# Complete Genome Sequence of *Pseudomonas syringae* pv. lapsa Strain ATCC 10859, Isolated from Infected Wheat

Jun Kong,[a] Hongshan Jiang,[b] Baiyun Li,[a] Wenjun Zhao,[b] Zhihong Li,[a] Shuifang Zhu[b]

College of Plant Protection, China Agricultural University, Beijing, China[a]; Institute of Plant Quarantine, Chinese Academy of Inspection and Quarantine, Beijing, China[b]

J.K. and H.J. contributed equally to this work.

*Pseudomonas syringae* pv. lapsa is a pathovar of *Pseudomonas syringae* that can infect wheat. The complete genome of *P. syringae* pv. lapsa strain ATCC 10859 contains a 5,918,899-bp circular chromosome with 4,973 coding sequences, 16 rRNAs, 69 tRNAs, and an average GC content of 59.13%. The analysis of this genome revealed several gene clusters that are related to pathogenesis and virulence.

Address correspondence to Zhihong Li, lizh@cau.edu.cn, or Shuifang Zhu, zhusf@caiq.gov.cn.

*Pseudomonas syringae* is a rod-shaped, polarly-flagellated, Gram-negative, aerobic bacterium, which comprises 64 pathovars with each corresponding to a specific susceptible host species, including tomato, wheat, apple, cocoa, etc. (1). Among these pathovars, *Pseudomonas syringae* pv. lapsa can specifically infect wheat, a widely cultivated food crop, causing yield reduction and economic loss. Though the whole-genome sequence of a model pathogen, i.e., *P. syringae* pv. tomato DC3000, has been published and well studied (2), most of the *P. syringae* strains have their genome sequences unrevealed or uncompleted. To gain a better understanding of the pathogenesis mechanism, as well as the virulence features of *P. syringae*, the whole-genome sequence of the *P. syringae* pv. lapsa, strain ATCC 10859 was determined.

Three paired-end DNA libraries were constructed, with average insert sizes of 380 bp, 490 bp, and 680 bp, respectively, for sequencing with an Illumina MiSeq platform. Meanwhile, a 10-kbp SMARTbell template library was constructed for sequencing in a PacBio RS II single-molecule real-time (SMRT) sequencing platform (3). A total of 4,282,609 (2 × 307-bp paired-end) reads, as well as 19,804 reads with a mean read length of 10,375 bp were obtained, corresponding to sequencing depths above 400× and 30×, respectively. *De novo* assembly of the SMRT reads was performed using the Hierarchical Genome Assembly Process (HGAP) assembly protocol version 3 (4) and the Gepard program (5), which yielded one complete circular chromosome sequence. The Illumina reads were adapter-trimmed using the Skewer program (6) before being used for error correction on the circular sequence. Genome annotations were done by the NCBI Prokaryotic Genome Annotation Pipeline (7). The genome contains a 5,918,899-bp circular chromosome with an average GC content of 59.13%. There are 4,973 coding sequences (CDSs), 16 rRNAs, 63 tRNAs, 1 ncRNA, and 55 pseudogenes.

Many studies have proved that *Pseudomonas* spp. can be used for biocontrol via secrete secondary metabolite (8, 9) and biofilm formation (10, 11). As expected, several secondary metabolism associated gene clusters, such as genes associated with pyrimidine synthesis (ACA40_02470, ACA40_02475, ACA40_10035, ACA40_01095, ACA40_19310, ACA40_20760) and genes associated with benzoate synthesis (ACA40_09900, ACA40_10125, ACA40_24870), were identified in this strain. Meanwhile, 6 fimbrial gene clusters (cup) (ACA40_01475, ACA40_07220, ACA40_09125, ACA40_13630, ACA40_15640, ACA40_22320), which were involved in distinct stages of biofilm formation (12), were identified in this strain. Due to the importance of flagella biosynthesis in the development of biofilm (13), the flagella-associated genes were also checked. In sum, 5 flagellar motor genes (motA:ACA40_02870, ACA40_22280, motB:ACA40_02875, motC:ACA40_16875, motD:ACA40_16870), as well as 4 other flagellar genes (FliY:ACA40_01295,FliL: ACA40_01630, FliK:ACA40_16680, FlhB:ACA40_16685) were found. These results indicate that ATCC 10859 is a potential antagonistic bacterium.

