant differences in structure between root types and along root axes. *Funct. Plant Biol.* 48:871–888.
- Kuźniar, A., Włodarczyk, K., Grządziel, J., Goraj, W., Gałazka, A., and Wolińska, A. 2020. Culture-independent analysis of an endophytic core microbiome in two species of wheat: *Triticum aestivum* L.(cv. ‘Honda’) and the first report of microbiota in *Triticum spelta* L.(cv. ‘Rokosz’). *Syst. Appl. Microbiol.* 43:126025.
- Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with bowtie 2. *Nat Methods.* 9 Available at: <https://doi.org/10.1038/nmeth.1923>.
- Lee, M. D. 2019. Applications and considerations of GToTree: a user-friendly workflow for phylogenomics. *Evol. Bioinforma.* 15:1176934319862245.
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.* 31:1674–1676.
- Liu, R., Liu, H., Feng, H., Wang, X., Zhang, C.-X., Zhang, K.-Y., et al. 2008. *Pseudomonas duriflava* sp. nov., isolated from a desert soil. *Int. J. Syst. Evol. Microbiol.* 58:1404–1408.
- Longley, R., Noel, Z. A., Benucci, G. M. N., Chilvers, M., Trail, F., and Bonito, G. 2020. Crop management impacts the soybean (*Glycine max*) microbiome. *Front. Microbiol.* 11:1116.
- Mariotte, P., Mehrabi, Z., Bezemer, T. M., De Deyn, G. B., Kulmatiski, A., Drigo, B., et al. 2018. Plant–soil feedback: bridging natural and agricultural sciences. *Trends Ecol. Evol.* 33:129–142.
- Montefusco, A., Nakamura, L., and Labeda, D. 1993. *Bacillus peoriae* sp. nov. *Int. J. Syst. Evol. Microbiol.* 43:388–390.
- Price, M. N., Dehal, P. S., and Arkin, A. P. 2010. FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS ONE.* 5:e9490.
- Richter, M., and Rosselló-Móra, R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci.* 106:19126–19131.
- Romanenko, L. A., Uchino, M., Falsen, E., Frolova, G. M., Zhukova, N. V., and Mikhailov, V. V. 2005. *Pseudomonas pachastrellae* sp. nov., isolated from a marine sponge. *Int. J. Syst. Evol. Microbiol.* 55:919–924.
- Sahlin, K., Vezzi, F., Nystedt, B., Lundeberg, J., and Arvestad, L. 2014. BESST-efficient scaffolding of large fragmented assemblies. *BMC Bioinformatics.* 15:281.
- Seemann, T. 2019. *barrnap 0.9: rapid ribosomal RNA prediction*. Available at: <https://github.com/tseemann/barrnap>.

- Sergaki, C., Lagunas, B., Lidbury, I., Gifford, M. L., and Schäfer, P. 2018. Challenges and approaches in microbiome research: from fundamental to applied. *Front. Plant Sci.* 9:1205.
- Trmčić, A., Martin, N., Boor, K., and Wiedmann, M. 2015. A standard bacterial isolate set for research on contemporary dairy spoilage. *J. Dairy Sci.* 98:5806–5817.
- Wang, L.-Y., Li, J., Li, Q. X., and Chen, S.-F. 2013. *Paenibacillus beijingensis* sp. nov., a nitrogen-fixing species isolated from wheat rhizosphere soil. *Antonie Van Leeuwenhoek.* 104:675–683.
- Wei, F., Zhao, L., Xu, X., Feng, H., Shi, Y., Deakin, G., et al. 2019. Cultivar-dependent variation of the cotton rhizosphere and endosphere microbiome under field conditions. *Front. Plant Sci.* 10:1659.
- Xiong, C., Zhu, Y., Wang, J., Singh, B., Han, L., Shen, J., et al. 2021. Host selection shapes crop microbiome assembly and network complexity. *New Phytol.* 229:1091–1104.
- Yoon, J.-H., Oh, H.-M., Yoon, B.-D., Kang, K. H., and Park, Y.-H. 2003. *Paenibacillus kribbensis* sp. nov. and *Paenibacillus terrae* sp. nov., biofloculants for efficient harvesting of algal cells. *Int. J. Syst. Evol. Microbiol.* 53:295–301.

