Genome Resource for *Pseudomonas* sp. Strain L22-9: A Potential Novel Species with Antifungal Activity

Jingling Liang, Sai Wang, Ayizekeranmu Yiming, Luoyi Fu, Wenhan Nie, Gongyou Chen, and Bo Zhu[†]

School of Agriculture and Biology, Shanghai Jiao Tong University/Key Laboratory of Urban Agriculture by Ministry of Agriculture of China, Shanghai 200240, China

Abstract

Pseudomonas is a complex genus with increasing numbers of new species. Recently, we isolated *Pseudomonas* sp. strain L22-9, which showed antifungal activity against several fungal phytopathogens. Here, we report the whole genome sequence of strain L22-9. Genomic analysis revealed that strain L22-9 contains one circular DNA chromosome of 6,730,360 bp length with 60.9% GC content. Bioinformatics analysis identified gene clusters in the genome that synthesize antimicrobial metabolites such as 2,4-diacetylphloroglucinol synthesis and hydrogen cyanide synthase. Further analysis suggests that strain L22-9 is a novel species of the genus *Pseudomonas*. This genome would be a valuable resource for future research in phytopathology.

Funding

This work was supported by the National Key R&D Program of China (2018YFD0201202 and 2017YFD0201108), Agri-X Interdisciplinary Fund of Shanghai Jiao Tong University (Agri-X2017010), Shanghai Committee of Science and Technology (19390743300), State Key Laboratory for Biology of Plant Diseases and Insect Pests (SKLOF201802), and Open Project Program of State Key Laboratory of Rice Biology (20190109).

Keywords

biological control, genomics, microbe-genome sequencing

Genome Announcement

Microbes exhibiting biocontrol properties have been of great research interest in recent decades, and the genus *Pseudomonas* is one representative (O'Brien 2017; Syed Ab Rahman et al. 2018; Weller 2007). Many *Pseudomonas* strains are natural inhabitants in the rhizosphere and can suppress phytopathogens and promote plant growth (Mercado-Blanco 2015). The mechanisms used by pseudomonads in their behavior as biocontrol agents include the synthesis of a wide range of antimicrobial metabolites such as pyoluteorin, 2,4-diacetylphloroglucinol (2,4-DAPG), pyrrolnitrin, and hydrogen cyanide (HCN) (Santoyo et al. 2012).

Novel *Pseudomonas* spp. have been continuously identified over the years. Recently, we have isolated *Pseudomonas* sp. strain L22-9 from a cabbage field. The strain was found to be antagonistic against several fungal pathogens in dual-culture assays. The complete genome analysis confirmed the antifungal properties of this strain. Interestingly, phylogenetic analysis and genomic evidence revealed that strain L22-9 is a potential novel species with antifungal properties and, thus, acts as a valuable resource in phytopathology.

The strain was cultivated overnight in 50 ml of Luria-Bertani broth in a shaking incubator at 180 rpm and 28°C. Bacterial cells were harvested by centrifugation; then, genomic DNA was extracted by using the Wizard Genomic DNA Purification Kit (Promega Corp., Madison, WI, U.S.A.). A Qubit 2.0 fluorometer (Life Technologies, Waltham, MA, U.S.A.) and NanoDrop 2000 UV-Vis spectrophotometer (Thermo Scientific, Waltham, MA, U.S.A.) were used to check the quantity and quality, respectively, of the extracted genomic DNA. The high-quality genomic DNA was transferred to Woosen Biotechnology, Hangzhou, China to perform the complete genome sequencing of strain L22-9 using Oxford Nanopore Technologies (ONT). The short reads from

[†]Corresponding author: B. Zhu; bzhu1981@sjtu.edu.cn

Data availability: The complete genome sequence of strain L22-9 has been deposited in GenBank under accession number CP048810.

The author(s) declare no conflict of interest.

Accepted for publication 23 July 2020.

