

Research Paper

Whole-Genome sequencing and phylogenomic analyses of *Alternaria japonica* and *A. telliensis* (Section *Japonicae*)

Roumaïssa Barkat^{1,2}, Nabahat Bessadat^{1,2}, Justine Colou¹, Philippe Simoneau¹✉

1. University of Angers, Institut Agro Rennes Angers, INRAE, UMR 1345 IRHS, SFR 4207 QUASAV, Beaucouzé Cedex, 49070, France.

2. Laboratory of Applied Microbiology, University Oran1 Ahmed Ben Bella, BP 1524, El M'Naouer 31000, Oran, Algeria.

✉ Corresponding author: Philippe Simoneau (simoneau@univ-angers.fr).

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <https://ivyspring.com/terms> for full terms and conditions.

Received: 2026.08.11; Accepted: 2026.09.11; Published: 2026.10.09

Abstract

Species of *Alternaria* are necrotrophic fungi responsible for leaf spot diseases on a wide range of cultivated and wild plants. Species delimitation within this genus, however, often relies on morphological features and a limited set of molecular markers, which can constrain phylogenetic resolution at finer taxonomic scales. This is the case for *Alternaria telliensis*, a species recently described from Algeria and subsequently reported in Iran, whose relationships within section *Japonicae* require clarification. To provide a higher resolution framework for this species, we performed whole-genome sequencing, assembly, and BUSCO completeness assessment for type strains of *A. telliensis* (CBS 145643) and *A. japonica* (CBS 118390), using Oxford Nanopore Technologies MinION sequencing. The genome assembly of *A. telliensis* spans 34.4 Mbp across 132 contigs, with an N50 of 2.2 Mbp and a GC content of 50.94%. The *A. japonica* assembly comprises 33.5 Mbp in 196 contigs, with an N50 of 1.6 Mbp and a GC content of 51.47%. Phylogenomic analysis places *A. telliensis* alongside *A. japonica* within section *Japonicae*. These genomic data support the classification of *A. telliensis* and contribute to clarifying taxonomic relationships within the genus *Alternaria*.

Keywords: *Alternaria telliensis*, *Alternaria japonica*, nanopore sequencing, genome assembly

Introduction

Alternaria telliensis was originally described from *Solanaceae* in northwestern Algeria and subsequently reported as a causal agent of cabbage leaf spot in Iran [1,2]. Preliminary characterization of this species relied on morphological analyses combined with multi-locus phylogenies (ITS, GPD, RPB2, TEF1), which placed it within section *Japonicae*, alongside *A. japonica*. However, when mono-locus phylogenies based on ITS or GPD were applied, *A. telliensis* grouped with members of section *Ulocladioides*.

Whole-genome sequencing provides a higher-resolution framework for addressing the taxonomic ambiguity. Phylogenomic analyses based on single-copy orthologous genes can capture broader evolutionary signal than analyses based on a few loci. The genus *Alternaria* was initially divided into 24 sub-generic sections [3] and now comprises 30 sections and 9 monotypic lineages. More than 280 *Alternaria* genome sequences are publicly available in NCBI

database. However, 73% are from members of section *Alternaria* and the remaining are divided into ten additional sections and one monotypic lineage. Unfortunately, no genome sequence from species of section *Japonicae* is currently available although species within this section might represent a threat for Brassicaceae crops.

This study aimed to generate whole-genome assemblies of the type strain of *A. telliensis* (CBS 145643) and *A. japonica* (CBS 118390) and reconstruct their phylogenomic relationships with selected *Alternaria* species.

Materials and Methods

Genomic DNA was extracted from frozen mycelium cultured for 7 days on Potato Dextrose Agar (PDA) medium on a pectocellulosic membrane after bead grinding, following a high-molecular weight

fungal DNA extraction protocol for filamentous fungi [4]. DNA quality and purity were assessed using a NanoDrop spectrophotometer and DNA concentration was confirmed by fluorometric quantification using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific).

