Journal of Plant Pathology
https://doi.org/10.1007/s42161-023-01545-4

GENOME RESOURCE PAPER (GRP)

Genome sequence and assembly of the causal agent of Cypress Canker Disease *Seiridium cupressi,* isolates BM-138–000234 and BM-138–000515

Edoardo Scali[1] · Gianni Della Rocca[2] · Roberto Danti[2] · Matteo Garbelotto[1] · Sara Barberini[2] · Angela Frascella[2] · Giovanni Emiliani[2]

Received: 13 June 2023 / Accepted: 30 October 2023

Keywords Comparative genomics · Emergent disease · Fungal genome · Plant pathogen

Seiridium cupressi (Nattrass, C. Booth & B. Sutton) Bonthond, Sand.-Den. & Crous, is an Ascomycete fungus belonging to the family Sporocadaceae. Together with other six *Seiridium* species, including *S. cardinale* (W. Wagener) B. Sutton & I.A.S. Gibson and *S. unicorne* (Cooke & Ellis) B. Sutton among them, *S. cupressi* is one of the causal agents of the Cypress Canker Disease pandemic (CCD) (Graniti 1998; Danti et al. 2013; Bonthond et al. 2018). The phytotoxic activity of *S. cupressi* has been studied, including the toxins that this species is able to produce, such as Cyclopaldic acid (Graniti et al. 1992; Sparapano & Evidente 1995) and Seiricuprolide (Ballio et al. 1988; Bartolucci et al. 1992). Nevertheless, genomic-level information that may enable us to better understand gene expression and the pathogenicity of CCD causal agents is still missing. In this study, two isolates of *S. cupressi*, BM-138–000515 and BM-138–000234, were retrieved from the Italian National Research Council of Sesto Fiorentino (Florence, Italy) CNR-IPSP-FTFC mycological collection (https://biomemory.cnr.it/collections/CNR-IPSP-FTFC/). These two isolates have been cultured from symptomatic tissue of *Cupressus sempervirens* located in Kos island (Grece) in 1984.

Each isolate was grown onto a cellophane disk placed on Petri dishes filled with Potato Dextrose Agar (PDA) for 14 days, at a constant temperature of 20 °C. Mycelia were collected, lyophilized and ground in liquid nitrogen using a sterile mortar and pestle. DNA was extracted using the Qiagen DNeasy plant kit, following the protocol provided by the manufacturer (Griffin et al. 2002). Quality and quantity of genomic DNA were assessed by gel electrophoresis and nanodrop (Desjardins & Conklin 2010). Short read sequencing was performed by Genartis SRL (Verona, Italy) using Illumina MiSeq (Caporaso et al. 2012). DNA libraries for these two samples were prepared with a KAPA Hyper Prep kit with a PCR-free protocol according to the manufacturer's instruction (Whitehorn et al. 2018). A total of 500 ng of DNA was fragmented by sonication using a Covaris S220 focused-ultrasonicator (Covaris, Woburn, MA). A Whole Genome Sequencing DNA library was generated according to the manufacturer's instruction (Roche), using 0.7X volume of AMPureXP beads (Jackson 2016) for the final size selection. Quality and size of WGS libraries were assessed by capillary electrophoretic analysis using the Agilent 4200 Tape station (Agilent Technologies). Libraries were preliminarily quantified using the Qubit BR dsDNA assay kit (Thermofisher) and then they were further quantified by Real-Time PCR, using the standard curve of the KAPA Library Quantification Kit (Kapa Biosystems). Libraries were pooled at equimolar concentrations and sequenced on a NovaSeq 6000 Illumina platform. DNA purity was measured using a NanoDrop Spectrophotometer, while DNA integrity was assessed using the Tape Station 4150 (Agilent Technologies). DNA quantity was measured using the Qubit dsDNA BR assay (ThermoFisher). DNA reads quality was determined using FastQC v0.11.7 (Andrews, 2010). Illumina reads were assembled using SPAdes v3.15.5 (Bankevich et al. 2012). The first draft assembled for isolate BM-138–000515 consisted of 795

✉ Matteo Garbelotto
matteog@berkeley.edu

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

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

Published online: 05 December 2023

Table 1 Summary statistics for *Seiridium cupressi*, isolates BM-138–000234 and BM-138–000515