Genome comparison between *P. syringae* pv. lapsa ATCC 10859 and *P. syringae* pv. tomato DC 3000 (2) revealed that about 14.7% (729/4,973) of the CDSs of the first strain cannot find homologs in the latter one. Not including many hypothetical genes annotated in this genome, 43 genes associated with type III secretion system (T3SS) were found. In comparison, there are 82, 64, and 49 T3SS genes annotated with the DC3000, 1448A, and B728a genomes, respectively.

**Nucleotide sequence accession number.** The genome sequence has been deposited at GenBank under the accession number CP013183. Strain ATCC 10859 is available from American Type Culture Collection (ATCC).

## ACKNOWLEDGMENTS

## REFERENCES

1. **Young JM**. 2010. Taxonomy of *Pseudomonas syringae*. J Plant Pathol **92**: 5–24.
2. **Buell CR, Joardar V, Lindeberg M, Selengut J, Paulsen IT, Gwinn ML, Dodson RJ, Deboy RT, Durkin AS, Kolonay JF, Madupu R, Daugherty S, Brinkac L, Beanan MJ, Haft DH, Nelson WC, Davidsen T, Zafar N, Zhou L, Liu J**. 2003. The complete genome sequence of the Arabidopsis and tomato pathogen Pseudomonas syringae pv. tomato DC3000. Proc Natl Acad Sci U S A **100**:10181–10186. http://dx.doi.org/10.1073/pnas.1731982100.
3. **Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M**. 2009. Real-time DNA sequencing from single polymerase molecules. Science **323**:133–138. http://dx.doi.org/10.1126/science.1162986.
4. **Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J**. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods **10**:563–569. http://dx.doi.org/10.1038/nmeth.2474.
5. **Krumsiek J, Arnold R, Rattei T**. 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics **23**:1026–1028.
6. **Jiang H, Lei R, Ding SW, Zhu S**. 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics **15**:182. http://dx.doi.org/10.1186/1471-2105-15-182.
7. **Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W**. 2013. Prokaryotic Genome Annotation Pipeline. The NCBI handbook, 2nd ed. NCBI, Bethesda MD.
8. **Duffy BK, Défago G**. 1997. Zinc improves biocontrol of fusarium crown and root rot of tomato by *Pseudomonas fluorescens* and represses the production of pathogen metabolites inhibitory to bacterial antibiotic biosynthesis. Phytopathology **87**:1250–1257. http://dx.doi.org/10.1094/PHYTO.1997.87.12.1250.
9. **Weller DM**. 2007. *Pseudomonas* biocontrol agents of soilborne pathogens: looking back over 30 years. Phytopathology **97**:250–256. http://dx.doi.org/10.1094/PHYTO-97-2-0250.
10. **Bianciotto V, Andreotti S, Balestrini R, Bonfante P, Perotto S**. 2001. Mucoid mutants of the biocontrol strain *Pseudomonas fluorescens* CHA0 show increased ability in biofilm formation on mycorrhizal and nonmycorrhizal carrot roots. Mol Plant Microbe Interact **14**:255–260. http://dx.doi.org/10.1094/MPMI.2001.14.2.255.
11. **Selin C, Habibian R, Poritsanos N, Athukorala SN, Fernando D, de Kievit TR**. 2010. Phenazines are not essential for *Pseudomonas chlororaphis* PA23 biocontrol of *Sclerotinia sclerotiorum*, but do play a role in biofilm formation. FEMS Microbiol Ecol **71**:73–83. http://dx.doi.org/10.1111/j.1574-6941.2009.00792.x.
12. **Vallet I, Olson JW, Lory S, Lazdunski A, Filloux A**. 2001. The chaperone/usher pathways of Pseudomonas aeruginosa: identification of fimbrial gene clusters (cup) and their involvement in biofilm formation. Proc Natl Acad Sci USA **98**:6911–6916. http://dx.doi.org/10.1073/pnas.111551898.
13. **O'Toole GA, Kolter R**. 1998. *Flagellar and twitching motility are necessary for Pseudomonas aeruginosa* biofilm development. Mol Microbiol **30**: 295–304.