Microbiology Resource Announcements

| Host-Microbial Interactions | Announcement

The draft genome sequence of *Diaporthe vaccinii*, isolated from diseased cranberries

Bhagya C. Thimmappa,[1] Matt Sarrasin,[1] B. Franz Lang,[1] Gertraud Burger[1]

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT We report the assembly and annotation of the nuclear genome of *Diaporthe vaccinii*, a pathogenic fungus isolated from diseased cranberries in Quebec, Canada. The genome was sequenced with the Illumina paired-end sequencing technology, assembled into 67 Mbp across 588 contigs, with an N50 of 386 Kbp and 97.5% BUSCO completeness.

KEYWORDS fruit-rot pathogen, nuclear genome, *Diaporthe vaccinii*, phytopathogen, *Vaccinium macrocarpon*

The North American cranberry is economically important but susceptible to fungal pathogens, collectively called the cranberry fruit rot (CFR) complex, causing significant crop losses (1–3). One CFR genus is *Diaporthe*, with the species *Diaporthe vaccinii* (anamorph *Phomopsis vaccinii*) causing, for example, stem cankers and soft rot in blueberries (4) and cranberries (*Vaccinium* spp.) (5–8).

Our previous work showed that the *D. vaccinii* isolate IS7, isolated from diseased cranberries, infects cranberry plants and leads to their death (9). Here, we report a high-quality draft genome of *D. vaccinii* IS7 that will facilitate further investigation of its pathogenicity.

The sample was collected by Richard Bélanger (Université Laval, Québec, Canada) in 2013 from the flesh of diseased cranberries in Saint-Louis-de-Blandford (Quebec-Canada) using the hyphal tip isolation method and kindly provided to us. The isolate was grown on potato dextrose agar medium and maintained at 4°C with subculturing every 6 months. For genome sequencing, IS7 was cultured on yeast-glycerol medium (YGM; yeast extract 0.4%, glycerol 2%, and pH 7) at room temperature for 3 days and disrupted in a blender (Hamilton Beach, model-51109C, 225 W for 3 min). DNA was isolated using the Qiagen DNeasy PowerPlant Pro Kit (Qiagen-GmbH) according to the manufacturer's protocol. Taxonomic identification was performed by PCR amplification of the Internal Transcribed Spacer (ITS) region (ITS1-5.8S-ITS2) using the fungus-specific primers BMB-CR forward 5′-GTACACACCGCCCGTCG-3′ (binds small-subunit rRNA) and ITS4 reverse 5′-TTCCWCCGCTTATTGATATGC-3′ (binds large-subunit rRNA) (10); the PCR product was purified using the QIAquick Gel Extraction Kit (Qiagen-GmbH) and Sanger sequenced by the Institut de Recherche en Immunologie et en Cancer (IRIC) Genomics Platform (Montreal-Canada) using the above primers. A blastn (v2.16.1) search (11) against the National Center for Biotechnology Information (NCBI) non-redundant database (update date: 30 December 2024) assigned IS7 to *D. vaccinii* (GenBank accession PP921214.1) with 100% identity and coverage.

For whole-genome sequencing, liquid YGM was inoculated with ~1 × 10^6 Colony-Forming Units (CFUs) of an IS7 hyphal suspension. The culture was grown under shaking for 3 days at room temperature. Genomic DNA was isolated using the DNeasy Plant Mini Kit (Qiagen-GmbH) according to the manufacturer's recommendation. Library preparation (KAPA-HyperPrep Kit, Roche-Switzerland) and Illumina MiSeq paired-end

Editor Leighton Pritchard, University of Strathclyde, Glasgow, United Kingdom

Address correspondence to Bhagya C. Thimmappa, bhagya.c.thimmappa@umontreal.ca.

The authors declare no conflict of interest.