Assembly Variables	BM-138–000515	BM-138–000234
Assembly length (bp)	46,207,062	45,973,548
Number of contigs	795	940
Largest contigs size (bp)	964,279	1,146,263
N50	292,075	359,772
N75	74,854	96,369
L50	48	40
L75	165	131
GC (%)	49.64	49.79
BUSCO completeness	98.6%	99.3%
Complete and single-copy	98.4%	99.1%
Complete and duplicated	0.2%	0.2%
Fragmented	0.1%	0.0%
Missing	1.3%	0.7%
Protein coding genes		
Number of predicted genes	13,376	13,446
Number of predicted secreted proteins	1,502	1,504
Number of predicted effector proteins	529	528
Number of predicted cytoplasmic effectors	177	175
Number of predicted apoplastic effectors	352	353

contigs (N50 = 289,630 bp and N75 = 146,648 bp), while isolate BM-138–000234 had 970 contigs (N50 = 359,772 bp and N75 = 196,592). In order to assess the genome assembly and integrity, BUSCO v.3.0.2 (Seppey et al. 2019) was utilized 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 and genome annotation have been performed with MAKER3 v3.01.03 (Cantarel et al. 2008). A total of 13,446 protein coding genes were obtained for the isolate BM-138–000234, while a total number of 13,376 protein coding genes were predicted for isolate BM-138–000515.

SignalP v5.0 (Almagro Armenteros et al. 2019), designed for the prediction of signal peptides, predicted 1,504 secreted proteins for BM-138–000515 isolate and 1,504 secreted proteins for BM-138–000234 isolate. Effector prediction analyses, carried out with EffectorP-fungi v3.0 (Sperschneider & Dodds 2022), predicted 528 secreted proteins may be effectors in BM-138–000234, including 175 cytoplasmic effectors and 353 apoplastic effectors. Regarding isolate BM-138–000515, 529 candidate effectors were predicted, including 177 cytoplasmic effectors and 352 apoplastic effectors. A phylogenomic comparative analysis was performed following the methods described in Baroncelli et al. (2022). Both isolates of *S. cupressi* were compared to phylogenetically related species belonging to the Sporocadaceae family and to the order Xylariales (Fig. 1). The genome of *S. cupressi* is the first *Seiridium* genome published, as well as the first genomic resource published regarding the CCD pathosystem, and will assist the scientific community in furthering its understanding of the mechanisms that make this disease a true world pandemic.

GenBank accession number

This whole-genome shotgun project has been deposited in GenBank. Genomic sequences of *S. cupressi* isolate BM-138–000234 have been deposited under the accession number JARVKI000000000 (BioProject: PRJNA953536; BioSample: SAMN34119420). Genomic sequences of *S. cupressi* isolate BM-138–000515 have been deposited under the accession number JARVKH000000000 (BioProject: PRJNA953536; BioSample: SAMN34119421).