FIGURES AND TABLES

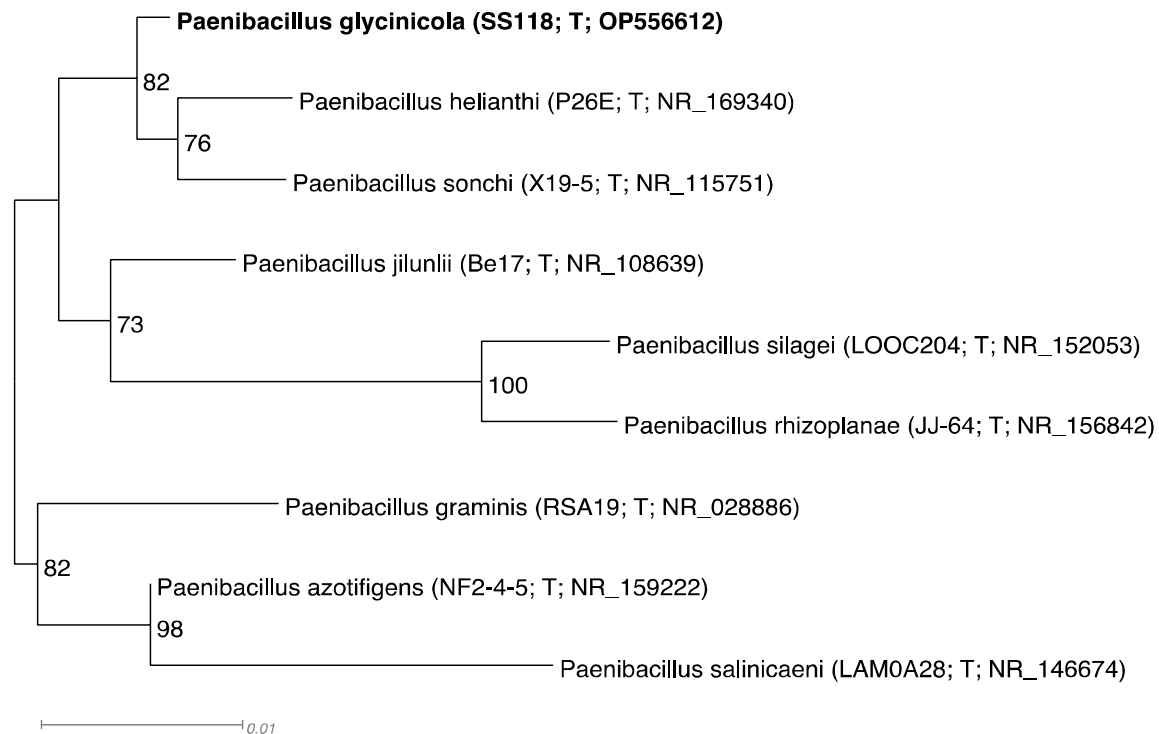


Figure 1. Phylogenetic 16S tree of *Paenibacillus glycinicola* sp. nov. and relatives generated using FastTree (Price et al. 2010). The eight species with the highest 16S identities to *P. glycinicola* were included in the tree, the tree is midpoint rooted. *Paenibacillus glycinicola* is marked in bold. Strain numbers and GenBank accession numbers follow species names, T stands for 'type'. Support values above 50% are given by the branches. *Paenibacillus glycinicola* is most closely related to *P. helianthi* and *P. sonchi* with 82% support. Branch lengths are proportional to the changes along the branches, a scale bar is provided.

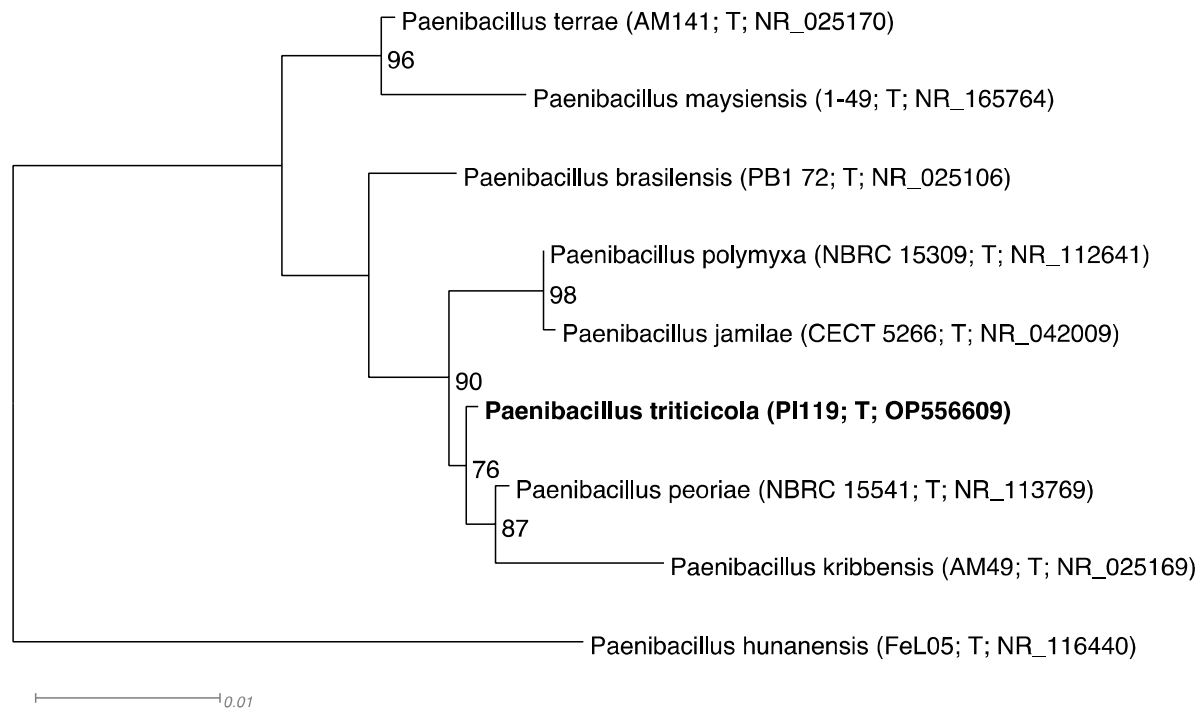


Figure 2. Phylogenetic 16S tree of *Paenibacillus tritricola* sp. nov. and relatives generated using FastTree (Price et al. 2010). The eight species with the highest 16S identities to *Paenibacillus tritricola* were included in the tree, the tree is midpoint rooted. *Paenibacillus tritricola* is marked in bold. Strain numbers and GenBank accession numbers follow species names, T stands for ‘type’. Support values above 50% are given by the branches. *Paenibacillus tritricola* is most closely related to *P. peoriae* and *P. kribbensis* with 76% support. Branch lengths are proportional to the changes along the branches, a scale bar is provided.

