

Contents lists available at [ScienceDirect](http://ScienceDirect.com)

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Data in Brief

Draft whole-genome sequence of the *Diaporthe helianthi* 7/96 strain, causal agent of sunflower stem cankerRiccardo Baroncelli ^a, Felice Scala ^{b,c}, Mariarosaura Vergara ^d, Michael R. Thon ^e, Michelina Ruocco ^{c,*}^a *Universitaire de Biodiversité et Ecologie Microbienne (LUBEM), University of Western Brittany, Avenue du Technopole, 29280 Plouzané, France*^b *Dipartimento di Agraria, via Università 100, 80055 portici (NA), Italy*^c *Istituto per la Protezione Sostenibile delle Piante, sede secondaria di Portici, (IPSP-CNR) via Università 133, 80055 Portici (NA), Italy*^d *Scuola Normale Superiore di Pisa, piazza dei Cavalieri 7, 56126 Pisa, Italy*^e *Instituto Hispano-Luso de Investigaciones Agrarias (CIALE), University of Salamanca, Campus de Villamayor, C/ Del Duero, 12; 37185 Villamayor (Salamanca), Spain*

ARTICLE INFO

Article history:

Received 12 October 2016

Received in revised form 3 November 2016

Accepted 9 November 2016

Available online 11 November 2016

ABSTRACT

Diaporthe helianthi is a fungus pathogenic to sunflower. Virulent strains of this fungus cause stem canker with important yield losses and reduction of oil content. Here we present the first draft whole-genome sequence of the highly virulent isolate *D. helianthi* strain 7/96, thus providing a useful platform for future research on stem canker of sunflower and fungal genomics. The genome sequence of the *D. helianthi* isolate 7/96 was deposited at DDBJ/ENA/GenBank under the accession number MAVT000000000 (BioProject PRJNA327798).

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Specifications [standardized info for the reader]

Organism/cell line/tissue	<i>Diaporthe helianthi</i> strain 7/96
Sex	N/A
Sequencer or array type	Illumina HiSeq 2000
Data format	Assembled
Experimental factors	Fungal strain pathogen on sunflower
Experimental features	De novo genome sequencing followed by assembly and annotation
Consent	N/A
Sample source location	<i>D. helianthi</i> 7/96 was isolated from diseased sunflowers in France

1. Direct link to deposited data

<https://www.ncbi.nlm.nih.gov/bioproject/327798>

2. Experimental design, materials and methods

Diaporthe helianthi Munt.-Cvet [1], anamorph *Phomopsis helianthi*, is one of the causal agent of stem canker of sunflower (*Helianthus*

annuus). The fungus belongs to division Ascomycota, subdivision Pezizomycotina, class Sordariomycetes, order Diaporthales, family Diaporthaceae. Disease develops on leaves by causing brown necrotic lesions at the edge of the leaf blade and then spreads through the veins to the petiole. The leaf dries and remains attached to the stem. From the junction of the leaf the pathogen gradually infects the stem causing diffuse necrosis. The main loss is the shriveling of capitula and lodging of stem, which breaks at different levels. The fungus overwinters as mycelium and perithecia on stem debris left on the soil. Ascospores are released as a yellowish jelly soluble in rainwater. The first contamination of the leaves by ascospores can occur in early spring.

D. helianthi 7/96 was isolated from diseased sunflowers in France. This isolate belongs to a group of highly virulent isolates characterized by the presence of a gene fragment having homologies with *Aspergillus terreus* lovastatin nonaketide polyketide synthase gene (*lovB*) [2].

The genome of *D. helianthi* strain 7/96 was sequenced using Illumina mate-paired sequencing technology by BGI HONG KONG CO (BGI-HONGKONG CO., Limited, 16 Dai Fu Street, Tai Po Industrial Estate, Tai Po, New Territories, Hong Kong). Mate-paired reads of 90 bp (1.80 Gbp; average coverage 38.80×) were assembled using Velvet V 1.2.10 [3]. The genome of *D. helianthi* consists of 13,383 sequence scaffolds with a total assembly length 46.50 Mbp (N50 = 6347 and N90 = 1502), 51.10% GC-content, and a maximum scaffold size of 54,299 bp. The completeness of the assembly was assessed using BUSCO v12 [4], which estimated the genome sequence to be 98.26%

* Corresponding author.

E-mail address: michelina.ruocco@ipsp.cnr.it (M. Ruocco).

complete. The genome was annotated using the MAKER2 pipeline [5]. Overall, 14,220 protein-coding gene models were predicted. Analysis with SignalP 4.1 [6] revealed that 1433 predicted proteins (9.92% of the proteome) contain a secretion signal peptide. Among those 45 are species-specific as they do not have any sequence similarity to proteins in public databases, based on BLAST searches (e-value threshold of $1e^{-5}$). Such features are characteristic of fungal effectors, which are proteins that have key roles in restricting the host defense [7]. A first comparative analysis within Diaportheales [8] and model organisms with publically available genomes [*Fusarium* [9], *Neurospora* [10], *Colletotrichum* [11], *Magnaporthe* [12]] evidenced that *D. helianthi* 7/96 genome contains a remarkable number (at least 40) of putative polyketide synthases (PKSs). Analysis of PKS gene family in strain 7/96 could be an intriguing starting point to unravel *D. helianthi* sunflower interactions.

2.1. Nucleotide sequence accession numbers

This Whole-Genome Shotgun project has been deposited in GenBank under the accession N^o: MAVT00000000 (BioProject PRJNA327798). The version described in this paper is: MAVT00000000.1.

