High-Quality Genome Resource of Xanthomonas hyacinthi Generated via Long-Read Sequencing

Stephen P. Cohen, Emily K. Luna, Jillian M. Lang, Janet Ziegle, Christine Chang, Jan E. Leach, Marion Fischer-Le Saux, Perrine Portier, Ralf Koebnik, Jonathan M. Jacobs

RESOURCE ANNOUNCEMENT

High-Quality Genome Resource of *Xanthomonas hyacinthi* Generated via Long-Read Sequencing

e-Xtra*

Stephen P. Cohen,[1] Emily K. Luna,[2] Jillian M. Lang,[2] Janet Ziegle,[3] Christine Chang,[3] Jan E. Leach,[2] Marion Fischer Le-Saux,[4,5] Perrine Portier,[4,5] Ralf Koebnik,[6] and Jonathan M. Jacobs[1,7,†]

[1] Department of Plant Pathology, The Ohio State University, Columbus, OH 43210, U.S.A.
[2] Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80523, U.S.A.
[3] Pacific Biosciences, Menlo Park, CA 94025, U.S.A.
[4] IRHS, Institut National de la Recherche Agronomique (INRA), Université d'Angers, Agrocampus-Ouest, SFR 4207 QuaSav, 49071, Beaucouzé, France
[5] CIRM-CFBP French Collection for Plant-associated Bacteria, IRHS, UMR 1345, INRA-ACO-UA, 42 rue Georges Morel, 49070, Beaucouzé, France
[6] Institut de Recherche pour le Développement (IRD), Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), Université Montpellier, IPME, 34000 Montpellier, France
[7] Infectious Disease Institute, The Ohio State University, Columbus, OH 43210, U.S.A.

Abstract

The bacterial plant pathogen *Xanthomonas hyacinthi* is the causal agent of yellow disease of *Hyacinthus* and other ornamental plant genera. There is no available complete genome for *X. hyacinthi*, limiting basic research for this pathogen. Here, we release a high-quality complete genome sequence for the *X. hyacinthi* type strain, CFBP 1156. Single-molecule real-time (SMRT) sequencing with a mean coverage of 306× revealed two contigs of 4,918,645 and 44,381 bp in size. This was the first characterized plant-disease-causing species of *Xanthomonas* and this genome provides a resource to better understand the biology of yellow disease of hyacinth.

Funding

Support was provided by the following sources: a United States Department of Agriculture National Institute of Food and Agriculture (NIFA) Postdoctoral Fellowship (2018-08122 to S. P. Cohen) and a joint National Science Foundation-NIFA PBI grant (2018-05040 to J. M. Lang, J. E. Leach, and J. M. Jacobs). This article is also based upon work from European Cooperation in Science and Technology (COST) Action CA16107 EuroXanth, supported by COST and a Fulbright Scholar Award to Belgium (J. M. Jacobs). This research was also supported by the collaborative French Xanthomonas Network with funds from CIRAD, INRA, and IRD.

Keywords

hyacinth, SMRT sequencing, *Xanthomonas hyacinthi*, yellow disease

Resource Announcement

Xanthomonas hyacinthi is a plant-pathogenic bacteria that causes yellow disease in *Hyacinthus*, *Scilla*, and other related ornamental plant genera (Janse and Miller 1983; Van Tuyl and Toxopeus 1980). *X. hyacinthi* was first isolated in 1883 by J. H. Wakker, who was among the first plant pathologists to recognize bacteria as a possible cause of plant disease (Beijer 1966; Wakker 1883). Because of its ability to cause severe symptoms, *X. hyacinthi* threatens production of ornamental bulbous plants. The type strain is available in culture collections as CFBP 1156, LMG 739, NCPPB 599, ICMP 189, ATCC 19314, and DSM 19077 (Vauterin et al. 1995). A low-quality draft genome of *X. hyacinthi* strain DSM 19077 was provided in 2015, with 15× coverage Illumina HiSeq data assembled into 1,415 contigs (Naushad et al. 2015). This draft and other incomplete genomic information allowed researchers to previously develop nucleic acid-based diagnostic methods for detecting *X. hyacinthi* (Back et al. 2015; van Doorn et al. 2001). A newer draft genome was recently generated with Illumina HiSeq data, with 104 contigs assembled representing 100× coverage (NCBI BioProject Accession PRJNA338630) (Merda et al. 2017). However, short-read

†Corresponding author: J. M. Jacobs; jacobs.1080@osu.edu

*The *e*-Xtra logo stands for "electronic extra" and indicates that one supplementary table is published online.

The author(s) declare no conflict of interest.

Accepted for publication 9 December 2019.

sequencing is not sufficient for assembly of *Xanthomonas* transcription activator-like effector (TALE) genes (Peng et al. 2016). Here, we report the complete genome sequence for *X. hyacinthi* strain CFBP 1156 generated by long-read sequencing.