Received 9 October 2024
Accepted 11 March 2025
Published 2 April 2025

TABLE 1 Nuclear genome statistics of *D. vaccinii* IS7

Genome feature	
Nr. of raw reads	16,891,320
Raw read length	300 bp
Estimated coverage	25×
Size of the assembly	67,097,566 bp
Nr. of contigs (>200 bp)	588
N50	386 Kbp
Size of the largest contig	1,775,149 bp
G + C content	46.4%
Genome completeness (BUSCO)	97.5%
Nr. of protein-coding genes	15,650
Mean gene length	2,099 bp
Mean exon length	736 bp
Annotation completeness (BUSCO)	99.8%

sequencing with a 300 bp read length were outsourced to the Genome Quebec Center in Montreal.

Reads were trimmed of adapters using Trimmomatic (v0.35) (12) and corrected with Rcorrector (v1.0.4) (13). Default parameters were used for all software unless specified. The genome was *de novo* assembled using the SPAdes assembler (v3.14.1) (14). Structural genome annotation was performed using an in-house pipeline, which masks repeats and searches protein sequence similarity in UniProtKB and RefSeq (downloaded November 2017) employing Spaln (v2.2.2) (15). The *ab initio* predictors used were Augustus (v3.3.2) (16), Snap (17), Genemark (v4.33) (18), and CodingQuarry (v2.0) (19). Functional assignments of protein-coding genes were performed by profile Hidden Markov Model (HMM) searches with HMMER3 (v3.4) (20) in UniProtKB/Swiss-Prot release 2022_02 and PFAM database (v37.1) (21). Genome assembly completeness was assessed with BUSCO (v5.1.0) (22) using the OrthoDB (v10) Sordariomycetes data set (23). Details of the assembly are compiled in Table 1.

ACKNOWLEDGMENTS

We thank Dr. Richard Bélanger (Université Laval, Québec, Canada) for providing cranberry pathogenic strains and Dr. Lila N. Salhi and Lise Forget (Université de Montréal) for their help with strain cultivation and maintenance.

This research was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC), project title "MycAtok Project: innovating cranberry farming with beneficial symbiotic microbes," funding number CRDPJ 514188–2017 to G.B. and B.F.L., and a J. A. DeSève Excellence Scholarship to B.C.T.

AUTHOR AFFILIATION

[1]Department of Biochemistry, Robert-Cedergren Center for Bioinformatics and Genomics, Université de Montréal, Montreal, Québec, Canada

AUTHOR ORCIDs

Bhagya C. Thimmappa http://orcid.org/0000-0002-4299-5478
Matt Sarrasin http://orcid.org/0000-0003-4849-0849
B. Franz Lang http://orcid.org/0000-0003-1035-5449
Gertraud Burger http://orcid.org/0000-0002-8679-8812

AUTHOR CONTRIBUTIONS

Bhagya C. Thimmappa, Formal analysis, Investigation, Methodology, Software, Writing – original draft | Matt Sarrasin, Formal analysis, Methodology | B. Franz Lang, Funding

acquisition, Supervision, Writing – review and editing | Gertraud Burger, Funding acquisition, Supervision, Writing – review and editing

DATA AVAILABILITY

Genome sequence reads (SRR28883308) and RNA-Seq data (SRR28883305, SRR28883306, and SRR28883307) were deposited in the Sequence Read Archive, and the Whole-Genome Shotgun project (JBAWTH000000000) at DDBJ/ENA/GenBank under BioProject PRJNA1084027. The version described in this paper is JBAWTH010000000.