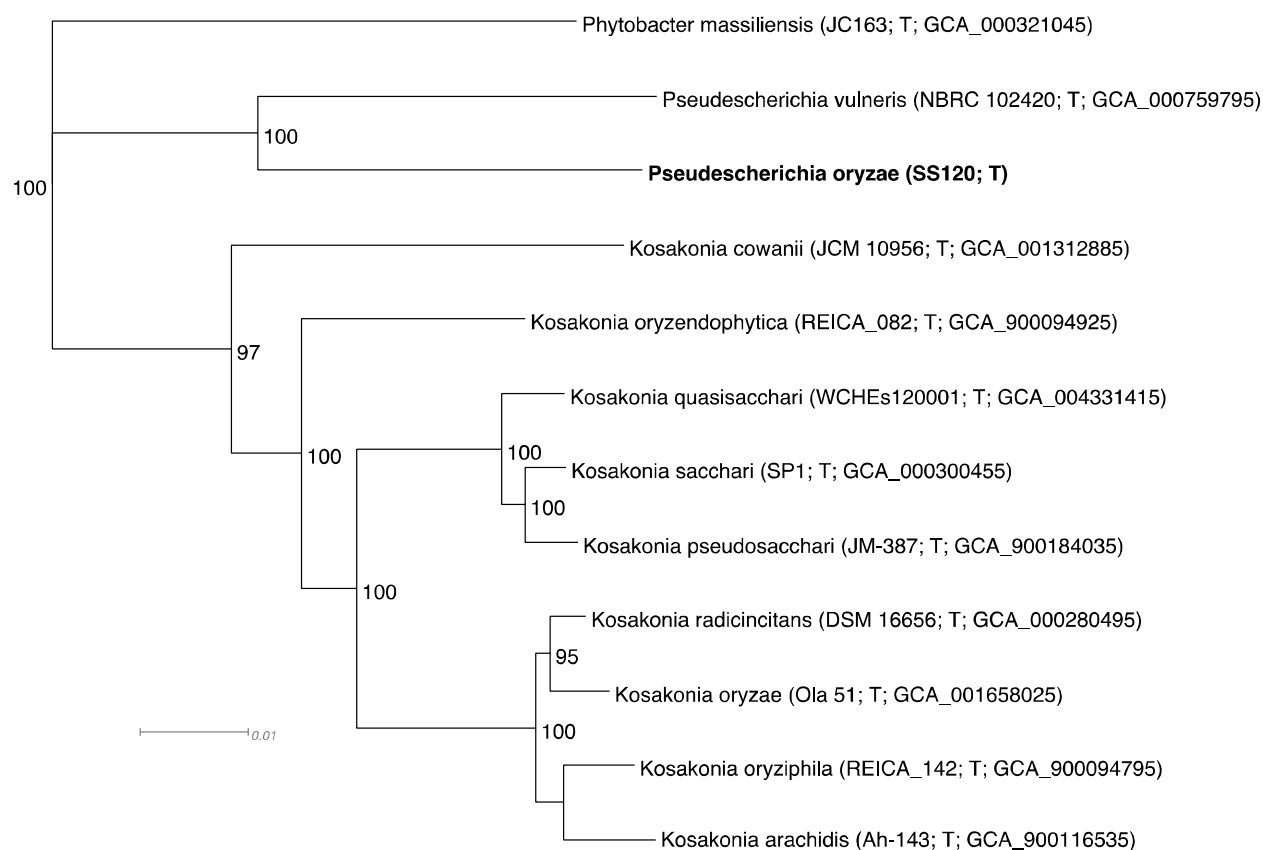


Figure 3. Phylogenomic tree of *Pseudoscherichia oryzae* sp. nov. and relatives inferred by GToTree (Lee 2019). Taxon sampling includes the closest relatives of *Pseudoscherichia oryzae* based on an *Enterobacteriaceae* species phylogenies (not shown). *Pseudoscherichia oryzae* is marked in bold. Strain numbers and GenBank accession numbers follow species names, T stands for ‘type’. Support values above 50% are given by the branches. *Pseudoscherichia oryzae* is most closely related to *Pseudoscherichia vulneris* with 100% support. Branch lengths are proportional to the changes along the branches, a scale bar is provided.

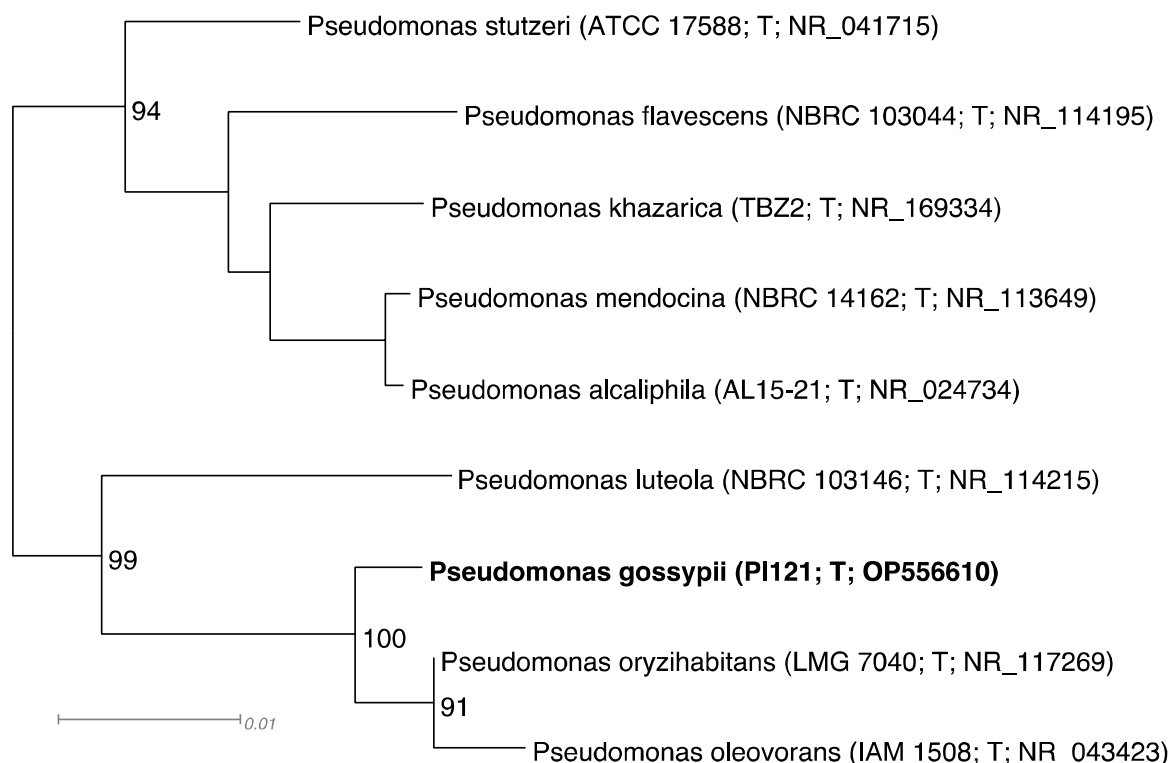


Figure 4. Phylogenetic 16S tree of *Pseudomonas gossypii* sp. nov. and relatives generated using FastTree (Price et al. 2010). The eight species with the highest 16S identities to *Paenibacillus gossypii* were included in the tree, the tree is midpoint rooted. *Pseudomonas gossypii* is marked in bold. Strain numbers and GenBank accession numbers follow species names, T stands for ‘type’. Support values above 50% are given by the branches. *Pseudomonas gossypii* is most closely related to *P. oryzihabitans* and *P. oleovorans* with 100% support. Branch lengths are proportional to the changes along the branches, a scale bar is provided.

