AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

# Draft Whole-Genome Sequences of *Xylella fastidiosa* subsp. *fastidiosa* Strains TPD3 and TPD4, Isolated from Grapevines in Hou-li, Taiwan

**Andreina I. Castillo,[a] Shu-Jen Tuan,[b] Adam C. Retchless,[a]* Fei-Ting Hu,[b] Hsun-Yin Chang,[b] Rodrigo P. P. Almeida[a]**

[a]Department of Environmental Science, Policy and Management, University of California, Berkeley, Berkeley, California, USA
[b]Department of Entomology, National Chung Hsing University, Taichung, Taiwan, Republic of China

**ABSTRACT** We report the draft assemblies of TPD3 and TPD4, two *Xylella fastidiosa* subsp. *fastidiosa* isolates infecting grapevines in Hou-li, Taiwan. TPD3 and TPD4 showed similar characteristics regarding genome size (2,483,503 bp and 2,491,539 bp, respectively), GC content (51.49% and 51.47%, respectively), and number of protein-coding sequences (2,394 and 2,413, respectively).

*Xylella fastidiosa* is a xylem-limited plant-pathogenic bacterium that causes disease in crops and in ornamental and shade tree species (1). *X. fastidiosa* is transmitted by xylem sap-feeding insects of the Cicadellidae (sharpshooter leafhopper) and Cercopoidea (spittlebugs) families (2, 3). Disease symptoms associated with *X. fastidiosa* infection in grapevines were first described in Los Angeles, CA, by Newton Pierce in 1880 and have been subsequently referred to as Pierce's disease (PD) of grapevine (4). The characteristic symptoms of PD include leaf scorching, gradual leaf chlorosis, and shriveling of grapes. Each year, *X. fastidiosa* infections cost $56.1 million in production losses, and $48.3 million in prevention costs are taken by nurseries and government agencies in the state of California (5). The symptoms described in California crops have also been reported on grapevines (*Vitis vinifera* L.) in central Taiwan in 2002 (3, 6). Sequencing of isolates obtained from infected grapevines in Taiwan showed that they share identical 16S rRNA sequences with *X. fastidiosa* PD strains from the Americas and are distantly related to *Xylella taiwanensis* (7). The results suggested that *X. fastidiosa* was imported to Asia from the Americas. Likewise, there have been multiple introductions of *X. fastidiosa* from the Americas to Europe (8, 9), as well as introductions of different subspecies within the American continent (10, 11). The expanding distribution and host range of *X. fastidiosa* bring forward relevant questions, mainly, what factors drive successful host infection and induction of symptoms in certain plant types, and how can better control, detection, and management strategies be developed? Since these factors have a genotypic component, whole-genome sequencing and subsequent analyses are expected to be helpful tools in answering these questions.

Two isolates were obtained from symptomatic grapevines in Hou-li, Taiwan (TPD3, 24°18′57.40″N, 120°41′53.30″E, and TPD4, 24°19′52.40″N, 120°42′03.90″E) in 2012. Petioles from symptomatic plants were wiped with 70% ethanol, sterilized by 0.6% NaOCl, and finally rinsed with sterile reverse-osmosis water (3). The sterile petioles were minced in 1 ml of PD2 broth (containing the following in g/liter: tryptone [4.0], soytone or phytone [2.0], trisodium citrate [1.0], disodium succinate [1.0], hemin chloride [0.01], $MgSO_4 \cdot 7H_2O$ [1.0], $KH_2PO_4$ [1.0], $K_2HPO_4$ [1.5], Bacto-agar [15.0], and bovine serum albumin fraction five [2.0]) using a razor blade in a petri dish. Samples were then grown in periwinkle wilt modified (PWG; containing phytone peptone [4.0 g], Trypticase peptone [1.0 g], $KzHPO_4$, [1.2 g], hemin chloride stock [0.1% bovine heroin chloride in 0.05 N NaOH, 10 ml], $KH_2PO_4$ [1.0 g], GELRITE [9 g], $MgSO_4 \cdot 7H_2O$, [0.4 g], phenol red

**Citation** Castillo AI, Tuan S-J, Retchless AC, Hu F-T, Chang H-Y, Almeida RPP. 2019. Draft whole-genome sequences of *Xylella fastidiosa* subsp. *fastidiosa* strains TPD3 and TPD4, isolated from grapevines in Hou-li, Taiwan. Microbiol Resour Announc 8:e00835-19. https://doi.org/10.1128/MRA.00835-19.

**Editor** Kenneth M. Stedman, Portland State University



Address correspondence to Rodrigo P. P. Almeida, rodrigoalmeida@berkeley.edu.

* Present address: Adam C. Retchless, Meningitis and Vaccine Preventable Diseases Branch, Division of Bacterial Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA.

A.I.C. and S.-J.T. contributed equally to this work.