TABLE 1. Genome statistics of *Pseudomonas* sp. strain L22-9

Feature	Chromosome
Genome size (bp)	6,730,360
G+C content (%)	60.9
Total genes	5,963
Coding sequences	5,877
Protein coding genes	5,661
RNAs	86
tRNAs	66
Ribosomal RNAs	16
ncRNAs	4
Pseudogenes	216
Genes with predicted functions	5,157
Annotated genes (databases)	
Clusters of Orthologous Group	5,157 (91%)
Gene Ontology	1,369 (24%)
Kyoto Encyclopedia of Genes and Genomes	3,669 (65%)
Carbohydrate-Active enZymes	61 (1.1%)
GenBank accession	CP048810

second-generation sequencing technology (Illumina Hiseq 2500) was also obtained to counter the error-prone long reads produced by ONT (Koren et al. 2012).

The ONT sequencing yielded 911.44 Mb of clean data (sequence depth $\geq 100\times$) comprising 89,632 reads with a mean read length of 10,168 bp and the N_{50} read length of 14,408 bp. The Illumina platform generated 5,821,728 reads and quality control was performed by FastQC (version 0.11.8; Babraham Bioinformatics, Cambridge, U.K.). The genome was assembled using Canu, version 1.8, including three steps: self-correction, trimming, and assembly (Koren et al. 2017). The draft genome assembly was then polished by Pilon, version 1.22 (Walker et al. 2014), with the second-generation sequencing data. BUSCO, version 4.1.1 (Seppey et al. 2019) was used to assess the completeness of the genome. The completeness of the genome assembly was estimated to be 99.2%.

The assembled genome was annotated by NCBI Prokaryotic Genomes Annotation Pipeline (Tatusova et al. 2016) while being submitted to GenBank. Functional analysis of the protein coding genes was performed by eggNOG-mapper v2 (Huerta-Cepas et al. 2017), which could retrieve annotation from Clusters of Orthologous Groups (COGs), Gene Ontology, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Carbohydrate-Active enZymes databases at the same time. Genes involved in secondary metabolite biosynthesis were identified using antiSMASH, version 5.1.2 (Blin et al. 2019). The average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) values between strain L22-9 and its closely related strains were calculated on ANI Calculator (Yoon et al. 2017) and Genome-to-Genome Distance Calculator 2.1 (Meier-Kolthoff et al. 2013), respectively.

Genome assembly showed that strain L22-9 has a single circular DNA chromosome with a length of 6,730,360 bp, and no plasmid was found in the genome. The overall genome features are displayed in Table 1. The genome of strain L22-9 was compared with the type strains of other *Pseudomonas* spp. A phylogenetic tree based on concatenated sequences of 16S ribosomal RNA, *gyrB*, *rpoB*, and *rpoD* genes of strain L22-9 and the representatives of related taxa showed that strain L22-9, *Pseudomonas kilonensis* DSM 13647^T, *P. brassicacearum* LMG 21623^T, *P. thivervalensis* DSM 13194^T, *P. mediterranea* CFBP 5447^T, and *P. corrugata* NCPPB 2445^T group together into a small clade (Fig. 1). The ANI values between strain L22-9 and these type strains ranged from 86.76 to 91.87%, and the dDDH values ranged from 32.6 to 46.3%. The highest ANI value (91.87%) and dDDH value (46.3%) were generated on comparison of studied strain L22-9 to reference strain *P. kilonensis* DSM 13647^T. These values are well below the threshold ANI value of 95% and dDDH value of 70%, respectively, within species, suggesting that strain L22-9 is a novel species of the genus *Pseudomonas*.

According to COG analysis, most of the coding genes (2,134 genes, 41%) were involved in metabolism, of which 550 genes (26%) related to amino acid transport and metabolism. The second most abundant coding genes (1,099 genes, 21%) annotated with “unknown functions”. The abundance of the poorly characterized genes implies that some unique features may exist in the genome of strain L22-9. In KEGG mapping results, genes were classified into six KEGG functional categories; namely, metabolism (78.09%), genetic information processing (5.53%),

Fig. 1. Phylogenetic tree based on concatenated sequences of 16S ribosomal RNA, *gyrB*, *rpoB*, and *rpoD* genes of strain L22-9 and the representatives of related taxa. Bootstrap values are expressed as a percentage of 1,000 replicates and only those higher than 70% are given at the branch points. The scale indicates the nucleotide substitutions per site. Species names are followed by the strain number and the GenBank accession number. *Pseudomonas stutzeri* ATCC 17588^T is selected as outgroup (^T indicates type strain).

environmental information processing (6.90%), cellular processes (5.29%), organismal systems (1.64%), and human diseases (3.95%).