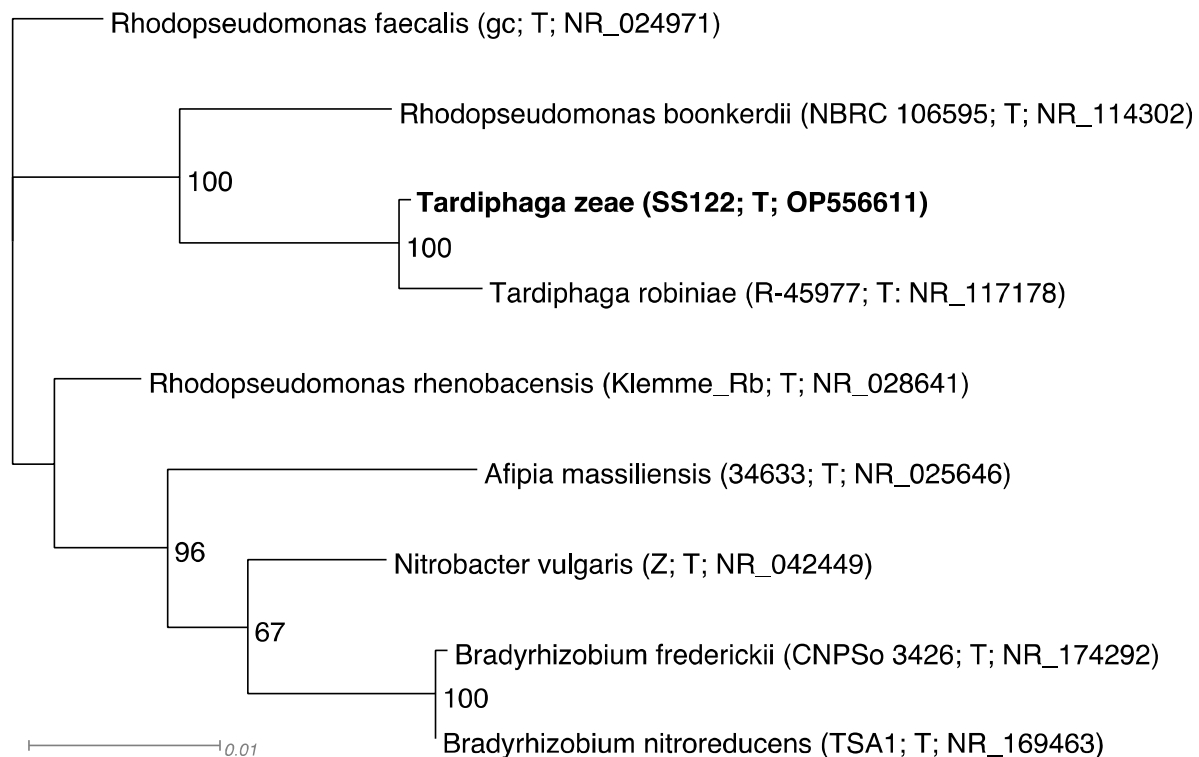


Figure 5. Phylogenetic 16S tree of *Tardiphaga zae* sp. nov. and relatives generated using FastTree (Price et al. 2010). The eight species with the highest 16S identities to *Tardiphaga zae* were included in the tree, the tree is midpoint rooted. *Tardiphaga zae* is marked in bold. Strain numbers and GenBank accession numbers follow species names, T stands for ‘type’. Support values above 50% are given by the branches. *Tardiphaga zae* is most closely related to *Tardiphaga robiniae* with 100% support. Branch lengths are proportional to the changes along the branches, a scale bar is provided.

