



Genome Sequence of a Lethal Strain of Xylem-Invasive *Verticillium nonalfalfae*

Jernej Jakše,^a Vid Jelen,^a Sebastjan Radišek,^b Ronnie de Jonge,^{c,d} Stanislav Mandelc,^{a*} Aljaž Majer,^a Tomaž Curk,^e Blaž Zupan,^e  Bart P. H. J. Thomma,^f Branka Javornik^a

^aCentre for Plant Biotechnology and Breeding, Department of Agronomy, Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia

^bSlovenian Institute of Hop Research and Brewing, Žalec, Slovenia

^cBioinformatics & Systems Biology, Ghent University, Ghent, Belgium

^dPlant-Microbe Interactions Department, Utrecht University, Utrecht, The Netherlands

^eLaboratory of Bioinformatics, Faculty of Computer and Information Sciences, University of Ljubljana, Ljubljana, Slovenia

^fLaboratory of Phytopathology, Department of Plant Sciences, Wageningen University and Research, Wageningen, The Netherlands

ABSTRACT *Verticillium nonalfalfae*, a soilborne vascular phytopathogenic fungus, causes wilt disease in several crop species. Of great concern are outbreaks of highly aggressive *V. nonalfalfae* strains, which cause a devastating wilt disease in European hops. We report here the genome sequence and annotation of *V. nonalfalfae* strain T2, providing genomic information that will allow better understanding of the molecular mechanisms underlying the development of highly aggressive strains.

Verticillium nonalfalfae isolates from hops express two types of aggressiveness, appearing as a mild or lethal form of the wilt disease. This disease is characterized by plant stunting and wilting, vascular browning, foliar chlorosis and necrosis, and, in the case of the lethal form, rapid plant withering and dieback (1). By comparative genomics of mild and lethal *V. nonalfalfae* hop isolates (BioProject no. PRJNA283258), we anticipated the dissection of the pathotype genomes to provide an insight into their genomic structure. This in turn might explain the increased virulence of the lethal strain, enable detection of virulence-associated factors, and elucidate the pathogenicity in *Verticillium* spp.

DNA from protoplasts prepared from *V. nonalfalfae* conidia that germinated overnight (2) was isolated following the cetyltrimethylammonium bromide (CTAB) protocol (3). Whole-genome sequencing was carried out by IGA Technology Services (Udine, Italy) and the Beijing Genomics Institute (BGI, Hong Kong) using Illumina Genome Analyzer IIx and HiSeq 2500 systems with a paired-end setup. For the reference genome assembly (strain T2), four libraries of different insert sizes (5,000-bp mate paired-end and 1,000-bp, 500- to 600-bp, and 370-bp paired-end inserts) were sequenced (BioSample no. SAMN03610599). In total, 98 million raw reads were generated, representing a 390× sequencing depth. Reference genome assembly was performed using the CLC Genomics Server (ver. 5.0.1) *de novo* assembly tool (Qiagen, USA). Sequencing data were assembled in 409 scaffolds with a G+C content of 55.35% and a total assembly size of 34.18 Mb. The *N*₅₀ length was 292,274 bp, the largest scaffold was 1,095,096 bp, and 50% of the assembly was represented by 34 scaffolds. There were 1,740 unknown nucleotides (nt) per 100 kb. In addition, RNA sequencing of mild and lethal *V. nonalfalfae* pathotypes from Slovenia was performed (Sequence Read Archive accession no. SRX1020679 and SRX1020629) using total RNA [enriched for the poly(A) mRNA fraction] isolated from fungal mycelia grown in xylem-simulating medium (4). For gene anno-

Received 21 November 2017 Accepted 1 December 2017 Published 11 January 2018

Citation Jakše J, Jelen V, Radišek S, de Jonge R, Mandelc S, Majer A, Curk T, Zupan B, Thomma BPHJ, Javornik B. 2018. Genome sequence of a lethal strain of xylem-invasive *Verticillium nonalfalfae*. *Genome Announc* 6:e01458-17. <https://doi.org/10.1128/genomeA.01458-17>.

Copyright © 2018 Jakše 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 Jernej Jakše, jernej.jakse@bf.uni-lj.si, or Branka Javornik, branka.javornik@bf.uni-lj.si.