X. hyacinthi strain CFBP 1156 was isolated from *Hyacinthus orientalis* in the Netherlands in 1958 and deposited in the French Collection for Plant-associated Bacteria (CIRM-CFBP) in 1969. From lyophilized stock preserved at CIRM-CFBP, a single colony was isolated and grown on rich nutrient agar media. Genomic DNA was extracted using the Genomic DNA buffer set and Genomic-tips following the manufacturer's instructions (Qiagen, Valencia, CA, U.S.A.). DNA was sheared three times with a Covaris g-TUBE at 5,500 rpm for 2 min to 8- to 30-kb fragments and sequenced using long-read single-molecule real-time (SMRT) sequencing on a PacBio Sequel I, and the genome was assembled with HGAP v4 (Pacific Biosciences, Menlo Park, CA, U.S.A.). QUAST 5.0.2 was used to assess genome quality (Gurevich et al. 2013). NCBI Prokaryotic Genome Annotation Pipeline version 4.9 was used for functional annotation of genes (Haft et al. 2018). SignalP 5.0 was used to identify genes with predicted signal peptides (Almagro Armenteros et al. 2019). AnnoTALE 1.4.1 was used to detect and annotate TALE genes (Grau et al. 2016).

Mean coverage for the *X. hyacinthi* strain CFBP 1156 genome was 306×. The genome was assembled into two contigs of 4,918,645 and 44,381 bp in size—a chromosome and a plasmid, respectively—with a total G+C content of 68.0%. The N_{50} and L_{50} were 4,918,645 and 1, respectively. Of the 4,392 genes predicted, 4,011 encoded proteins, 63 encoded functional RNAs, and 318 were considered pseudogenes. Genes encoding predicted secretion system (SS) proteins were identified, including T1SS, T2SS, T3SS, and T4SS. In all, 10 genes were annotated as type III effectors or avirulence genes, including one locus (FZ025_20815) encoding a TALE previously undetected in draft genomes. The TALE was assigned to a novel class by AnnoTALE and designated as TalHQ1. Signal peptides for Sec/SPI, Sec/SPII, and Tat/SPI were detected in 496, 87, and 207 genes, respectively (Supplementary Table S1).

Raw reads and the complete genome were uploaded to NCBI Sequence Read Archive and GenBank under BioProject accession PRJNA562936.

Literature Cited

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. 37:420-423.

Back, C.-G., Lee, S.-Y., Lee, B.-J., Yea, M.-C., Kim, S.-M., Kang, I.-K., Cha, J.-S., and Jung, H.-Y. 2015. Development of a species-specific PCR assay for three *Xanthomonas* species, causing bulb and flower diseases, based on their genome sequences. Plant Pathol. J. 31:212-218.

Beijer, J. J. 1966. Dr. J. H. Wakker (1859–1927). Neth. J. Plant Pathol. 72:38-45.

Grau, J., Reschke, M., Erkes, A., Streubel, J., Morgan, R. D., Wilson, G. G., Koebnik, R., and Boch, J. 2016. AnnoTALE: Bioinformatics tools for identification, annotation, and nomenclature of TALEs from *Xanthomonas* genomic sequences. Sci. Rep. 6:21077.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072-1075.

Haft, D. H., DiCuccio, M., Badretdin, A., Brover, V., Chetvernin, V., O'Neill, K., Li, W., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Zheng, C., Thibaud-Nissen, F., Geer, L. Y., Marchler-Bauer, A., and Pruitt, K. D. 2018. RefSeq: An update on prokaryotic genome annotation and curation. Nucleic Acids Res. 46: D851-D860.

Janse, J. D., and Miller, H. J. 1983. Yellow disease in *Scilla tubergeniana* and related bulbs caused by *Xanthomonas campestris* pv. *hyacinthi*. Neth. J. Plant Pathol. 89:203-206.

Merda, D., Briand, M., Bosis, E., Rousseau, C., Portier, P., Barret, M., Jacques, M.-A., and Fischer-Le Saux, M. 2017. Ancestral acquisitions, gene flow and multiple evolutionary trajectories of the type three secretion system and effectors in *Xanthomonas* plant pathogens. Mol. Ecol. 26:5939-5952.

Naushad, S., Adeolu, M., Wong, S., Sohail, M., Schellhorn, H. E., and Gupta, R. S. 2015. A phylogenomic and molecular marker based taxonomic framework for the order *Xanthomonadales*: Proposal to transfer the families *Algiphilaceae* and *Solimonadaceae* to the order *Nevskiales* ord. nov. and to create a new family within the order *Xanthomonadales*, the family *Rhodanobacteraceae* fam. nov., containing the genus *Rhodanobacter* and its closest relatives. Antonie Leeuwenhoek 107:467-485.

Peng, Z., Hu, Y., Xie, J., Potnis, N., Akhunova, A., Jones, J., Liu, Z., White, F. F., and Liu, S. 2016. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. BMC Genomics 17:21.

van Doorn, J., Hollinger, T. C., and Oudega, B. 2001. Analysis of the type IV fimbrial-subunit gene *fimA* of *Xanthomonas hyacinthi*: Application in PCR-mediated detection of yellow disease in hyacinths. Appl. Environ. Microbiol. 67:598-607.

Van Tuyl, J. M., and Toxopeus, S. J. 1980. Breeding for resistance to yellow disease of hyacinths. I. Investigations on F1's from diallel crosses. Euphytica 29:555-560.

Vauterin, L., Hoste, B., Kersters, K., and Swings, J. 1995. Reclassification of *Xanthomonas*. Int. J. Syst. Evol. Microbiol. 45:472-489.

Wakker, J. 1883. Vorläufige Mittheilungen über Hyacinthenkrankheiten. Bot. Centralbl. 14:315-317.