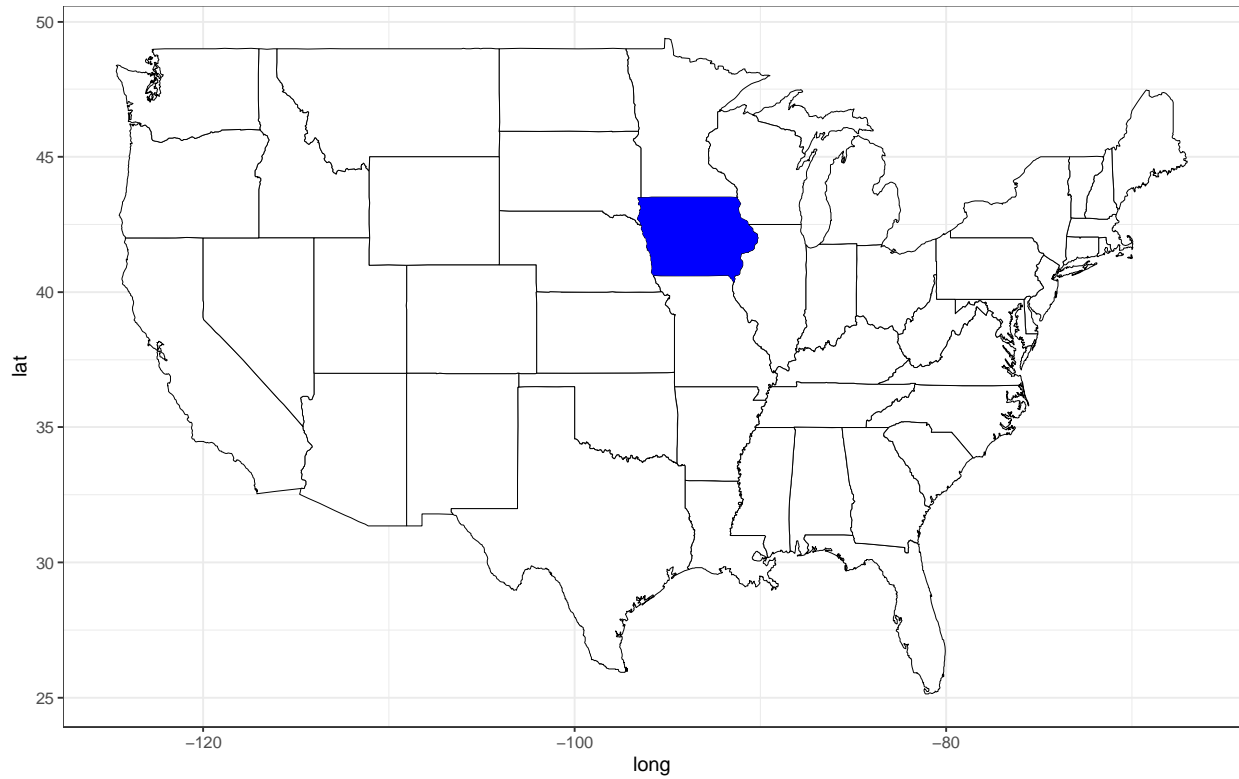


Figure 6. Geographic distribution of *Paenibacillus glycinicola* sp. nov. in the United States from culturing. Dark blue indicates the state of origin of the type strain. See text for details.

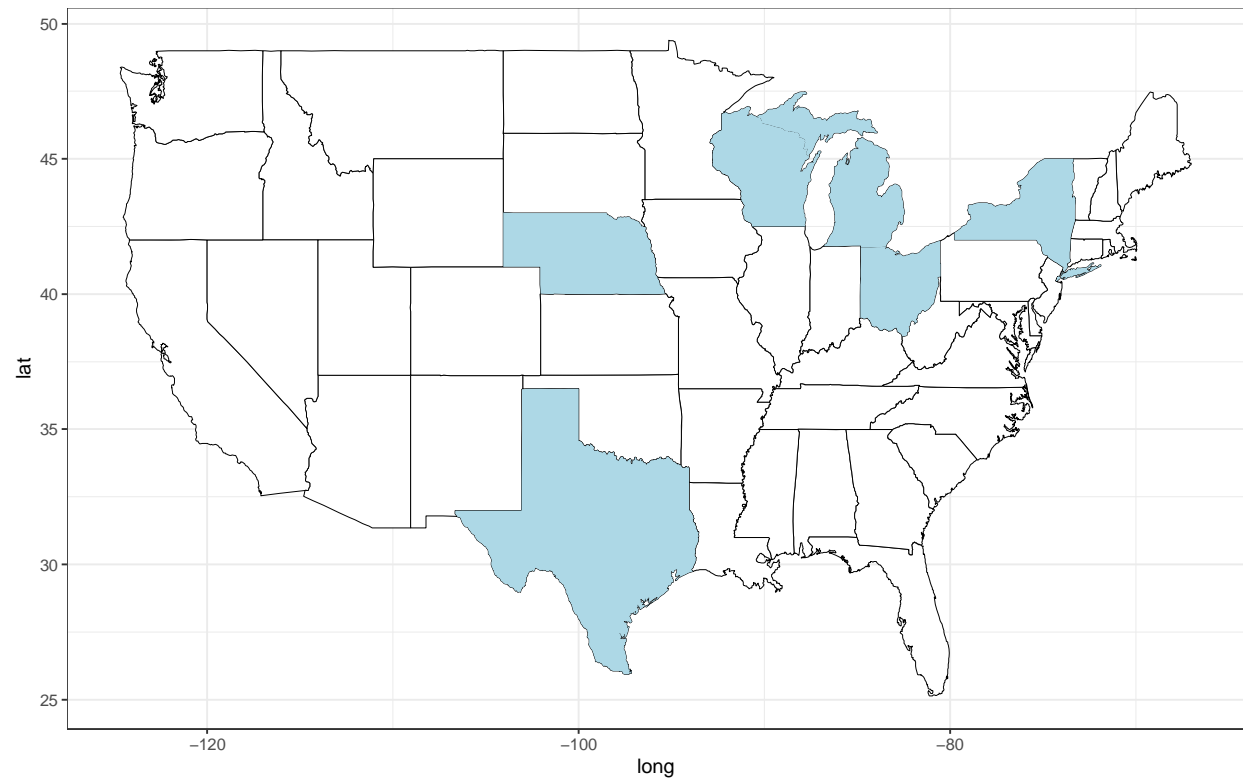


Figure 7. Geographic distribution of *Paenibacillus trititicola* sp. nov. in the United States from culturing. Light blue indicates the state of origin of *Paenibacillus trititicola* strains based on genome assemblies from GenBank and the Indigo internal collection. The type of *Paenibacillus trititicola* was isolated in Australia. See text for details.

