GENOME RESOURCE PAPER (GRP)

Sequence and assembly of the genome of *Seiridium unicorne*, isolate CBS 538.82, causal agent of cypress canker disease

Edoardo Scali[1] · Gianni Della Rocca[2] · Roberto Danti[2] · Matteo Garbelotto[1] · Sara Barberini[2] · Riccardo Baroncelli[3] · Giovanni Emiliani[2]

According to the most recent taxonomy, Cypress Canker Disease (CCD) is caused by seven Ascomycete fungi belonging to the genus *Seiridium*: *S. cardinale* (W. Wagener) B. Sutton & I.A.S. Gibson, *S. cupressi* (Guba) Boesew., *S. unicorne* (Cooke & Ellis) B. Sutton, *S. neocupressi* (Bonthond, Sandoval-Denis & Crous), *S. cancrinum* (Bonthond, Sandoval-Denis & Crous), *S. pseudocardinale* (Wijayaw., Camporesi, McKenzie & K.D. Hyde) and *S. kenyanum* (Bonthond, Sandoval-Denis & Crous) (Bonthond et al. 2018). Differently from other *Seiridium* species which appear to be associated mostly with hosts in the Cupressaceae family, *S. unicorne* has been isolated from a wider range of plant families, including the Anacardiaceae, Caprifoliaceae, Cornaceae, Cupressaceae, Hamamelidaceae, Rosaceae and Vitaceae (Guba 1961; Boeswinkel 1983; Cho and Shin 2004).

Presently, CCD is found in all continents, therefore it is regarded as a remarkable example of a true pandemic caused by a tree disease, but there are differences depending on the number of *Seiridium* species involved in local outbreaks (Danti and Della Rocca 2017). Understanding the evolutionary relationships among the various *Seiridium* species responsible for CCD may provide knowledge necessary for the improved management of the disease (Möller and Stukenbrock 2017). The production of reliable and high quality genomic data is the starting point to achieve this objective. Here, we present the first genomic draft of *S. unicorne* isolate CBS 538.82. This fungal strain was isolated in New Zealand from *Cryptomeria japonica*. The fungus was grown in Petri dishes on cellophane placed on top of Potato Dextrose Agar (PDA). Petri dishes were kept at a constant temperature of 20 °C for 14 days before harvesting the mycelia for DNA extraction performed as follows. First, the mycelium was lyophilized and then, the lyophilized fungal mycelium was manually ground using a mortar and pestle in the presence of liquid nitrogen in the mortar. DNA extraction was performed with the Qiagen DNeasy plant kit, according to manufacturer instructions for genomic fungal DNA (Griffin et al. 2002). DNA fragment integrity was assessed using the Tape Station 4150 (Desjardins and Conklin 2010). Whole Genome Sequencing (WGS) was carried out using the Illumina MiSeq platform (Caporaso et al. 2012) at Genartis SRL (Verona, Italy). The KAPA Hyper Prep kit with a PCR-free protocol was followed for the preparation of the DNA library, according to manufacturer instructions (Whitehorn et al. 2018). A Covaris S220 focused-ultrasonicator (Covaris, Woburn, MA) was used for the fragmentation of 500 ng of DNA. An AMPureXP beads volume of 0.7X (Jackson 2016) was used to generate a WGS DNA library, according to manufacturer instructions (Whitehorn et al. 2018). A capillary electrophoretic analysis on Agilent 4200 Tape station (Walker et al. 2023) was used to assess the quality of the library. The preliminary quantification of the DNA library was performed using the Qubit BR dsDNA assay kit (Thermo Fisher Scientific). A further quantification was performed with the Real-Time PCR KAPA Library Quantification Kit (Kapa Biosystems). DNA Sequencing was performed

✉ Roberto Danti
roberto.danti@ipsp.cnr.it

[1] Department of Environmental Science, Policy and Management (ESPM), University of California, Berkeley, Berkeley, California, USA

[2] Institute for Sustainable Plant Protection (IPSP), National Research Council (CNR), Sesto Fiorentino, Italy

[3] Department of Agricultural and Food Sciences (DISTAL), University of Bologna, Bologna, Italy

Table 1 Summary statistics of *Seiridium unicorne* isolate CBS 538.82 genome assembly

Assembly Variables	Statistics
Assembly length (Mbp)	45.55
Number of contigs	581
Largest scaffold size (bp)	1,097,952
N50	390,618
N90	136,290
L50	36
L90	105
GC (%)	50.25
BUSCO completeness	3,799 (99.6%)
Complete and single-copy	3,793 (99.4%)
Complete and duplicated	6 (0.2%)
Fragmented	1 (0.0%)
Missing	17 (0.4%)
Protein coding genes	
Number of predicted genes	13,576
Number of predicted secreted proteins	1,550
Number of predicted effector proteins	3,155
Number of predicted cytoplasmic effectors	2,678
Number of predicted apoplastic effectors	477

with the NovaSeq 6000 Illumina platform from libraries pooled at equimolar concentrations. The quality of raw reads was determined with FastQC v0.11.7 (Andrews

2010). Finally, the genome was assembled using SPAdes v3.15.5 (Bankevich et al. 2012).

