

Genome Sequence Data of the Soybean Pathogen *Stagonosporopsis vannaccii*: A Resource for Studies on *Didymellaceae* Evolution

Giovanni Cafà,¹ Thaís Regina Bouffleur,² Renata Rebellato Linhares de Castro,² Nelson Sidnei Massola, Jr.,² and Riccardo Baroncelli^{3,†}

¹ CABI Europe-UK, Bakeham Lane, Egham, Surrey TW20 9TY, U.K.

² University of São Paulo, ESALQ, Department of Plant Pathology and Nematology, Piracicaba/SP, Brazil

³ Instituto Hispano-Luso de Investigaciones Agrarias (CIALE), University of Salamanca, Calle del Duero, 12; 37185 Villamayor (Salamanca), Spain

Abstract

The genus *Stagonosporopsis* is classified within the Didymellaceae family and has around 40 associated species. Among them, several species are important plant pathogens responsible for significant losses in economically important crops worldwide. *Stagonosporopsis vannaccii* is a newly described species pathogenic to soybean. Here, we present the draft whole-genome sequence, gene prediction, and annotation of *S. vannaccii* isolate LFN0148 (also known as IMI 507030). To our knowledge, this is the first genome sequenced of this species and represents a new useful source for future research on fungal comparative genomics studies.

Genome Announcement

Genus *Stagonosporopsis* is a monophyletic fungus genus within the Didymellaceae (Pleosporales, Dothideomycetes, Ascomycota) comprising around 40 species, of which 28 are recognized based on molecular data (Crous et al. 2019; Marin-Felix et al. 2019). With new species being described recently, i.e., *Stagonosporopsis tanacetii* (Vaghefi et al. 2012) and *Stagonosporopsis vannaccii* (Crous et al. 2019), the *Stagonosporopsis* genus includes several plant-pathogenic fungi of economically important cultures worldwide, such as *Stagonosporopsis curcubitacearum*, which is the causal agent of the gummy stem blight, one of the major diseases that affects plants of the family Curcubitaceae (Newark et al. 2020), and two quarentenary pathogens, *Stagonosporopsis andigena* and *Stagonosporopsis chrysanthemy*.

Stagonosporopsis vannaccii is a newly described fungal species and an emerging pathogen of soybean (*Glycine max*) (Crous et al. 2019). According to the data collected from culture collections and GenBank, *S. vannaccii* is restricted to its host in Brazil. *Stagonosporopsis vannaccii* IMI 507030 was isolated from anthracnose-like symptoms on pods of soybean in Sinop, Mato Grosso, central Brazil in 2017 and pathogenicity to its host was confirmed by artificial inoculation (Crous et al. 2019). As reported by the authors who described the species, it is very likely that the disease caused by *S. vannaccii* in soybean is being confused with damping-off caused by *Colletotrichum* spp. in the field (Crous et al. 2019).

The genome of *S. vannaccii* LFN0148 was sequenced with Illumina MiSeq paired-end sequencing technology (v2 cartridge; 300 cycles). Paired reads of 150 bp (3.29 Gbp; average

Funding

This work was funded by the National Science and Technology Development Council grant CNPq Process 70328/1997-2, Escalera de Excelencia grant CLU-2018-04, the São Paulo Research Foundation grant FAPESP 2017/09178-8, and the Institutional Program for internalization grant CAPES/Print Process 88887.368016/2019-00.

Keywords

Dothideomycetes, Pleosporales, plant-pathogenic fungus, stem blight, genomics

†Corresponding author: R. Baroncelli; riccardobaroncelli@gmail.com

The author(s) declare no conflict of interest.

Accepted for publication 18 April 2020.