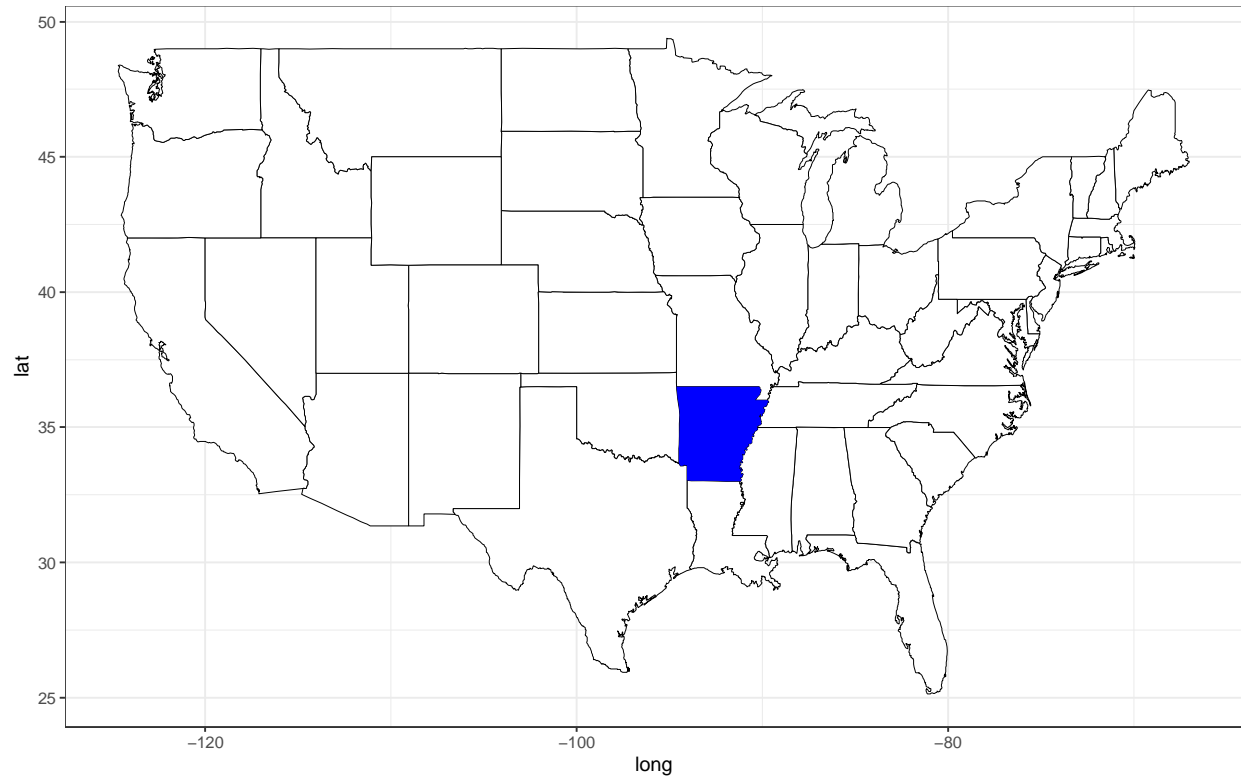


Figure 8. Geographic distribution of *Pseudoscherichia oryzae* sp. nov. in the United States from culturing. Dark blue indicates the state of origin of the type strain. See text for details.

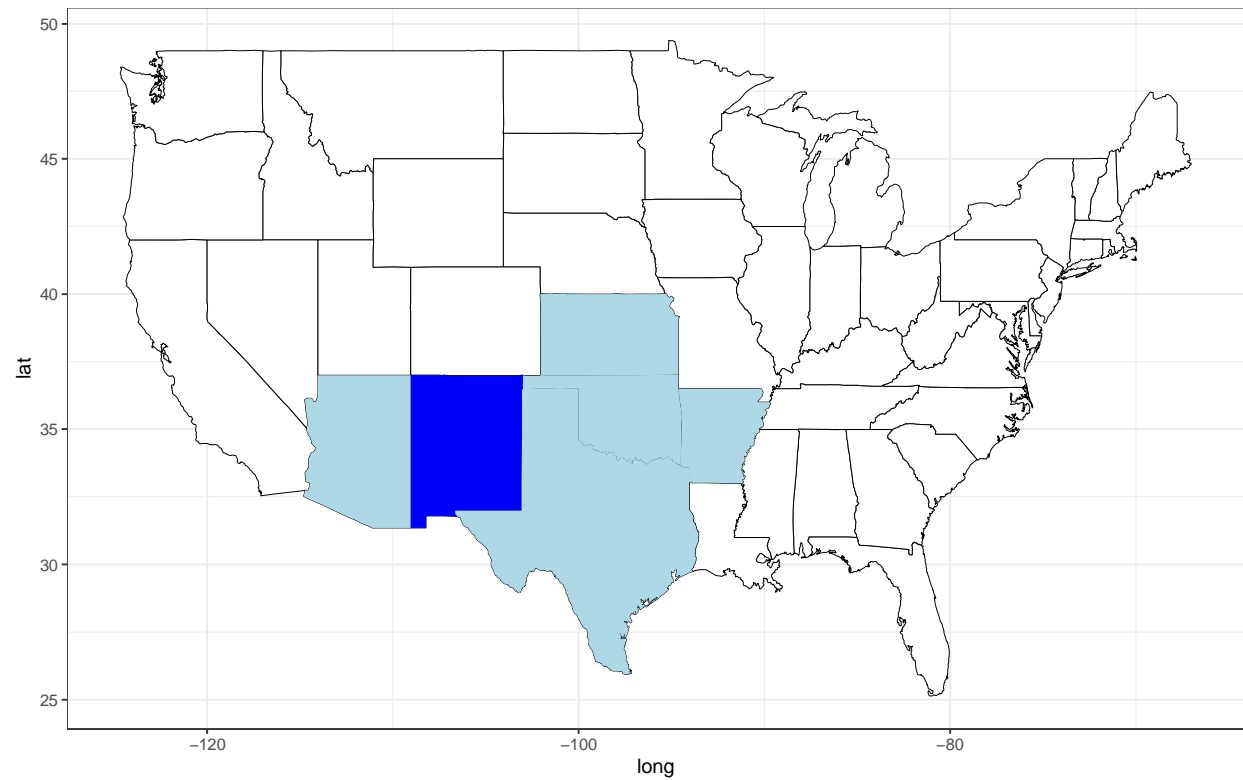


Figure 9. Geographic distribution of *Pseudomonas gossypii* sp. nov. in the United States from culturing. Dark blue indicates the state of origin of the type strain, light blue is the state of origin of additional *Pseudomonas gossypii* strains based on genome assemblies from GenBank and the Indigo internal collection. See text for details.

