



Complete Genome Sequences of Walnut-Associated *Xanthomonas euroxantha* Strains CPBF 367 and CPBF 426 Obtained by Illumina/Nanopore Hybrid Assembly

Miguel Teixeira,^{a,b} Leonor Martins,^{a,b} Camila Fernandes,^c Cátia Chaves,^a Joana Pinto,^a  Fernando Tavares,^{a,b}  Nuno A. Fonseca^a

^aCIBIO-Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO-Laboratório Associado, Universidade do Porto, Vairão, Portugal

^bFCUP-Faculdade de Ciências, Departamento de Biologia, Universidade do Porto, Porto, Portugal

^cINIAV, Instituto Nacional de Investigação Agrária e Veterinária, Unidade Estratégica de Investigação e Serviços de Sistemas Agrários e Florestais e Sanidade Vegetal, Oeiras, Portugal

ABSTRACT We present the complete genome sequences of two *Xanthomonas euroxantha* strains isolated from buds of a walnut tree. The whole-genome sequences of strains CPBF 367 and CPBF 426 consist of two circular chromosomes of 4,923,218 bp and 4,883,254 bp and two putative plasmids of 45,241 bp and 17,394 bp, respectively. These data may contribute to the understanding of *Xanthomonas* species-specific adaptations to walnut.

Xanthomonas is a genetically diverse genus of bacteria comprising etiological agents of several plant diseases affecting crops worldwide and causing large yield losses and negative economic impacts (1–5). *Xanthomonas arboricola* is a species known to cause pathologies in several fruit trees, including walnut trees (*Juglans regia* L.). The occurrence of multiple *Xanthomonas* lineages within the same walnut host has revealed a consortium of walnut-associated *Xanthomonas* bacteria distinct from *Xanthomonas arboricola* pv. *juglandis* (5). More recently, based on phenotypic and genomic analyses, such as average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH), it has been proposed that some of these bacteria are members of a new species designated *Xanthomonas euroxantha*, which includes both pathogenic and nonpathogenic strains on walnut (6). Here, we sequenced the whole genome of two strains belonging to this new *Xanthomonas* species, CPBF 367 (LMG 31036=CCOS 1890) and CPBF 426 (LMG 31038=CCOS1892). These genomes should contribute to unveiling the ecology, evolution, and virulence of strains belonging to this new *Xanthomonas* species and disclose niche-specific adaptations.

Both strains were isolated in April 2016 in Loures, Portugal, from asymptomatic dormant buds of a walnut tree known to develop symptoms of walnut bacterial blight during the growing season. Characteristic mucoid yellow colonies were grown on yeast dextrose carbonate (YDC) agar medium, and pure colonies were further cultured in liquid medium from which an aliquot was cryopreserved at -80° in YDC with 30% glycerol as previously detailed (7).

DNA extraction for sequencing was carried out from cryopreserved bacteria recovered on bacterial culture medium M2 (yeast extract, 2 g liter⁻¹; Bacto peptone, 5 g liter⁻¹; NaCl, 5 g liter⁻¹; KH₂PO₄, 0.45 g liter⁻¹; Na₂HPO₄ 12H₂O, 2.39 g liter⁻¹) at 28°C and 100 rpm for 48 h. DNA was extracted using the E.Z.N.A. bacterial DNA purification kit (Omega Bio-tek, Norcross, GA) and sequenced with Illumina and Oxford Nanopore Technologies (ONT) MinION platforms. Illumina sequencing was outsourced to GATC Biotech, AG (Konstanz, Germany), using an Illumina HiSeq instrument with a standard 2 × 150-bp paired-end library protocol. This resulted in 8,413,466 and 6,494,807 reads,

Citation Teixeira M, Martins L, Fernandes C, Chaves C, Pinto J, Tavares F, Fonseca NA. 2020. Complete genome sequences of walnut-associated *Xanthomonas euroxantha* strains CPBF 367 and CPBF 426 obtained by Illumina/Nanopore hybrid assembly. Microbiol Resour Announc 9:e00902-20. <https://doi.org/10.1128/MRA.00902-20>.