Table 1. Summary statistics of the *Stagonosporopsis vannaccii* IMI 507030 genome

Variables	Statistics
Number of scaffolds	483
Assembly length (bp)	30,958,390
N ₅₀	204,158
L ₅₀	48
GC (%)	53.26
BUSCO completeness (%)	97.9
Number of predicted genes	12,024
Number of predicted secreted proteins	1,304

coverage about 50x) were assembled using SPAdes v 3.13.1 (Bankevich et al. 2012). The completeness of the assembly was assessed using BUSCO 3.0.2 (Simão et al. 2015), which estimated the genome sequence to be 97.9% complete. The assembled genome was manually curated, and low/high coverage scaffolds were removed. The nuclear genome of *S. vannaccii* consists of 483 sequence scaffolds, with a total assembly length 30.96 Mbp (N₅₀ = 204,158 and L₅₀ = 48), 53.26% GC content, and a maximum scaffold size of 803,877 bp (Table 1). Surprisingly, the genome obtained is 25% smaller than the one of *S. tanacetii* DAR 70020 that was previously sequenced (Wingfield et al. 2015), the higher fragmentation of the *S. tanacetii* DAR 70020 (also known as CBS 131484) assembly may suggest a higher content of repetitive genomic regions compared with *S. vannaccii*.

The nuclear genome was annotated using the MAKER pipeline v2.31.10 (Holt and Yandell 2011) based on GeneMark-ES v. 4.48 (Borodovsky and Lomsadze 2011), Augustus v. 2.5.5 (Stanke and Morgenstern, 2005), as well as evidence-based gene prediction using the expressed sequence tags and protein sequences of other available Didymellaceae genomes (Grigoriev et al. 2014; Verma et al. 2016). Overall, 12,024 protein-coding gene models were predicted. Analysis with SignalP 4.1 (Petersen et al. 2011) revealed that 1,304 predicted proteins (10.84% of the proteome) contain a secretion signal peptide. Protein clustering analyses (Emms and Kelly 2015) with genomes belonging to the family Didymellaceae, available in the MycoCosm portal (Grigoriev et al. 2014), revealed that 836 are proteins encoded by *S. vannaccii* not shared with closely related sequenced species. Among these, only 36 have a secretion signal and have no similarity with any known proteins. Such features are characteristic of fungal species-specific effectors, which are proteins that might be involved in the suppression or evasion of host defense.

In the present study, we provide the draft genome sequence of *S. vannaccii*, a fungal species belonging to the family Didymellaceae. To our knowledge, this is the first *S. vannaccii* genome sequenced and the first genome of its genus with available gene prediction. The sequence represents a new useful source for future research on fungal comparative genomics studies.

Nucleotide sequence accession numbers

This Whole-Genome Shotgun project has been deposited in GenBank under the accession number VWXS00000000 (BioProject: PRJNA564422; BioSample: SAMN12712260; Short Read Archive: SRR11555624). The version described in this paper is VWXS00000000.1.

Literature Cited

- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19:455-477.
- Borodovsky, M., and Lomsadze, A. 2011. Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES. *Curr. Protoc. Bioinformatics* Chapter 4: Unit 4.6.1-10.
- Crous, P. W., Wingfield, M. J., Lombard, L., Roets, F., Swart, W. J., Alvarado, P., Carnegie, A. J., Moreno, G., Luangsaard, J., Thangavel, R., Alexandrova, A. V., Baseia, I. G., Bellanger, J. M., Bessette, A. E., Bessette, A. R., De la Peña-Lastra, S., García, D., Gené, J., Pham, T. H. G., Heykoop, M., Malysheva, E., Malysheva, V., Martín, M. P., Morozova, O. V., Noisripoom, W., Overton, B. E., Rea, A. E., Sewall, B. J., Smith, M. E., Smyth, C. W., Tسانathai, K., Visagie, C. M., Adamčík, S., Alves, A., Andrade, J. P., Aninat, M. J., Araújo, R. V. B., Bordallo, J. J., Bouffleur, T., Baroncelli, R., Barreto, R. W., Bolin, J., Cabero, J., Caboň, M., Cafà, G., Caffot, M. L. H., Cai, L., Carlavilla, J. R., Chávez, R., de