**Received** 7 August 2019
**Accepted** 24 October 2019
**Published** 21 November 2019

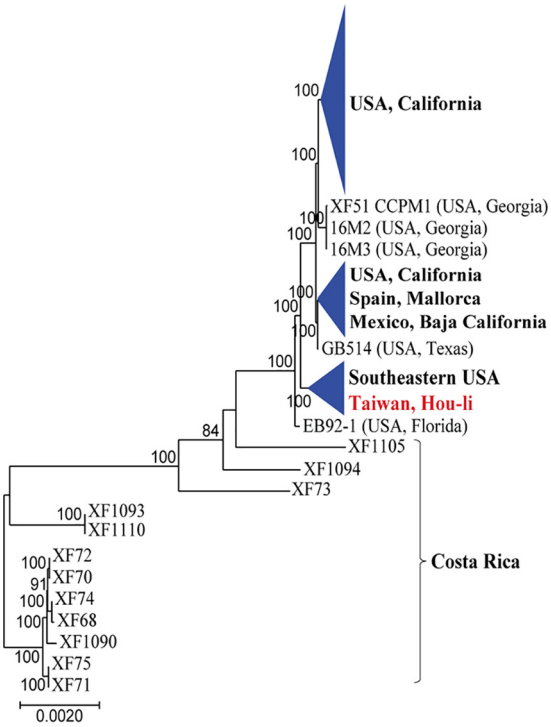


**FIG 1** Phylogeny build using the core-genome alignment for worldwide *X. fastidiosa* subsp. *fastidiosa* isolates. Both samples from Hou-li, Taiwan (red), cluster within the clade, including samples from the southeastern United States, suggesting that they descend from this population.

stock [0.2% phenol red in distilled water, 10 ml], glutamine stock [8.0% {wt/vol}, 50 ml], and bovine serum albumin fraction-five stock [20% {wt/vol}, 15 ml]) medium plates, and isolation of a pure clone was done using a streak plate technique done three times in consecutive petri dishes. Subsequently, DNA was extracted using the Qiagen DNeasy blood and tissue kit (spin columns) using the Gram-negative bacterial DNA extraction protocol. The DNA provided for Illumina library preparation was at a concentration of 10.06 ng/$\mu$l and an $A_{260}/A_{280}$ purity ratio of 1.55 for TPD3. In the case of TPD4, the DNA concentration was 9.67 ng/$\mu$l, and the $A_{260}/A_{280}$ purity ratio was 1.42. Fragments were selected to be 400 bp for both TPD3 and TPD4. Isolates were sequenced using an Illumina HiSeq 2000 platform at the University of California, Berkeley Vincent J. Coates Genomics Sequencing Laboratory (California Institute for Quantitative Biosciences [QB3]). The quality of raw FASTQ paired reads was evaluated using FastQC (12). Low-quality reads and adapter sequences were removed from all paired raw reads using seqtk v1.2 (13) and cutadapt v1.14 (14), respectively. Briefly, seqtk uses Mott's algorithm to perform quality trimming. The algorithm is described online at http://www.phrap.org/phredphrap/phred.html. In addition, cutadapt searches for provided adapter sequences (detected by FastQC) in all reads and removes them when it finds them. A total of 9,914,835 paired fastq reads were used for the TPD3 assembly, and a total of 10,054,122 paired fastq reads were used for the TPD4 assembly. Preprocessed Illumina reads were assembled *de novo* with SPAdes v3.13 (15, 16) using the -careful parameter and -k of 21, 33, 55, and 77. The assembled contigs were reordered using Mauve's contig mover function (17) and the Temecula1 assembly (GenBank accession number GCF_000007245) as a reference. The TPD3 genome (818× sequencing depth of coverage) had a GC content of 51.49%, with a genome length of 2,422,083 bp distributed among 377 contigs ranging in size from 203 bp to 94,897 bp, an $N_{50}$ of 38,877 kb, and an $L_{50}$ of 20 kb. On the other hand, the TPD4 genome (828× sequencing depth of coverage) had a GC content of 51.47%, with a genome length of 2,427,175 bp distributed among 393 contigs ranging in size from 200 bp to 86,168 bp, an $N_{50}$ value of

31,881 kb, and an $L_{50}$ of 27 kb. The assembled and reordered genomes were individually annotated using PGAP (18). The TPD3 genome was predicted to have a total of 2,394 coding sequences (CDS), 51 tRNAs, 3 rRNAs, and 1 transfer-messenger RNA (tmRNA). The TPD4 genome was predicted to have 2,413 CDS, 51 tRNAs, 3 rRNAs, and 1 tmRNA.

Roary v3.11.2 was used to create an alignment of genes shared in 99% to 100% of the isolates in a data set (core-gene alignment) (19). The core-genome alignment was used to build a maximum likelihood (ML) tree using RaxML (20). The tree was built using the GTRCAT substitution model. Tree topology and branch support were assessed using 1,000 bootstrap replicates. The core-genome phylogenetic analyses of isolates TPD3 and TPD4 show a clear clustering within other *X. fastidiosa* subsp. *fastidiosa* isolates originating from the United States, thus providing further evidence for an introduction from the Americas into Taiwan (Fig. 1).