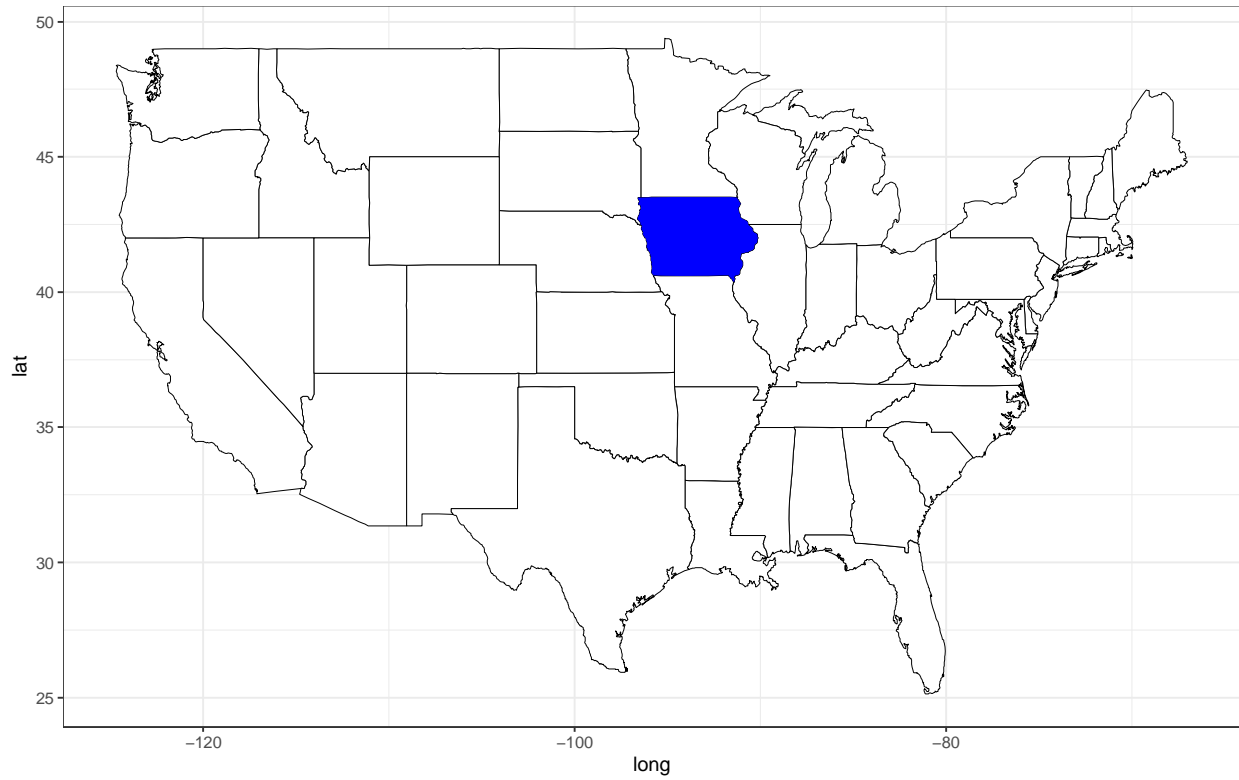


Figure 10. Geographic distribution of *Tardiphaga zeae* sp. nov. in the United States from culturing. Dark blue indicates the state of origin of the type strain. See text for details.