The assembled genomic draft consisted of 581 contigs (N50 = 390,618, N90 = 136,290). The integrity of the genome assembly was evaluated with BUSCO v3.0.2, using the Sordariomycetes odb10 model dataset (Manni et al. 2021). Assembly statistics were produced using QUAST v5.2.0 (Gurevich et al. 2013) and results are reported in Table 1. Gene prediction was performed with MAKER v3.01.03 (Cantarel et al. 2008), using both Augustus (Stanke and Morgenstern 2005) and Genemark-ES (Lukashin 1998). Augustus gene prediction was performed using the *Fusarium* spp. model. A total of 13,576 protein-coding genes were predicted. The presence of eukaryotic amino acid signal peptides was assessed using SignalP v5.0 (Almagro Armenteros et al. 2019). A total number of 1,550 secreted proteins was predicted. Fungal effector prediction was performed with EffectorP-fungi 3.0 (Sperschneider and Dodds 2022). The total number of predicted effector proteins was 3,155 and, among these, 2,678 were classified as cytoplasm predicted effectors and 477 were classified as apoplastic predicted effectors. A phylogenetic inference of orthologs was made with OrthoFinder (Emms and Kelly 2019) by comparing

Fig. 1 A phylogenetic tree obtained using OrthoFinder. Predicted protein-coding genes from the genome draft of *S. unicorne* CBS 538.82 were compared with published predicted proteins from fungi belonging to the order Xylariales and to the *Sporocadaceae* family. *Botrytis cinerea* and *Diplodia corticola* were selected as outgroups. *S. unicorne* CBS 538.82 is identified by the bold text and the asterisk at the end. The *Seiridium* species cluster is highlighted in red
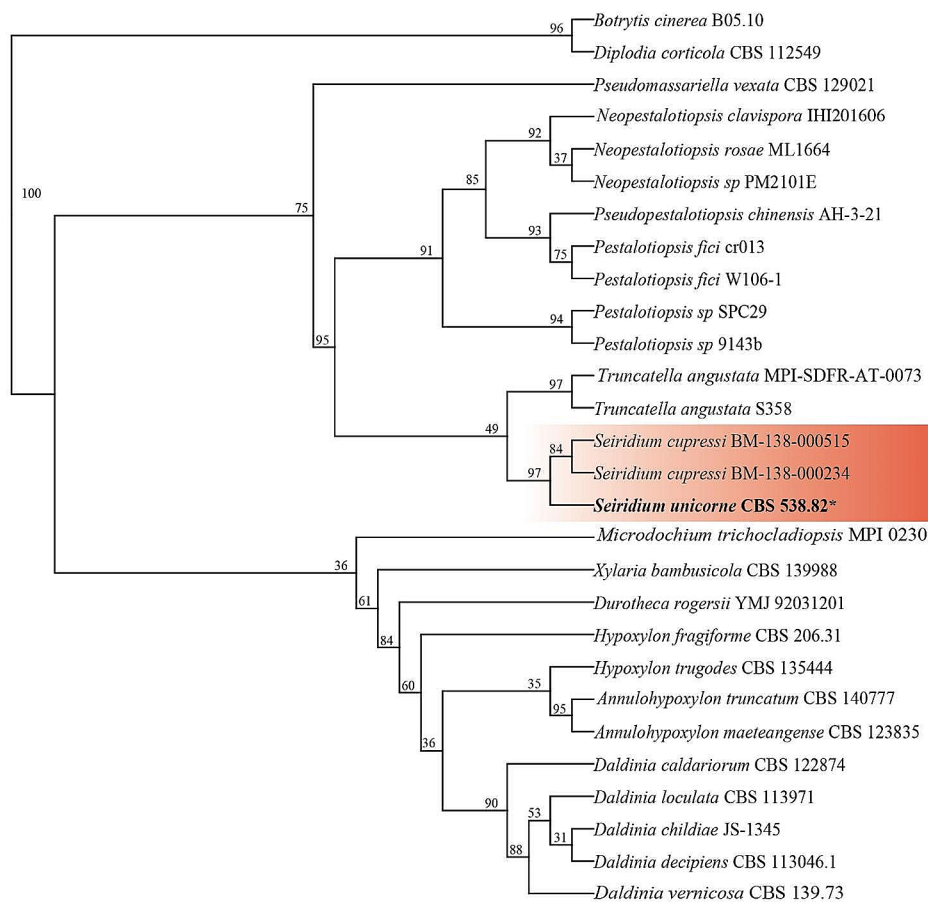
proteins from available Xylariales, *Sporocadaceae* and from two isolates of *Seiridium cupressi* published in a previous genome announcement (Scali et al. 2023).

Phylogenomic relationship among select fungal species (Fig. 1), including *S. unicorne* and its known close relatives, has been constructed inferred with STAG (Emms and Kelly 2018).