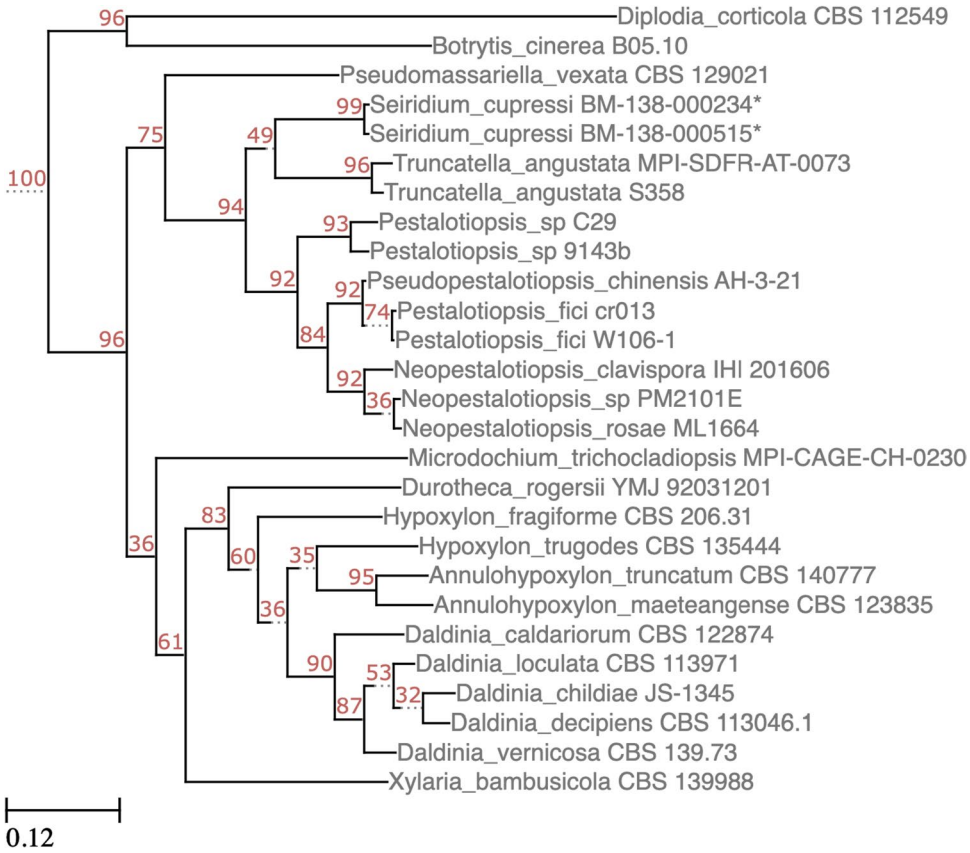

Fig. 1 Phylogenetic tree comparing the newly sequenced isolates of *S. cupressi* BM-138–000515 and BM-138–000234 with those of closely available fungi belonging to Xylariales order and Sporocadaceae family. *Botrytis cinerea* and *Diplodia corticola* have been selected as outgroups. Asterisks identify the isolates whose genomes have been sequenced in this study

Data Availability All data are deposited and publicly available using the accession numbers provided below. This is clearly stated at the end of the manuscript.

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):20–423. https://doi.org/10.1038/s41587-019-0036-z

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

Ballio A, Evidente A, Graniti A, Randazzo G, Sparapano L (1988) Seiricuprolide, a new phytotoxic macrolide from a strain of *Seiridium cupressi* infecting cypress. Phytochemistry 27(10):3117–3121

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

Baroncelli R, Cobo-Díaz JF, Benocci T, Peng M, Battaglia E, Haridas S, Andreopoulos W, LaButti K, Pangilinan J, Lipzen A, Koriabine M, Bauer D, Floch GL, Mäkelä MR, Drula E, Henrissat B, Grigoriev IV, Crouch JA, De Vries RP, Thon MR (2022) Genome evolution and transcriptome plasticity associated with adaptation to monocot and eudicot plants in Colletotrichum fungi [Preprint]. Microbiology. https://doi.org/10.1101/2022.09.22.508453

Bartolucci C, Cerrini S, Lamba D, Evidente A, Randazzo G (1992) Absolute configuration of seiricuprolide, a new phytotoxin from Seiridium cupressi. Acta Crystallogr C 48(1):83–86

Bonthond G, Sandoval-Denis M, Groenewald JZ, Crous PW (2018) *Seiridium* (*Sporocadaceae*): An important genus of plant pathogenic fungi. Persoonia - Molecular Phylogeny and Evolution of 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

Danti R, Dela Rocca G, Panconesi A (2013) Cypress canker. Infectious forest diseases. CABI, Wallingford UK, pp 359–375

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

Graniti A (1998) Cypress canker: A pandemic in progress. Annu Rev Phytopathol 36(1):91–114

Graniti A, Sparapano L, Evidente A (1992) Cyclopaldic acid, a major phytotoxic metabolite of Seiridium cupressi, the pathogen of a canker disease of cypress. Plant Pathol 41(5):563–568

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

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. Bioline. https://dx.doi.org/10.17504/protocols.io.f3ebqje

Manni M, Berkeley MR, Seppey M, Zdobnov EM (2021) BUSCO: Assessing genomic data quality and beyond. Curr Protoc 1(12). https://doi.org/10.1002/cpz1.323

Seppey M, Manni M, Zdobnov EM (2019) BUSCO: Assessing genome assembly and annotation completeness. In M. Kollmar (Ed.), Gene Prediction (Vol. 1962, pp. 227–245). Springer New York. https://doi.org/10.1007/978-1-4939-9173-0_14

Sparapano L, Evidente AA (1995) Biological activity of cyclopaldic acid, a major toxin of Seiridium cupressi, its six derivatives, and isocyclopaldic acid. Nat Toxins 3(3):156–165

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

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