



# Complete Genome Sequences of *Curtobacterium*, *Pantoea*, *Erwinia*, and Two *Pseudomonas* sp. Strains, Isolated from Apple Flower Stigmas from Connecticut, USA

 Zhouqi Cui,<sup>a</sup> Blaire Steven,<sup>b</sup> Quan Zeng<sup>a</sup>

<sup>a</sup>Department of Plant Pathology and Ecology, The Connecticut Agricultural Experiment Station, New Haven, Connecticut, USA

<sup>b</sup>Department of Environmental Sciences, The Connecticut Agricultural Experiment Station, New Haven, Connecticut, USA

**ABSTRACT** The genome sequences of 5 bacterial strains isolated from apple flower stigmas are reported. The strains represent species of *Curtobacterium*, *Pantoea*, and *Erwinia* and two species of *Pseudomonas*. These data will provide information for future taxonomic studies and information for investigating the metabolic and functional characteristics of apple flower-colonizing bacteria.

The five bacterial strains, *Curtobacterium* sp. 24E2, *Erwinia* sp. 18B1, *Pantoea* sp. 1B4, *Pseudomonas* sp. 15A4, and *Pseudomonas* sp. 1079, were isolated from apple flower stigma samples collected on the apple cultivar ‘Early Macoun’ (*Malus x domestica* NY75414-1) planted at Lockwood Farm in Hamden, CT (41.406°N, 72.906°W) (1). Our previous data showed that *Enterobacteriaceae* and *Pseudomonadaceae* are the two predominant families of bacteria on apple stigma (2, 3). Within the *Enterobacteriaceae*, the genera *Pantoea* and *Erwinia* were predominant, and within the *Pseudomonadaceae*, *Pseudomonas* was the principal genus (3). The genus *Curtobacterium*, belonging to the phylum *Actinobacteria*, was much less abundant in the microbiome and was investigated as a representative of the apple stigma “rare biosphere” (3). These four strains were identified in our previous study by 16S rRNA gene comparisons (1). Acquisition of the full-genome sequence information of these strains will advance future investigations concerning the metabolic and functional requirements of the flower microbiome.

The stigma portion of an apple flower was dissected and immersed into 200  $\mu$ l of 0.5 $\times$  phosphate-buffered saline (PBS) in a sterile 1.5-ml microcentrifuge tube. The stigma sample was then sonicated for 5 min, followed by vortexing for 30 s. Then, 5 ml of PBS was spread onto a lysogeny broth (LB) agar plate to collect stigma resident bacteria (1). A single colony of each strain was inoculated into LB broth and incubated at 28°C overnight with shaking. Genomic DNA was extracted using the E.Z.N.A. (Omega, GA, USA) bacterial DNA kit according to the manufacturer’s instructions. The quality and quantity of isolated DNA were determined using the high-sensitivity D5000 ScreenTape system (Agilent Technologies, Santa Clara, CA) and the QuBit double-stranded DNA (dsDNA) broad-range assay (Thermo Fisher Scientific, Waltham, MA), respectively. For Nanopore sequencing, the ligation sequencing kit (SQK-LSK109; Oxford Nanopore Technologies) was used with  $\sim$ 500 ng of input DNA for library construction. The library for *Pseudomonas* sp. 1079 was sequenced with a Flongle adaptor, and the other four strains were loaded together into an R9.4 flow cell with specific barcodes on the Oxford Nanopore MinION device. All base calling was performed with the “high-accuracy” model as integrated in MinKNOW v3.1.13 software.

Quality control of raw reads was performed on the raw sequencing data using LongQC (4). Genome assembly was performed using Flye, which is a long-read *de novo* genome assembly pipeline (5). Polishing was performed using the Burrows-Wheeler Aligner (BWA) v0.7.17 (6) and Racon v1.4.19 (7) with parameters specific for Nanopore

**Citation** Cui Z, Steven B, Zeng Q. 2021. Complete genome sequences of *Curtobacterium*, *Pantoea*, *Erwinia*, and two *Pseudomonas* sp. strains, isolated from apple flower stigmas from Connecticut, USA. *Microbiol Resour Announc* 10:e00154-21. <https://doi.org/10.1128/MRA.00154-21>.

**Editor** David A. Baltrus, University of Arizona  
**Copyright** © 2021 Cui et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Blaire Steven, Blaire.Steven@ct.gov, or Quan Zeng, Quan.Zeng@ct.gov.