Editor David A. Baltus, University of Arizona

Copyright © 2020 Teixeira 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 Fernando Tavares, ftavares@fc.up.pt.

Received 1 August 2020

Accepted 12 October 2020

Published 5 November 2020

TABLE 1 Summary of sequencing data and genome statistics

Characteristic	Data for strain:	
	CPBF 367	CPBF 426
ENA/GenBank accession number	GCA_903989455	GCA_903989465
No. of ONT reads	18,857	19,997
ONT read length (nt) (mean \pm SD)	6,386 \pm 8,299	5,895 \pm 6,300
No. of Illumina reads	8,413,466	6,494,807
Illumina read length (nt)	2 \times 150	2 \times 150
Mean coverage (\times)	532	415
G+C content (%)	65.81	65.85
N_{50} (bp)	4,923,218	4,883,254
No. of sequences	2	2
Genome size (bp)	4,968,459	4,900,648
Chromosome size (bp)	4,923,218	4,883,254
Plasmid size (bp)	45,241	17,394
No. of genes (total)	4,174	4,157
No. of CDSs (total)	4,094	4,080
No. of genes (coding)	4,014	4,010
No. of CDSs (with protein)	4,014	4,010
No. of genes (RNA)	80	77
No. of rRNAs (5S, 16S, 23S)	2, 2, 2	2, 2, 2
No. of complete rRNAs (5S, 16S, 23S)	2, 2, 2	2, 2, 2
No. of tRNAs	56	53
No. of ncRNAs	18	18
No. of pseudogenes	80	70
No. of CDSs (without protein)	80	70

with 506 \times and 394 \times coverage for CPBF 367 and CPBF 426, respectively. For Nanopore sequencing, libraries were prepared with the SQK-LSK109 kit and multiplexed using the EXP-NBD104 barcoding kit. Sequencing was performed on a MinION sequencer using an R9.4.1 flow cell. Reads were base called and demultiplexed using Guppy v3.4.1 (high accuracy base-calling mode) and produced for CPBF 367 (18,857 reads with a mean length of 6,386 nucleotides [nt] [\pm 8,299] and a maximum read length of 92,631 nt) and CPBF 426 (19,997 reads with a mean length of 5,895 nt [\pm 6,300] and a maximum read length of 74,948 nt). The reads were assembled *de novo* following a hybrid Nanopore-Illumina approach using Unicycler v0.4.8 (8) and annotated with PGAP v2020-03-30.build4489 (9). Assembly metrics were calculated using QUAST v5.0.2 (10). The assembly quality and completeness were assessed using BUSCO v4.0.6 (parameters: lineage xanthomonadales_odb10) (11), which indicated that the CPBF 367 and CPBF 426 genomes are both >99.8% complete. Default parameters were used for all software unless otherwise noted.

The assembly of CPBF 367 has a total length of 4,968 Mb, comprising 2 circular contigs of 4,923,218 and 45,241 bp with a G+C content of 65.81%. The annotation indicates 4,014 coding genes in a total of 4,174 genes from 4,094 coding DNA sequences (CDSs), with 2 complete sets of rRNAs (5S, 16S, and 23S rRNAs), 56 tRNAs, 18 noncoding RNAs (ncRNAs), and 80 pseudogenes. For CPBF 426, the assembly has a total length of 4,900 Mb, comprising 2 circular contigs of 4,883,254 and 17,394 bp with a G+C content of 65.85%. The annotation revealed 4,010 coding genes in 4,157 genes from 4,080 CDSs, with 2 complete sets of rRNAs, 53 tRNAs, 18 ncRNAs, and 70 pseudogenes. Table 1 contains a summary of the two complete genome sequences.