Funding information

This work was supported by a dedicated grant from the Italian Ministry of Economy and Finance to the National Research Council for the project Innovazione e Sviluppo del Mezzogiorno - Conoscenze Integrate per Sostenibilità ed Innovazione del Made in Italy Agroalimentare - Legge n. 191/2009.

References

- [1] M. Muntañola-Cvetkovic, M. Mihaljcevi, M. Petrof, On the identity of the causative agent of a serious *Phomopsis-Diaporthe* disease in sunflower plants. *Nova Hedwigia* 34 (1981) 417–435.
- [2] M. Vergara, C. Cristani, C. Regis, G. Vannacci, A coding region in *Diaporthe helianthi* reveals genetic variability among isolates of different geographic origin. *Mycopathologia* 158 (2004) 123–130.
- [3] D.R. Zerbino, E. Birney, Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18 (5) (2008) 821–829.
- [4] F.A. Simão, R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31 (19) (2015) 3210–3212.
- [5] C. Holt, M. Yandell, MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12 (2011) 491.
- [6] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8 (2011) 785–786.
- [7] R. Baroncelli, J.M. Sanz-Martín, G.E. Rech, S.A. Sukno, M.R. Thon, Draft genome sequence of *Colletotrichum sublineola*, a destructive pathogen of cultivated sorghum. *Genome Announcement* 2 (3) (2014) e00540-14.
- [8] A. Morales-Cruz, K.C. Amrine, B. Blanco-Ulate, D.P. Lawrence, R. Travadon, P.E. Rolshausen, K. Baumgartner, D. Cantu, Distinctive expansion of gene families associated with plant cell wall degradation, secondary metabolism, and nutrient uptake in the genomes of grapevine trunk pathogens. *BMC Genomics* 19 (16) (2015) 469.
- [9] L.J. Ma, H.C. van der Does, K.A. Borkovich, J.J. Coleman, M.J. Daboussi, A. Di Pietro, M. Dufresne, M. Freitag, M. Grabherr, B. Henrissat, P.M. Houterman, S. Kang, W.B. Shim, C. Woloshuk, X. Xie, J.R. Xu, J. Antoniw, S.E. Baker, B.H. Bluhm, A. Breakspear, D.W. Brown, R.A. Butchko, S. Chapman, R. Coulson, P.M. Coutinho, E.G. Danchin, A. Diener, L.R. Gale, D.M. Gardiner, S. Goff, K.E. Hammond-Kosack, K. Hilburn, A. Hua-Van, W. Jonkers, K. Kazan, C.D. Kodira, M. Koehrsen, L. Kumar, Y.H. Lee, L. Li, J.M. Manners, D. Miranda-Saavedra, M. Mukherjee, G. Park, J. Park, S.Y. Park, R.H. Proctor, A. Regev, M.C. Ruiz-Roldan, D. Sain, S. Sakthikumar, S. Sykes, D.C. Schwartz, B.G. Turgeon, I. Wapinski, O. Yoder, S. Young, Q. Zeng, S. Zhou, J. Galagan, C.A. Cuomo, H.C. Kistler, M. Rep, Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464 (7287) (2010) 367–373.
- [10] J.E. Galagan, S.E. Calvo, K.A. Borkovich, E.U. Selker, N.D. Read, D. Jaffe, W. FitzHugh, L.J. Ma, S. Smirnov, S. Purcell, B. Rehman, T. Elkins, R. Engels, S. Wang, C.B. Nielsen, J. Butler, M. Endrizzi, D. Qui, P. Ianakiev, D. Bell-Pedersen, M.A. Nelson, M. Werner-Washburne, C.P. Selitrennikoff, J.A. Kinsey, E.L. Braun, A. Zelter, U. Schulte, G.O. Kothe, G. Jedd, W. Mewes, C. Staben, E. Marcotte, D. Greenberg, A. Roy, K. Foley, J. Naylor, N. Stange-Thomann, R. Barrett, S. Gnerre, M. Kamal, M. Kamvysseles, E. Mauceli, C. Bielke, S. Rudd, D. Frishman, S. Krystofova, C. Rasmussen, R.L. Metznerberg, D.D. Perkins, S. Kroken, C. Cogoni, G. Macino, D. Catcheside, W. Li, R.J. Pratt, S.A. Osmani, C.P. DeSouza, L. Glass, M.J. Orbach, J.A. Berglund, R. Voelker, O. Yarden, M. Plamann, S. Seiler, J. Dunlap, A. Radford, R. Aramayo, D.O. Natvig, L.A. Alex, G. Mannhaupt, D.J. Ebbole, M. Freitag, I. Paulsen, M.S. Sachs, E.S. Lander, C. Nusbaum, B. Birren, The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422 (6934) (2003) 859–868.
- [11] R. Baroncelli, D.B. Amby, A. Zapparata, S. Sarrocco, G. Vannacci, G. Le Floch, R.J. Harrison, E. Holub, S.A. Sukno, S. Sreenivasaprasad, M.R. Thon, Gene family expansions and contractions are associated with host range in plant pathogens of the genus *Colletotrichum*. 2016. *BMC Genomics* 17 555, <http://dx.doi.org/10.1186/s12864-016-2917-6>.
- [12] D. RA, T. NJ, E. DJ, F. ML, M. TK, O. MJ, M. Thon, R. Kulkarni, X. JR, H. Pan, R. ND, L. YH, I. Carbone, D. Brown, O. YY, N. Donofrio, J. JS, S. DM, S. Djonovic, E. Kolomiets, C. Rehmeier, L. W, M. Harding, S. Kim, L. MH, H. Bohnert, S. Coughlan, J. Butler, S. Calvo, M. LJ, R. Nicol, S. Purcell, C. Nusbaum, G. JE, B. BW, The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 434 (7036) (2005) 980–986.