**Data availability.** All raw reads and information regarding each strain have been submitted under BioProject number PRJNA549761. TPD3 is submitted under BioSample number SAMN12097273, and TPD4 is submitted under BioSample number SAMN12097274. The accession numbers are VJWG00000000 (assembly accession number GCA_007845655) for TPD3 and VJWH00000000 (assembly accession number GCA_007845705) for TPD4.

## ACKNOWLEDGMENTS

This research was supported by the Bureau of Animal and Plant Health Inspection and Quarantine, Council of Agriculture, Executive Yuan, Taiwan (grants 101-AS-10.2.1-Q-B4 and 102-AS-10.2.1-Q-B5) and by the California Department of Food and Agriculture Pierce's Disease Research Program.

## REFERENCES

1. Rapicavoli J, Ingel B, Blanco-Ulate B, Cantu D, Roper C. 2018. *Xylella fastidiosa*: an examination of a re-emerging plant pathogen. Mol Plant Pathol 19:786–800. https://doi.org/10.1111/mpp.12585.
2. Overall LM, Rebek EJ. 2017. Insect vectors and current management strategies for diseases caused by *Xylella fastidiosa* in the southern United States. J Integr Pest Manag 8:12. https://doi.org/10.1093/jipm/pmx005.
3. Tuan SJ, Hu FT, Chang HY, Chang PW, Chen YH, Huang TP. 2016. *Xylella fastidiosa* transmission and life history of two Cicadellinae sharpshooters, *Kolla paulula* and *Bothrogonia ferruginea* (Hemiptera: Cicadellidae), in Taiwan. J Econ Entomol 109:1034–1040. https://doi.org/10.1093/jee/tow016.
4. Hopkins DL, Purcell AH. 2002. *Xylella fastidiosa*: cause of Pierce's disease of grapevine and other emergent diseases. Plant Dis 86:1056–1066. https://doi.org/10.1094/PDIS.2002.86.10.1056.
5. Tumber KP, Alston JM, Fuller KB. 2014. Pierce's disease costs California $104 million per year. Calif Agric (Berkeley) 68:20–29. https://doi.org/10.3733/ca.v068n01p20.
6. Su CC, Chang CJ, Chang CM, Shih HT, Tzeng KC, Jan FJ, Kao CW, Deng WL. 2013. Pierce's disease of grapevines in Taiwan: isolation, cultivation and pathogenicity of *Xylella fastidiosa*. J Phytopathol 161:389–396. https://doi.org/10.1111/jph.12075.
7. Su CC, Deng WL, Jan FJ, Chang CJ, Huang H, Shih HT, Chen J. 2016. *Xylella taiwanensis* sp. nov., causing pear leaf scorch disease. Int J Syst Evol Microbiol 66:4766–4771. https://doi.org/10.1099/ijsem.0.001426.
8. Saponari M, Giampetruzzi A, Loconsole G, Boscia D, Saldarelli P. 2019. *Xylella fastidiosa* in olive in Apulia: where we stand. Phytopathology 109:175–186. https://doi.org/10.1094/PHYTO-08-18-0319-FI.
9. Cella E, Angeletti S, Fogolari M, Bazzardi R, De L, Ciccozzi M, Cella E, Angeletti S, Fogolari M, Bazzardi R. 2018. Two different Xylella fastidiosa strains circulating in Italy: phylogenetic and evolutionary analyses. J Plant Interact 13:428–432. https://doi.org/10.1080/17429145.2018.1475022.
10. Coletta-Filho HD, Francisco CS, Lopes JRS, Muller C, Almeida R. 2017. Homologous recombination and *Xylella fastidiosa* host-pathogen associations in South America. Phytopathology 107:305–312. https://doi.org/10.1094/PHYTO-09-16-0321-R.
11. Nunney L, Hopkins DL, Morano LD, Russell SE, Stouthamer R. 2014. Intersubspecific recombination in *Xylella fastidiosa* strains native to the United States: infection of novel hosts associated with an unsuccessful invasion. Appl Environ Microbiol 80:1159–1169. https://doi.org/10.1128/AEM.02920-13.
12. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
13. Li H. 2012. Seqtk: a toolkit for processing sequences in FASTA/Q formats. https://github.com/lh3/seqtk.
14. Marcel M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.
15. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham SON, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev M, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
16. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembly single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20:714–737. https://doi.org/10.1089/cmb.2013.0084.
17. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. 2009. Reordering contigs of draft genomes using the Mauve Aligner. Bioinformatics 25:2071–2073. https://doi.org/10.1093/bioinformatics/btp356.
18. 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.
19. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421.
20. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.