**Received** 26 February 2021

**Accepted** 19 April 2021

**Published** 13 May 2021

**TABLE 1** Characteristics and accession numbers for genomes of the apple stigma bacterial isolates

Data for:		<i>Erwinia</i> sp. 18B1	<i>Pantoea</i> sp. 1B4	<i>Pseudomonas</i> sp. 15A4	<i>Pseudomonas</i> sp. 1079
Mean raw sequence length (bp)	7,304	6,911	6,430	7,384	6,627
Total sequences (Gb)	2.17	2.05	2.10	1.37	0.57
Mean coverage (×)	555	384	391	229	91
Flow cell barcode sequences	TCCATCCCTCCGATA GATGAAAC	ACGTAACCTGGTTT GTTCCCTGAA	AAGGATTCATCCCCAC GGTAACAC	TAGGGAACACCGATA GAATCCGAA	<sup>a</sup>
Total sequence length (bp)	3,805,409	5,153,201	5,091,746	5,702,273	5,898,439
No. of scaffolds	1	5	6	1	2
Raw read N <sub>50</sub> (bp)	15,616	13,014	12,702	15,569	14,902
Assembly N <sub>50</sub> (bp)	3,805,409	4,898,912	4,115,195	5,702,273	5,894,865
GC content (%)	70.7	56.4	55.1	60.5	60.2
No. of genes	3,743	5,039	4,823	5,039	5,343
No. of protein-coding genes	3,043	3,661	3,711	4,318	4,389
No. of 16S rRNAs	4	7	7	5	6
No. of tRNAs	47	81	77	64	68
Completeness (%) <sup>b</sup>	87.33	87.18	96.23	98.57	97.27
Contamination (%) <sup>b</sup>	5.08	6.90	0.45	0.03	0.27
Closest relative (ANI [%], 16S rRNA gene identity [%])	<i>Curtobacterium citreum</i> (85.5, 90.75)	<i>Erwinia aphidicola</i> (98.7, 95.00)	<i>Pantoea agglomerans</i> (98.3, 99.87)	<i>Pseudomonas graminis</i> (91.7, 99.93)	<i>Pseudomonas camis</i> (98.5, NA <sup>c</sup> )
Genome GenBank accession no.	CP068987	JAEUXA0000000000	JAEUWZ0000000000	CP068986	JAEUXB0000000000
Raw read SRA no.	SRR13499755– SRR13499826	SRR13499599– SRR13499670	SRR13499518– SRR13499596	SRR13500215– SRR13500259	SRR13499671– SRR13499754

<sup>a</sup> This isolate was sequenced on a Flongle flow cell, while the other strains were barcoded and sequenced on an R9.4 flow cell.<sup>b</sup> Based on marker gene analysis in the CheckM software package.<sup>c</sup> NA, no available 16S rRNA gene sequence in the NCBI database.

read alignment and mapping. The polished output was further processed with Medaka v1.0.3 (<https://github.com/nanoporetech/medaka>). Quality assessment was determined with QUAST (8). The annotation was performed with the NCBI Prokaryotic Genome Annotation Pipeline v4.12 (PGAP) (9), with completeness and contamination checked using CheckM (10). The genome information of these five strains is listed in Table 1. The closest phylogenetic neighbors of these strains were identified by whole-genome average nucleotide identity (ANI) using FastANI v1.32 and 16S rRNA gene homology using blastn with the NCBI database. Default parameters were used for all software unless otherwise specified.

**Data availability.** The data of the raw reads and complete genome sequences of these five strains have been deposited in the SRA and GenBank (Table 1), respectively, under the BioProject accession number [PRJNA693803](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA693803).

## ACKNOWLEDGMENTS

This work was supported by USDA-NIFA-Organic Transitions grant 2017-51106-27001, a Northeastern IPM Center partnership grant, and a USDA specialty crop block grant (SCBG) through the Department of Agriculture, State of Connecticut.

## REFERENCES

1. Cui Z, Huntley RB, Schultes NP, Steven B, Zeng Q. 2021. Inoculation of stigma-colonizing microbes to apple stigmas alters microbiome structure and reduces the occurrence of fire blight disease. *Phytobiomes J* <https://doi.org/10.1094/PBIOMES-04-20-0035-R>.
2. Steven B, Huntley RB, Zeng Q. 2018. The influence of flower anatomy and apple cultivar on the apple flower phytobiome. *Phytobiomes J* 2:171–179. <https://doi.org/10.1094/PBIOMES-03-18-0015-R>.
3. Cui Z, Huntley RB, Zeng Q, Steven B. 2021. Temporal and spatial dynamics in the apple flower microbiome in the presence of the phytopathogen *Erwinia amylovora*. *ISME J* 15:318–329. <https://doi.org/10.1038/s41396-020-00784-y>.
4. Fukasawa Y, Ermini L, Wang H, Carty K, Cheung M-S. 2020. LongQC: a quality control tool for third generation sequencing long read data. *G3 (Bethesda)* 10:1193–1196. <https://doi.org/10.1534/g3.119.400864>.
5. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
6. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 13033997. <https://arxiv.org/abs/1303.3997>.
7. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27:737–746. <https://doi.org/10.1101/gr.214270.116>.
8. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
9. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
10. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.