Journal of Plant Pathology (2024) 106:1389–1392
https://doi.org/10.1007/s42161-024-01665-5

GE□□ME RES□□R□E □□□ER □GR□□

Genome assembly of two California isolates of *Seiridium cardinale* (BM-138-000494 and BM-138-000479)

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

Received: 10 April 2024 / Accepted: 29 May 2024 / Published online: 17 June 2024

Cypress Canker Disease (CDD) is a pandemic affecting plants in the family Cupressaceae (Danti and Della Rocca 2017). CCD is caused by at least seven fungal species belonging to the genus *Seiridium* (Bonthond et al. 2018). Among these species, *S. cardinale* (Wag.) Sutt. et Gibs., (Ascomycota, Sordariomycetes, *Sporocadaceae*), has been the most studied (Graniti 1998) and has proven to be the most virulent both in California (USA), where the pathogen is likely to be native, and in the Mediterranean region, where the pathogen was introduced and has become invasive (Della Rocca et al. 2013). The two isolates of *S. cardinale* here described, BM-138-000479 and BM-138-000494, were originally isolated in California from symptomatic tissues of two *xCupressocyparis leylandii* trees. Isolate BM-138-000479 was isolated in San Joaquin County, while isolate BM-138-000494 was isolated in Marin county (California, US). Isolates BM-138-000479 and BM-138-000494 are representative of the β-tubulin haplotype A and B groups of the pathogen, respectively (Della Rocca et al. 2011). These two groups are clearly differentiated at the genomic level, as shown by AFLP analysis (Della Rocca et al. 2013).

Sequencing and bioinformatics analyses were conducted as described in two previous genome announcements (Scali et al. 2023, 2024 in press) to produce a first draft of genome assembly for both isolates. In brief, fungal cultures were grown on a cellophane layer on top of Potato Dextrose Agar (PDA) for 14 days at 20 °C. Mycelia were then collected, lyophilized, and ground in liquid nitrogen using a sterile mortar and pestle. DNA was extracted using the Qiagen DNeasy plant kit, following the manufacturer's instruction for lyophilized fungal cultures (Griffin et al. 2002). Gel electrophoresis and a nanodrop spectrophotometer (Desjardins and Conklin 2010) were then employed to assess the quality and the quantity of the genomic DNA. Short reads sequencing was performed with the Illumina MiSeq platform (Caporaso et al. 2012) by Genartis SRL (Verona, Italy). The KAPA Hyper Prep kit with a PCR-free protocol was employed for the preparation of DNA libraries (Whitehorn et al. 2018). DNA fragmentation was achieved by sonication of 500 ng of DNA, using a Covaris S220 focused-ultrasonicator (Covaris, Woburn, MA). Before conducting Whole Genome Sequencing, a preparatory step involved treating the DNA with a 0.7X volume of AMPureXP beads (Jackson 2016), in accordance with the manufacturer's instructions (Roche). Agilent 4200 Tape stations (Agilent technologies) were used to assess the quality and size of WGS libraries. The Qubit dsDNA assay kit (Thermofisher) and Real-Time PCR (Kapa Biosystems) were used for preliminary DNA quantification. DNA libraries were pooled at equimolar concentrations and were then sequenced on a NovaSeq 6000 Illumina platform. Quality checks of short reads were performed with fastQC v0.11.7 (Andrews 2010). Raw short-reads were then polished by removing the adapters using Cutadapt v2.4 (Martin 2011). SPAdes v3.15.5 (Bankevich et al. 2012) was used to assemble the Illumina short-reads. The first genomic assembly draft for BM-138-000494 gave a total number of 257 contigs with a size > = 500 bp (N50 = 532,845 and N90 = 137,466). A total of 252 contigs 500 bp or more in size (N50 = 483,201, N90 = 168,572) were assembled instead for isolate BM-138-000479. A quality assessment of the assembled genomes was performed with BUSCO v.3.0.2, using the Sordariomycetes odb10 model dataset (Manni et al. 2021). QUAST v5.2.0 (Gurevich et al. 2013) produced the assembly statistics for both isolates