OrthoANI v1.40 (12) indicates an average nucleotide identity value higher than 97%, assigning these strains to *X. euroxanthea*, together with a previously reported strain, CPBF 424 (13, 14). Interestingly, while *X. euroxanthea* CPBF 424 holds an apparently functional T3SS and was shown to be pathogenic in walnut plantlets, *X. euroxanthea* strains CPBF 367 and CPBF 426 were deficient for most T3SS operon genes, suggesting that these strains are nonpathogenic in walnut, which was confirmed in walnut

pathogenicity assays for strain CPBF 367 (6). Altogether, the distinct genomic repertoires of these *X. euroxanthea* strains (CPBF 367 and CPBF 426), in comparison to those of *X. arboricola* pv. *juglandis* strains, will provide valuable genomic data for a better understanding of the pathogenicity and virulence of walnut-associated *Xanthomonas*.

Data availability. The raw data and assembled/annotated sequences have been deposited in the European Nucleotide Archive (ENA). The study accession number is [PRJEB39139](https://doi.org/10.1093/nar/gkx569). Illumina reads, MinION reads, and the assembled genome are accessible under the accession numbers [ERX2780809](https://doi.org/10.1093/nar/gkx569), [ERX4296808](https://doi.org/10.1093/nar/gkx569), and [GCA_903989455](https://doi.org/10.1093/nar/gkx569), respectively, for CPBF 367 and [ERX2780811](https://doi.org/10.1093/nar/gkx569), [ERX4296809](https://doi.org/10.1093/nar/gkx569), and [GCA_903989465](https://doi.org/10.1093/nar/gkx569), respectively, for CPBF 426. Supplementary data are accessible via the BioStudies database with the accession number [S-BSST433](https://doi.org/10.1093/nar/gkx569).

ACKNOWLEDGMENTS

Funding for this work was provided by the EU Horizon 2020 Research and Innovation program (grant agreement 668981), European Structural & Investment Funds (ESIFs) through COMPETE 2020, National Funds through FCT (project EVOXANT-PTDC/BIA-EVF/3635/2014-POCI-01-0145-FEDER-016600), and the Operational Thematic Program for Competitiveness and Internationalization (POCI) under the PORTUGAL 2020 Partnership Agreement through the European Regional Development Fund (FEDER) project PORBIOTA-POCI-01-0145-FEDER-022127. Leonor Martins was supported by a Ph.D. fellowship from FCT (SFRH/BD/137079/2018).

REFERENCES

- Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: progress toward understanding and management. *Plant Dis* 98:1600–1610. <https://doi.org/10.1094/PDIS-08-14-0831-FE>.
- Moragrega C, Özaktan H. 2010. Apical necrosis of Persian (English) walnut (*Juglans regia*): an update. *J Plant Pathol* 92:S1.67–S1.71.
- Frutos D. 2010. Bacterial diseases of walnut and hazelnut and genetic resources. *J Plant Pathol* 92:S1.79–S1.85.
- Hajri A, Meyer D, Delort F, Guillaumès J, Brin C, Manceau C. 2010. Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant Pathol* 59:1014–1022. <https://doi.org/10.1111/j.1365-3059.2010.02362.x>.
- Fernandes C, Albuquerque P, Cruz L, Tavares F. 2018. Genotyping and epidemiological metadata provides new insights into population structure of *Xanthomonas* isolated from walnut trees. *bioRxiv* <https://doi.org/10.1101/397703>.
- Martins L, Fernandes C, Blom J, Dia NC, Pothier JF, Tavares F. 2020. *Xanthomonas euroxanthea* sp. nov., a new xanthomonad species including pathogenic and non-pathogenic strains of walnut. *Int J Syst Evol Microbiol* In press. <https://doi.org/10.1099/ijsem.0.004386>.
- Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Dis* 101:858–865. <https://doi.org/10.1094/PDIS-10-16-1481-RE>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- 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>.
- 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>.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
- Lee I, Kim YO, Park SC, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103. <https://doi.org/10.1099/ijsem.0.000760>.
- Konstantinidis KT, Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* 102:2567–2572. <https://doi.org/10.1073/pnas.0409727102>.
- Fernandes C, Blom J, Pothier JF, Tavares F. 2018. High-quality draft genome sequence of *Xanthomonas* sp. strain CPBF 424, a walnut-pathogenic strain with atypical features. *Microbiol Resour Announc* 7:e00921-18. <https://doi.org/10.1128/MRA.00921-18>.