*Present address: Stanislav Mandelc, Lek d.d., Sandoz, Mengeš, Slovenia.

tation, we followed the *V. dahliae* annotation pipeline (5), in which various evidences of genic regions are combined into a final gene model, as follows: i) Cufflinks (6) identification of transcripts with transcriptome sequencing (RNA-seq) data and PASA assembly of transcripts (7), ii) protein alignments of 15 fungal species obtained by AAT (8) and Exonerate software (9), and iii) *ab initio* prediction using GeneMark-ES (10) and Augustus (11) predictions. EvidenceModeler software (7) was then used to obtain gene models. The finalized genome annotation includes 9,269 protein coding genes. Most of the genes have a coding sequence (CDS) length in the range of 1,000 to 2,000 bp, with the majority being <5,000 nucleotides (nt). Most genes have <5 exons, with a prevalence of 2 exons. The CEGMA pipeline (12) identified 445 of the 458 clusters of orthologous groups (COGs) to be present (completeness, 97.2%).

Additionally, lower-coverage sequencing (paired-end) of 2 lethal and 3 mild strains was performed and accumulated from 9.6 to 43.9 million reads (BioSample no. SAMN03612289, SAMN03611609, SAMN03610662, SAMN03612288, and SAMN03612257). Comparative genomics of mild and lethal *V. nonalfalfae* pathotypes revealed a 0.547-kb lethal pathotype-specific region, which is absent in mild pathotypes.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [PDEZ00000000](https://doi.org/10.1093/nar/gk916). The version described in this paper is the first version, PDEZ01000000.

ACKNOWLEDGMENT

This work was supported by the Slovenian Research Agency, Ljubljana, Slovenia (grant P4-0077 to Branka Javornik).

REFERENCES

1. Radišek S, Jakše J, Javornik B. 2006. Genetic variability and virulence among *Verticillium albo-atrum* isolates from hop. *Eur J Plant Pathol* 116:301–314. <https://doi.org/10.1007/s10658-006-9061-0>.
2. Pantou MP, Typas MA. 2005. Electrophoretic karyotype and gene mapping of the vascular wilt fungus *Verticillium dahliae*. *FEMS Microbiol Lett* 245:213–220. <https://doi.org/10.1016/j.femsle.2005.03.011>.
3. Möller EM, Bahnweg G, Sandermann H, Geiger HH. 1992. A simple and efficient protocol for isolation of high-molecular-weight DNA from filamentous fungi, fruit bodies, and infected-plant tissues. *Nucleic Acids Res* 20:6115–6116. <https://doi.org/10.1093/nar/20.22.6115>.
4. Neumann MJ, Dobinson KF. 2003. Sequence tag analysis of gene expression during pathogenic growth and microsclerotia development in the vascular wilt pathogen *Verticillium dahliae*. *Fungal Genet Biol* 38: 54–62. [https://doi.org/10.1016/S1087-1845\(02\)00507-8](https://doi.org/10.1016/S1087-1845(02)00507-8).
5. de Jonge R, Bolton MD, Kombrink A, van den Berg GCM, Yadeta KA, Thomma BP. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. *Genome Res* 23:1271–1282. <https://doi.org/10.1101/gr.152660.112>.
6. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28:511–515. <https://doi.org/10.1038/nbt.1621>.
7. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EvidenceModeler and the program to assemble spliced alignments. *Genome Biol* 9:R7. <https://doi.org/10.1186/gb-2008-9-1-r7>.
8. Huang XQ, Adams MD, Zhou H, Kerlavage AR. 1997. A tool for analyzing and annotating genomic sequences. *Genomics* 46:37–45. <https://doi.org/10.1006/geno.1997.4984>.
9. Slater GS, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6:31. <https://doi.org/10.1186/1471-2105-6-31>.
10. Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M. 2005. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res* 33:6494–6506. <https://doi.org/10.1093/nar/gki937>.
11. Stanke M, Waack S. 2003. Gene prediction with a hidden Markov model and new intron submodel. *Bioinformatics* 19:ii215–ii225. <https://doi.org/10.1093/bioinformatics/btg1080>.
12. Parra G, Bradnam K, Ning Z, Keane T, Korf I. 2009. Assessing the gene space in draft genomes. *Nucleic Acids Res* 37:289–297. <https://doi.org/10.1093/nar/gkn916>.