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

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

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

Table 1 Summary statistics for *Seiridium cardinale*, isolates BM-138-000494 and BM-138-000479

Assembly variables	BM-138-000494	BM-138-000479
Assembly length (bp)	46,140,912	46,205,179
Number of contigs > = 500 bp	257	252
Largest contig size (bp)	1,425,233	2,139,086
N50	532,845	483,201
N90	137,466	168,572
L50	29	30
L90	93	92
GC content	50.27%	50.2%
BUSCO completeness	98.3%	98.2%
Complete and single-copy	98.2%	98.1%
Complete and duplicate	0.1%	0.1%
Fragmented	0.4%	0.4%
Missing	1.3%	1.4%
Protein coding genes		
Number of predicted genes	14,281	14,050
Number of predicted secreted proteins	1536	1520
Number of predicted effector proteins	3398	3323
Number of predicted cytoplasmic effector	2883	2821
Number of predicted apoplastic effector	515	502

and results are reported in Table 1. Annotation of fungal genomes and prediction of protein coding genes were performed with Maker v3.01.03 (Cantarel et al. 2008). A total number of 14,281 protein coding genes were predicted for *S. cardinale* isolate BM-138-000494, while 14,050 protein coding genes were predicted for isolate BM-138-000479. Predicted proteins were further annotated with SignalP v6.0 (Teufel et al. 2022). A total of 1536 and 1520 secreted proteins were predicted for isolates BM-138-000494 and BM-138-000479, respectively. Analyses of effectors were performed with EffectorP-fungi v3.0 (Sperschneider and Dodds 2022) and resulted in the prediction of 3398 effectors, 2883 cytoplasmic effectors and 515 apoplastic effectors for isolate BM-138-000494. For isolate BM-138-000479, 3323 effectors, 2821 cytoplasmic effectors, and 502 apoplastic effectors were predicted. Phylogenetic analyses were performed by clustering the proteomes of these two species with the publicly available proteomes of other *Seiridium* species (Scali et al. 2023, 2024 in press) and of other fungi in the order Xylariales. The proteomes of *Botrytis cinerea* and *Diplodia corticola* were selected as outgroups. Results of this analysis are shown in Fig. 1. The cluster of the three *Seiridium* spp., highlighted in red, appears to be monophyletic and inclusive of two distinct subclades: one contains *S. cardinale*, and the other includes both *S. unicorne* and *S. cupressi*, which thus appear to be sister taxa, more distantly related to *S. cardinale* than to each other.

GenBank accession number

Both assembled genomic drafts have been deposited in GenBank. Genomic sequence for the isolate BM-138-000494 has been deposited under accession number JARVKK000000000 (BioProject: PRJNA953536, BioSample: SAMN34119418). Genomic sequence for isolate BM-138-000479 has been deposited under accession number JARVKM000000000 (BioProject: PRJNA953536, BioSample: SAMN34119416). For both isolates, GeneBank submission has been accompanied with the submission of the genome annotation.