Some secondary metabolites play a vital role in the suppression of plant pathogens by antagonistic microorganisms. antiSMASH software identified a gene cluster belonging to type III polyketide synthase and it showed 100% similarity with the gene cluster that exhibited 2,4-DAPG synthesis (Bangera and Thomashow 1999). In addition, genes encoding HCN synthase (*hcnA*, *hcnB*, and *hcnC*) (Laville et al. 1998) were also found in the genome of strain L22-9. This genomic information provided evidence for the antifungal activity of strain L22-9.

Literature Cited

- Bangera, M. G., and Thomashow, L. S. 1999. Identification and characterization of a gene cluster for synthesis of the polyketide antibiotic 2, 4-diacetylphloroglucinol from *Pseudomonas fluorescens* Q2-87. *J. Bacteriol.* 181:3155-3163.
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H., and Weber, T. 2019. antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47:W81-W87.
- Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., Von Mering, C., and Bork, P. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* 34:2115-2122.
- Koren, S., Schatz, M. C., Walenz, B. P., Martin, J., Howard, J. T., Ganapathy, G., Wang, Z., Rasko, D. A., McCombie, W. R., and Jarvis, E. D. 2012. Hybrid error correction and *de novo* assembly of single-molecule sequencing reads. *Nat. Biotechnol.* 30:693-700.
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. 2017. Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27:722-736.
- Laville, J., Blumer, C., Von Schroetter, C., Gaia, V., Dfago, G., Keel, C., and Haas, D. 1998. Characterization of the *hcnABC* gene cluster encoding hydrogen cyanide synthase and anaerobic regulation by ANR in the strictly aerobic biocontrol agent *Pseudomonas fluorescens* CHA0. *J. Bacteriol.* 180:3187-3196.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.-P., and Gker, M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinf.* 14:60.
- Mercado-Blanco, J. 2015. *Pseudomonas* strains that exert biocontrol of plant pathogens. Pages 121-172 in: *Pseudomonas*: Volume 7: New Aspects of *Pseudomonas* Biology. J.-L. Ramos, J. B. Goldberg, and A. Filloux, eds. Springer Netherlands, Dordrecht, The Netherlands.
- O'Brien, P. A. 2017. Biological control of plant diseases. *Australas. Plant Pathol.* 46: 293-304.
- Santoyo, G., Orozco-Mosqueda, M. C., and Govindappa, M. 2012. Mechanisms of biocontrol and plant growth-promoting activity in soil bacterial species of *Bacillus* and *Pseudomonas*: A review. *Biocontrol Sci. Technol.* 22:855-872.
- Seppy, M., Manni, M., and Zdobnov, E. M. 2019. BUSCO: Assessing genome assembly and annotation completeness. Pages 227-245 in: *Gene Prediction*. M. Kollmar, ed. Methods in Molecular Biology, Vol. 1962. Humana, New York, NY, U.S.A.

- Syed Ab Rahman, S. F., Singh, E., Pieterse, C. M., and Schenk, P. M. 2018. Emerging microbial biocontrol strategies for plant pathogens. *Plant Sci.* 267:102-111.
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetverin, V., Nawrocki, E. P., Zaslavsky, L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., and Ostell, J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44:6614-6624.
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., and Young, S. K. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963.
- Weller, D. M. 2007. *Pseudomonas* biocontrol agents of soilborne pathogens: Looking back over 30 years. *Phytopathology* 97:250-256.
- Yoon, S.-H., Ha, S., Lim, J., Kwon, S., and Chun, J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Leeuwenhoek* 110: 1281-1286.