Whole-genome sequencing was performed using the Oxford Nanopore Technologies MinION (ONT) platform with the Native Barcoding Kit V14 (SQK-NBD114.24) and the flowcell FLO-MIN114 with R10.4 pores [5]. The basecalling was performed on MinKNOW (v25.03.9) with the Dorado software (v7.8.3). Nanopore reads were filtered for quality and length using NanoFilt with a minimum quality threshold of $Q \geq 10$ and a minimum read length of 1,000 bp. Read quality was assessed using NanoStat and NanoPlot from the NanoPack toolkit [6]. De novo assembly was performed with Flye (v2.9.6) using raw long reads, an estimated genome size of 31 Mb and three iterations [7]. The initial assemblies were improved through three successive polishing cycles with Racon (v1.5.0) [8]. For each polishing cycle, reads were aligned to the assembly using minimap2 (v2.24) in Oxford Nanopore mode [9].

Assembly quality before and after polishing was evaluated using QUAST v5.3.0 [10]. Assembly completeness was estimated using BUSCO (v6.0.0) against the dothideomycetes_odb10 lineage dataset (3,786 BUSCO groups) [11], using miniprot as the gene predictor. The assemblies of *A. telliensis* and *A. japonica* contained 3,766 (99.5%) and 3,769 (99.6%) complete BUSCOs, respectively. Overall assembly statistics were generated using SeqKit [12].

Structural gene prediction was performed using Funannotate v1.8.17. The soft-masked genome assembly was used as input. Protein evidence was generated by aligning a reference protein database containing 560,728 proteins against the genome using DIAMOND and Exonerate. Ab initio gene predictors (Augustus, GlimmerHMM, and SNAP) were trained on BUSCO derived gene models (fungi_odb10 lineage dataset) identified within the assembly. Gene model consensus was generated using EvidenceModeler (EVM), combining predictions from Augustus (weight = 1), high-quality Augustus predictions (weight = 2), GlimmerHMM (weight = 1), and SNAP (weight = 1). Gene models shorter than 50 amino acids, spanning assembly gaps, or corresponding to transposable elements were removed. tRNA genes were predicted using tRNAscan-SE.

A comparative genomic dataset was constructed using selected *Alternaria* genome assemblies retrieved from the NCBI GenBank database. The dataset included representatives of ten *Alternaria* sections

(plus one monotypic lineage) and *Stemphylium lycopersici* (GCA 003268315.1) as outgroup. The assemblies used for comparison are listed in Table S1. All comparative analyses were performed on the concatenated alignment of 5,399 single-copy orthologous genes in the comparative dataset. Orthologous gene clustering was performed using OrthoFinder (v2.5.2) [13]. Protein sequence similarity searches were conducted with DIAMOND [14], and orthogroup delineation used Markov clustering of protein similarity relationships [15]. Multiple sequence alignments were generated for each inferred orthogroup using MAFFT v7 [16].

The species phylogeny was reconstructed under the maximum-likelihood framework using IQ-TREE v2.2.2.7 [17]. The best-fit substitution model was Q.plant+F+I+R8, selected according to the Bayesian Information Criterion in ModelFinder [18]. Branch support was assessed using ultrafast bootstrap analysis with 1,000 replicates and the SH-aLRT test with 1,000 replicates [19,20].

Results and Discussion

Nanopore read distributions were visualized for *Alternaria telliensis* and *A. japonica* using NanoPlot (Figure S1). The Nanopore read dataset for *A. telliensis* contained 382,467 reads and 1,260,768,324 bases, whereas the *A. japonica* dataset contained 570,917 reads and 1,780,512,442 bases. Estimated sequencing coverages were approximately 36.6x for *A. telliensis* and 53.1x for *A. japonica* (Table 1). The two datasets had similar mean read qualities, with Q22.5 for *A. telliensis* and Q22 for *A. japonica*.

The *A. telliensis* genome assembly had a total length of 34,408,234 bp distributed across 132 contigs, with a GC content of 50.94%. Its largest contig was 3,841,795 bp, and its N50 and N90 were 2,186,759 bp and 525,007 bp, respectively. The *A. japonica* assembly had a total length of 33,509,493 bp distributed across 196 contigs, with a GC content of 51.47%. Its largest contig was 5,995,486 bp, and its N50 and N90 were 1,655,298 bp and 351,347 bp, respectively. BUSCO analyses revealed highly complete assemblies, the *A. telliensis* genome contained 99.5% complete BUSCOs, whereas the *A. japonica* genome contained 99.6% complete BUSCOs. Neither assembly contained ambiguous bases, with N's per 100 kbp equal to zero. These results are consistent with what has been reported for species of the genus *Alternaria* whose genome sizes varied from 31 Mbp (*A. brassicicola*) to 40 Mbp (*A. metachromatica*) with an estimated number of genes varying from 10,000 (*A. dauci*) to 13,000 (*A. linariae*) [21].