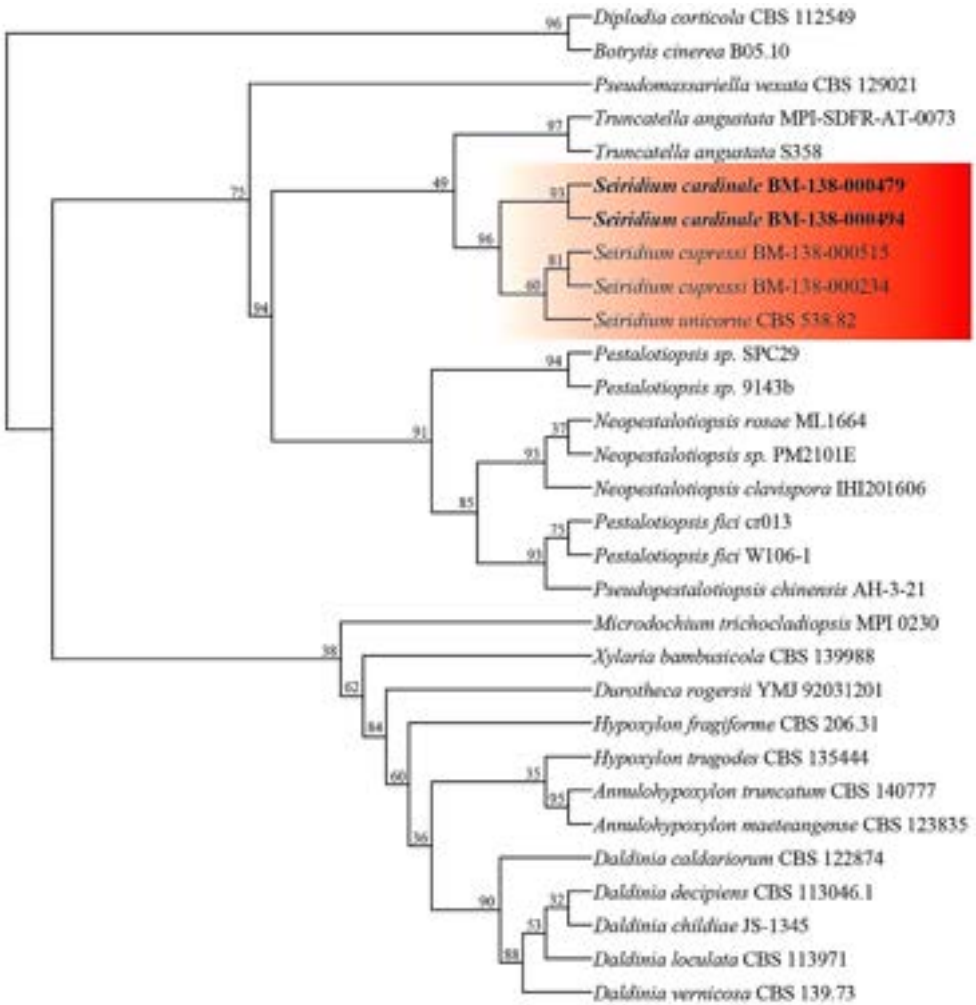

Fig. 1 Phylogenetic tree inferred with OrthoFinder. *S. cardinale* isolates BM-138-000479 and BM-138-000494 have been highlighted in bold. Clusters of *Sieridium* species have been highlighted with a red label

Funding Open access funding provided by IPSP - BARI within the CRUI-CARE Agreement.

Data Availability All data are deposited and publicly available using the accessionnumbers provided above.

Conflict of Interest Authors declare no conflict of interest.

References

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

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

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

Della Rocca G, Eyre CA, Danti R, Garbelotto M (2011) Sequence and simple-sequence repeat analyses of the Fungal Pathogen *Seiridium Cardinale* Indicate California is the most likely source of the Cypress Canker Epidemic for the Mediterranean Region. Phytopathology® 101(12):1408–1417. https://doi.org/10.1094/PHYTO-05-11-0144

Della Rocca G, Osmundson T, Danti R, Doulis A, Pecchioli A, Donnarumma F, Casalone E, Garbelotto M (2013) AFLP analyses of California and Mediterranean populations of *Seiridium Cardinale* provide insights on its origin, biology and spread pathways. Forest Pathol 43(3):211–221. https://doi.org/10.1111/efp.12019

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

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

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

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17(1):10. https://doi.org/10.14806/ej.17.1.200

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

Scali E, Rocca GD, Danti R, Garbelotto M, Barberini S, Baroncelli R, Emiliani G (2024) Sequence and assembly of the genome of Seiridium unicorne, isolate CBS 538.82, causal agent of Cypress Canker Disease. Journal of Plant Pathology. (in press)

Scali E, Rocca GD, 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. J Plant Pathol. https://doi.org/10.1007/s42161-023-01545-4

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

Teufel F, Armenteros A, Johansen JJ, Gíslason AR, Pihl MH, Tsirigos SI, Winther KD, Brunak O, von Heijne S, G., Nielsen H (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40(7) Article 7. https://doi.org/10.1038/s41587-021-01156-3

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