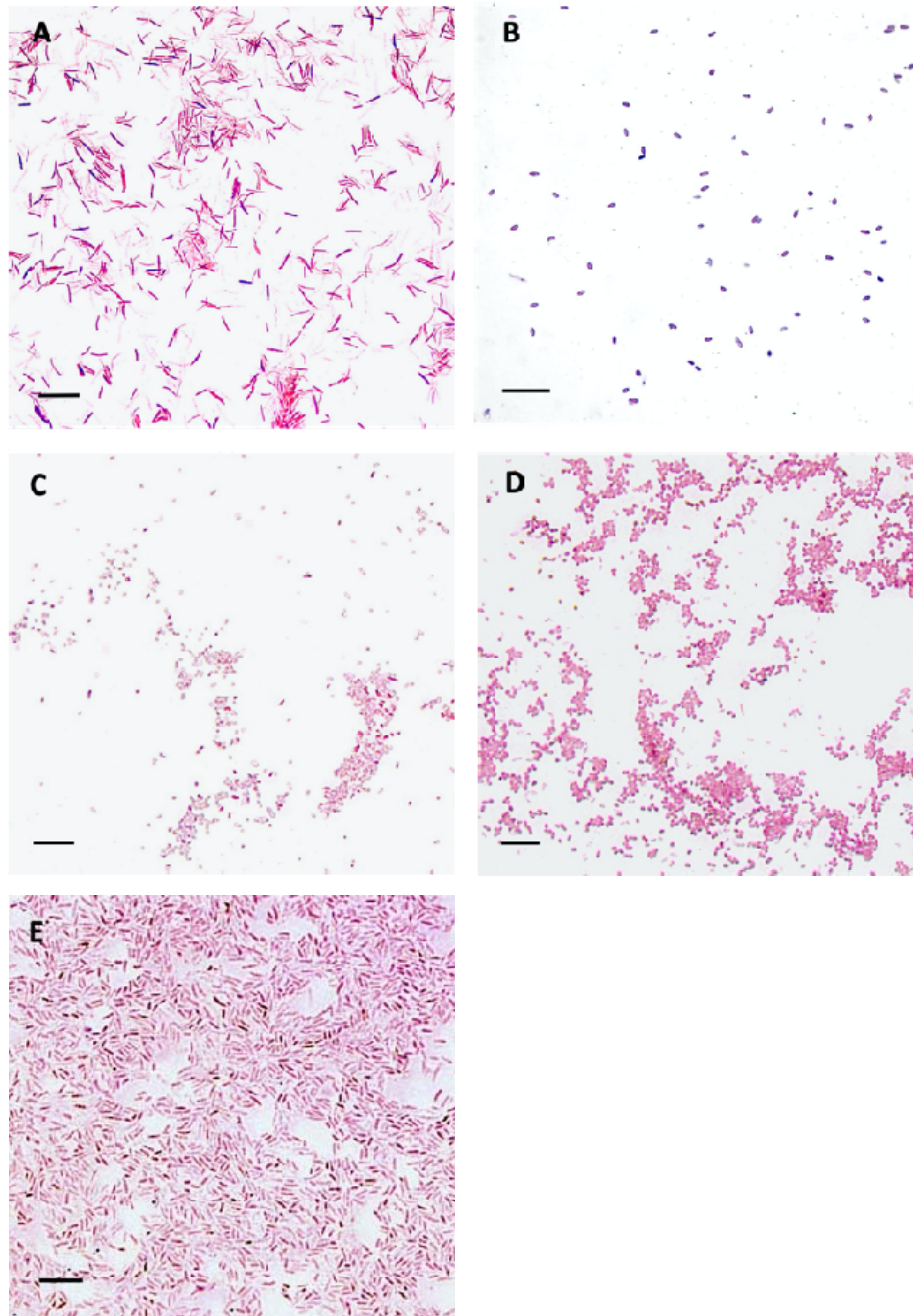


Figure 11. Morphology of new species depicted following Gram stain under bright field microscopy. A. *Paenibacillus glycinicola* sp. nov. strain SS118. B. *Paenibacillus triticiicola* sp. nov. strain PI119. C. *Pseudoscherichia oryzae* sp. nov. strain SS120. D. *Pseudomonas gossypii* sp. nov. PI121. E. *Tardiphaga zeae* sp. nov. strain SS122. Bar = 10 μ m.

Table 1. Substrates of novel species based on culturing.

Species	Substrate ¹
<i>Paenibacillus glycinicola</i>	<i>Glycine max</i>
<i>Paenibacillus trititicola</i>	<i>Allium cepa</i> , garden soil, milk, <i>Panicum virgatum</i> , sweet sorghum, <i>Triticum aestivum</i> , <i>Zea mays</i>
<i>Pseudescherichia oryzae</i>	<i>Oryza sativa</i>
<i>Pseudomonas gossypii</i>	<i>Bos taurus</i> , <i>Capsicum annuum</i> , <i>Gossypium hirsutum</i> , <i>Zea mays</i>
<i>Tardiphaga zeae</i>	<i>Zea mays</i>

¹ No disease symptoms reported for strains isolated from plants.

Table 2. Distinguishing characteristics of *Paenibacillus glycinicola* sp. nov. strain SS118 and closely related species.

	<i>Paenibacillus glycinicola</i> SS118	<i>Paenibacillus helianthi</i> P26E ^T	<i>Paenibacillus sonchi</i> X19-5 ^T
Motility	-	+	+
Temperature Range:			
Growth at 30°C	+	nr	+
Growth at 37°C	-	nr	+
Growth at 42°C	-	nr	-
Enzyme Activity:			
Oxidase	-	nr	-
Catalase	-	+	-
Urease	-	nr	nr
Utilization of:			
Fructose	+	+	+
D-glucose	+	+	-
Maltose	+	+	-
Sucrose	-	+	-

+, positive; -, negative; w, weak positive; nr, not reported.

Data from this study, Ambrosini et al. (2018) and Hong et al. (2009).

Table 3. Distinguishing characteristics of *Paenibacillus triticicola* sp. nov. strain PI119 and closely related species.

	<i>Paenibacillus triticicola</i> PI119	<i>Paenibacillus peoriae</i> NBRC 15541 ^T	<i>Paenibacillus</i> <i>kribbensis</i> AM49 ^T
Temperature Range:			
Growth at 42°C	-	nr	nr
Enzyme Activity:			
Oxidase	-	-	-
Catalase	+	+	+
Urease	-	nr	-
Utilization of:			
Dextrin	+	-	-
N-Acetyl-D-Glucoside	-	+	+
L-Rhamnose	-	+	+
<i>myo</i> -Inositol	-	+	+

+, positive; -, negative; w, weak positive; nr, not reported.

Data from this study, Montefusco et al. (1993), Wang et al. (2013), and Yoon et al. (2003).

Table 4. Distinguishing characteristics of *Pseudoscherichia oryzae* sp. nov. strain SS120 and closely related species.

	<i>Pseudoscherichia oryzae</i> SS120	<i>Pseudoscherichia vulneris</i> NBRC 102420 ^T
Motility	-	+
Temperature Range:		
Growth at 30°C	+	nr
Growth at 37°C	+	nr
Growth at 42°C	+	nr
Enzyme Activity:		
Catalase	w	+
Utilization of:		
D-glucose	w	+
Lactose	w	+
Sucrose	-	+
D-mannitol	w	+
D-sorbitol	+	-
Raffinose	-	+
Glycerol	+	-

+, positive; -, negative; w, weak positive; nr, not reported.

Data from this study and Brenner et al. (1982).

Table 5. Distinguishing characteristics of *Pseudomonas gossypii* sp. nov. strain PI121 and closely related species.

	<i>P. gossypii</i> PI121	<i>P. oryzihabitans</i> LMG7040	<i>P. oleovorans</i> IAM1508 ^T	<i>P. luteola</i> NBRC103146 ^T
Temperature Range:				
Growth at 42°C	-	-	+	+
NaCl tolerance:				
8%	w	-	-	-
Enzyme Activity:				
Oxidase	-	-	nr	-
Catalase	w	+	nr	+
Urease	-	+	nr	+
Utilization of:				
Maltose	w	+	-	+
Gluconate	+	+	-	+
D-Sorbitol	+	+	-	-
Propionic Acid	w	-	-	nr
D-mannose	+	+	-	+
D-galactose	+	+	-	nr
Mannitol	+	+	-	+

+, positive; -, negative; w, weak positive; nr, not reported.

Data from this study, Hauser et al. (2004), Liu et al. (2008), and Romanenko et al. (2005).

Table 6. Distinguishing characteristics of *Tardiphaga zae* sp. nov. strain SS122 and closely related species.

	<i>Tardiphaga zae</i> SS122	<i>Tardiphaga robiniae</i> LMG 26467 ^T	<i>Rhodopseudomonas</i> <i>bookkerdii</i> NBRC 106595 ^T
Motility	+	-	+
Enzyme Activity:			
Catalase	-	+	+
Utilization of:			
D-glucose	-	w	-
Malate	+	-	+

+, positive; -, negative; w, weak positive; nr, not reported.

Data from this study and De Meyer et al. (2012).