A comparison of gene prediction results between *S. unicorne* CBS 538.82 and *S. cupressi* (Scali et al. 2023), indicates that the number of predicted effector proteins is significantly greater in the first species. Predicted effectors for *S. unicorne* CBS 538.82 represent 23.24% of the total predicted protein coding genes, while the percentages of predicted effectors were 3.95% and 3.92% for *S. cupressi* isolate BM-138-000515 and *S. cupressi* isolate BM-138-000234, respectively. This result is interesting and may be explained by the wider host range of *Seiridium unicorne*, compared to that of *S. cardinale* and *S. cupressi* (Guba 1961; Boesewinkel 1983; Cho and Shin 2004). Further investigations will allow us to better understand the co-evolutionary dynamics underlying this difference in the number of effectors (Zhao et al. 2022).

Data availability All data are deposited and publicly available using the accession numbers provided above.

Declarations

Conflict of interest Authors declare no conflict of interest.

References

Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, Von Heijne G, Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37(4):420–423. https://doi.org/10.1038/s41587-019-0036-z

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19(5):455–477

Boesewinkel H (1983) New records of the three fungi causing cypress canker in New Zealand, Seiridium Cupressi (Guba) comb. Nov. And S. Cardinale on Cupressocyparis and S. unicorne on Cryptomeria and Cupressus. Trans Br Mycological Soc 80(3):544–547

Bonthond G, Sandoval-Denis M, Groenewald JZ, Crous PW (2018) Seiridium (Sporocadaceae): an important genus of plant pathogenic fungi. Persoonia - Mol Phylogeny Evol Fungi 40(1):96–118. https://doi.org/10.3767/persoonia.2018.40.04

Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M (2008) MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res 18(1):188–196. https://doi.org/10.1101/gr.6743907

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6(8):1621–1624. https://doi.org/10.1038/ismej.2012.8

Cho W-D, Shin H-D (2004) List of plant diseases in Korea. Korean Society of Plant Pathology

Danti R, Della Rocca G (2017) Epidemiological history of Cypress Canker Disease in Source and Invasion sites. Forests 8(4):121. https://doi.org/10.3390/f8040121

Desjardins P, Conklin D (2010) NanoDrop microvolume quantitation of nucleic acids. JoVE (Journal Visualized Experiments) 45:e2565

Emms DM, Kelly S (2018) STAG: Species Tree Inference from All Genes [Preprint]. Evolutionary Biology. https://doi.org/10.1101/267914

Emms DM, Kelly S (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol 20:1–14

Griffin DW, Kellogg CA, Peak KK, Shinn EA (2002) A rapid and efficient assay for extracting DNA from fungi. Lett Appl Microbiol 34(3):210–214. https://doi.org/10.1046/j.1472-765x.2002.01071.x

Guba E (1961) Monograph of Monochaetia and. Pestalotia. Harvard

Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29(8):1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Jackson B (2016) Clean-up using AMPure XP beads

Lukashin A (1998) GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res 26(4):1107–1115. https://doi.org/10.1093/nar/26.4.1107

Manni M, Berkeley MR, Seppey M, Zdobnov EM (2021) Curr Protocols 1(12). https://doi.org/10.1002/cpz1.323. BUSCO: Assessing Genomic Data Quality and Beyond

Möller M, Stukenbrock EH (2017) Evolution and genome architecture in fungal plant pathogens. Nat Rev Microbiol 15(12):756–771. https://doi.org/10.1038/nrmicro.2017.76

Scali E, Della Rocca G, Danti R, Garbelotto M, Barberini S, Frascella A, Emiliani G (2023) Genome sequence and assembly of the causal agent of Cypress Canker Disease *Seiridium Cupressi*, isolates BM-138-000234 and BM-138-000515. Journal of Plant Pathology, in press

Sperschneider J, Dodds PN (2022) EffectorP 3.0: prediction of apoplastic and cytoplasmic effectors in Fungi and Oomycetes. Mol Plant-Microbe Interactions® 35(2):146–156. https://doi.org/10.1094/MPMI-08-21-0201-R

Stanke M, Morgenstern B (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res 33(suppl2):W465–W467

Walker JE, Oliver JC, Stewart AM, Beh ST, Arce RA, Glass MJ, Vargas DE, Qiji SH, Intorcia AJ, Borja CI, Cline MP, Hemmingsen SJ, Krupp AN, McHattie RD, Mariner MR, Lorenzini I, Aslam S, Tremblay C, Beach TG, Serrano GE (2023) Measuring up: a comparison of TapeStation 4200 and Bioanalyzer 2100 as Measurement Tools for RNA quality in Postmortem Human Brain samples. Int J Mol Sci 24(18):13795. https://doi.org/10.3390/ijms241813795

Whitehorn H, Appel M, Solutions RS (2018) KAPA HyperPrep Kits offer a flexible. high-efficiency library preparation solution for PCR-free human whole-genome sequencing

Zhao Y, Zhu X, Chen X, Zhou J-M (2022) From plant immunity to crop disease resistance. J Genet Genomics 49(8):693–703. https://doi.org/10.1016/j.jgg.2022.06.003

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.