REFERENCES

1. Stiles CM, Oudemans PV. 1999. Distribution of cranberry fruit-rotting fungi in new jersey and evidence for nonspecific host resistance. Phytopathology 89:218–225. https://doi.org/10.1094/PHYTO.1999.89.3.218
2. Olatinwo RO, Hanson EJ, Schilder AMC. 2003. A first assessment of the cranberry fruit rot complex in Michigan. Plant Dis 87:550–556. https://doi.org/10.1094/PDIS.2003.87.5.550
3. Conti M, Cinget B, Labbe C, Bélanger RR. 2020. First report of Godronia cassandrae as a major cranberry fruit rot pathogen in Eastern Canada. Plant Dis. https://doi.org/10.1094/PDIS-06-20-1193-PDN
4. Hilário S, Santos L, Alves A. 2021. Diversity and pathogenicity of *Diaporthe* species revealed from a survey of blueberry orchards in Portugal. Agriculture 11:1271. https://doi.org/10.3390/agriculture11121271
5. Michalecka M, Bryk H, Seliga P. 2017. Identification and characterization of *Diaporthe vaccinii* shear causing upright dieback and viscid rot of cranberry in Poland. Eur J Plant Pathol 148:595–605. https://doi.org/10.1007/s10658-016-1114-4
6. Farr DF, Castlebury LA, Rossman AY. 2002. Morphological and molecular characterization of *Phomopsis vaccinii* and additional isolates of Phomopsis from blueberry and cranberry in the eastern United States. Mycologia 94:494–504. https://doi.org/10.1080/15572536.2003.11833214
7. van Bruggen AHC, West JS, van der Werf W, Potting RPJ, Gardi C, Koufakis I, Zelenev VV, Narouei - Khandan H, Schilder A, Harmon P. 2018. Input data needed for a risk model for the entry, establishment and spread of a pathogen (*Phomopsis vaccinii*) of blueberries and cranberries in the EU. Annals of Applied Biology 172:126–147. https://doi.org/10.1111/aab.12414
8. Lombard L, Leeuwen G, Guarnaccia V, Polizzi G. 2014. *Diaporthe* species associated with vaccinium, with specific reference to Europe. Phytopathol Mediterr 53:287–299.
9. Thimmappa BC, Salhi LN, Forget L, Sarrasin M, Bustamante Villalobos P, Henrissat B, Lang BF, Burger G. 2024. A biofertilizing fungal endophyte of cranberry plants suppresses the plant pathogen *Diaporthe*. Front Microbiol 15:1327392. https://doi.org/10.3389/fmicb.2024.1327392
10. Conserved primer sequences for PCR amplification of fungal rDNA. 2018. Duke University, Vilgalys Mycology Lab.
11. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421
12. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
13. Song L, Florea LR. 2015. Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. Gigascience 4:48. https://doi.org/10.1186/s13742-015-0089-y
14. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, 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
15. Gotoh O. 2008. A space-efficient and accurate method for mapping and aligning cDNA sequences onto genomic sequence. Nucleic Acids Res 36:2630–2638. https://doi.org/10.1093/nar/gkn105
16. Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res 32:W309–12. https://doi.org/10.1093/nar/gkh379
17. Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics 5:59. https://doi.org/10.1186/1471-2105-5-59
18. Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res 33:W451–4. https://doi.org/10.1093/nar/gki487
19. Testa AC, Hane JK, Ellwood SR, Oliver RP. 2015. CodingQuarry: highly accurate hidden Markov model gene prediction in fungal genomes using RNA-seq transcripts. BMC Genomics 16:170. https://doi.org/10.1186/s12864-015-1344-4
20. Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–37. https://doi.org/10.1093/nar/gkr367
21. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A. 2021. Pfam: The protein families database in 2021. Nucleic Acids Res 49:D412–D419. https://doi.org/10.1093/nar/gkaa913
22. Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. Methods Mol Biol Clifton NJ:227–245. https://doi.org/10.1007/978-1-4939-9173-0_14
23. Kuznetsov D, Tegenfeldt F, Manni M, Seppey M, Berkeley M, Kriventseva EV, Zdobnov EM. 2023. OrthoDB V11: annotation of orthologs in the widest sampling of organismal diversity. Nucleic Acids Res 51:D445–D451. https://doi.org/10.1093/nar/gkac998