- Castro, R. R. L., Delgat, L., Deschuyteneer, D., Dios, M. M., Domínguez, L. S., Evans, H. C., Eyssartier, G., Ferreira, B. W., Figueiredo, C. N., Liu, F., Fournier, J., Galli-Terasawa, L. V., Gil-Durán, C., Glienke, C., Gonçalves, M. F. M., Gryta, H., Guarro, J., Himaman, W., Hywel-Jones, N., Iturrieta-González, I., Ivanushkina, N. E., Jargeat, P., Khalid A. N., Khan, J., Kiran, M., Kiss, L., Kochkina, G. A., Kolařík, M., Kubátová, A., Lodge, D. J., Loizides, M., Luque, D., Manjón, J. L., Marbach, P. A. S., Massola, N. S., Jr., Mata, M., Miller, A. N., Mongkolsamrit, S., Moreau, P. A., Morte, A., Mujic, A., Navarro-Ródenas, A., Németh, M. Z., Nóbrega, T. F., Nováková, A., Olariaga, I., Ozerskaya, S. M., Palma, M. A., Petters-Vandresen, D. A. L., Piontelli, E., Popov, E. S., Rodríguez, A., Requejo, Ó., Rodrigues, A. C. M., Rong, I. H., Roux, J., Seifert, K. A., Silva, B. D. B., Sklenář, F., Smith, J. A., Sousa, J. O., Souza, H. G., De Souza, J. T., Švec, K., Tanchaud, P., Tanney, J. B., Terasawa, F., Thanakitpipattana, D., Torres-García, D., Vaca, I., Vaghefi, N., Iperen, A. L., Vasilenko, O. V., Verbeken, A., Yilmaz, N., Zamora, J. C., Zapata, M., Jurjević, Ž., and Groenewald, J. Z. 2019. Fungal planet description sheets: 951–1041. *Persoonia* 43:223-425.
- Emms, D. M., and Kelly, S. 2015. OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16:157.
- Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otilar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., and Shabalov, I. 2014. MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 42:D699-D704.
- Holt, C., and Yandell, M. 2011. MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491.
- Marin-Felix, Y., Hernández-Restrepo, M., Iturrieta-González, I., García, D., Gené, J., Groenewald, J. Z., Cai, L., Chen, Q., Quaedvlieg, W., Schumacher, R. K., Taylor, P. W. J., Ambers, C., Bonthond, G., Edwards, J., Krueger-Hadfield, S. A., Luangsa-Ard, J. J., Morton, L., Moslemi, A., Sandoval-Denis, M., Tan, Y. P., Thangavel, R., Vaghefi, N., Cheewangkoon, R., and Crous, P. W. 2019. Genera of phytopathogenic fungi: GOPHY 3. *Stud. Mycol.* 94:1-124.
- Newark, M. J., Li, P., Yang, X. P., Paret, M. L., and Dufault, N. S. 2020. Comparing *Stagonosporopsis* spp. fungicide resistance profiles in Florida and East China cucurbit production systems. *Plant Dis.* 104:129-136.
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* 8:785-786.
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210-3212.
- Stanke M, Morgenstern B. 2005. AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 33:W465-W467.
- Vaghefi, N., Pethybridge, S. J., Ford, R., Nicolas, M. E., Crous, P. W., and Taylor, P. W. 2012. *Stagonosporopsis* spp. associated with ray blight disease of *Asteraceae*. *Australas. Plant Pathol.* 41:675-686.
- Verma, S., Gazara, R. K., Nizam, S., Parween, S., Chattopadhyay, D., and Verma, P. K. 2016. Draft genome sequencing and secretome analysis of fungal phytopathogen *Ascochyta rabiei* provides insight into the necrotrophic effector repertoire. *Sci. Rep.* 6:24638.
- Wingfield, B. D., Ades, P. K., Al-Naemi, F. A., Beirn, L. A., Bihon, W., Crouch, J. A., de Beer, Z. W., De Vos, L., Duong, T. A., Fields, C. J., Fourie, G., Kanzi, A. M., Malapi-Wight, M., Pethybridge, S. J., Radwan, O., Rendon, G., Slippers, B., Santana, Q. C., Steenkamp, E. T., Taylor, P. W., Vaghefi, N., van der Merwe, N. A., Veltri, D., and Wingfield, M. J. 2015. IMA Genome-F 4: Draft genome sequences of *Chrysoporthe austroafricana*, *Diplodia scrobiculata*, *Fusarium nygamai*, *Leptographium lundbergii*, *Limonomyces culmigenus*, *Stagonosporopsis tanacetii*, and *Thielaviopsis punctulata*. *IMA Fungus* 6:233-248.