Table 1 Assembly statistics for *Alternaria telliensis* and *A. japonica*.

Strain	Contigs	Largest contig (bp)	Total length (bp)	GC content (%)	N50 (bp)	Coverage(x)	Number of predicted genes
<i>A. japonica</i>	196	5,995,486	33,509,493	51.47	1,655,298	53.1	9,522
<i>A. telliensis</i>	132	3,841,795	34,408,234	50.94	2,186,759	36.6	9,703

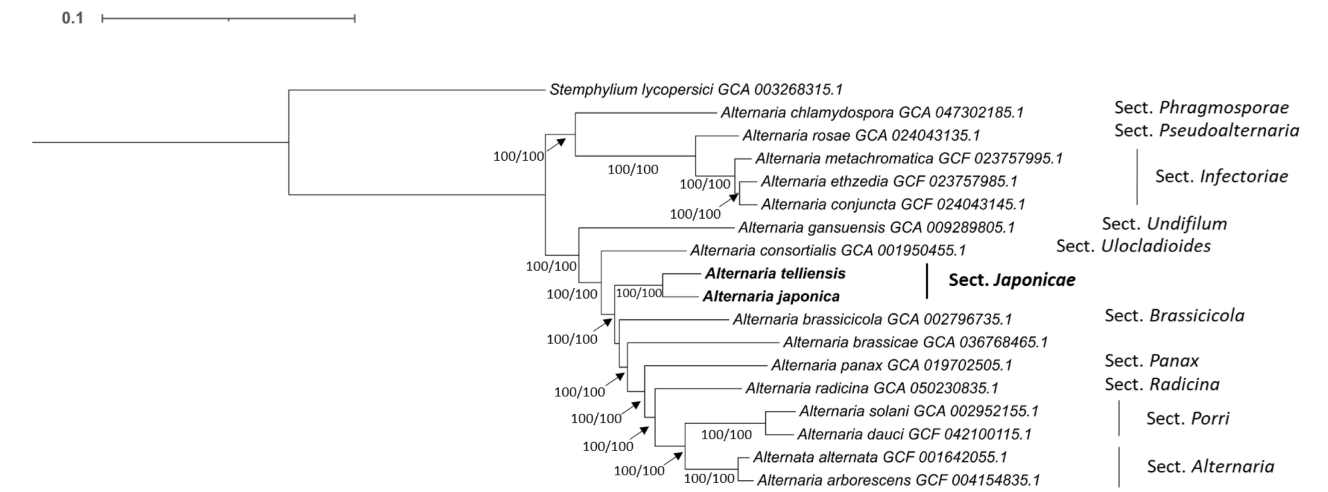


Figure 1. Phylogenomic tree constructed by the maximum likelihood (ML) method from the concatenated alignment of single-copy orthologous genes, illustrating the relationships between the different species of the *Alternaria* sections. *Stemphylium lycopersici* (GCA 003268315.1) was used as outgroup. Numbers at nodes indicate SH-aLRT support (%) / ultrafast bootstrap (UFBoot) support (%), based on 1,000 replicates each.

The maximum-likelihood phylogenomic tree reconstructed from the concatenated alignment of single-copy orthologous genes across the comparative dataset was consistent with the established sectional classification within the genus *Alternaria*, with all major sections recovered as well as the monotypic lineage *A. brassicae*. *A. telliensis* and *A. japonica* were placed within section *Japonicae* with maximum bootstrap support (100%) (Figure 1). This placement supports the assignment of *A. telliensis* to section *Japonicae*.

This study constitutes the first genome sequence report for two *Alternaria* species of section *Japonicae* that are both considered dangerous for Brassicaceae plants. Although *A. japonica* has been reported worldwide, records of *A. telliensis* are recent and, up to now, restricted to Algeria and Iran. Future comparative studies with available genome sequences from other important pathogens of Brassicaceae such as *A. brassicae* and *A. brassicicola* may contribute to identify key determinants of pathogenicity for these *Alternaria* species.

Supplementary Material

Supplementary figure and table.
<https://www.jgenomics.com/v14p0030s1.pdf>

Acknowledgments

This research is part of R Barkat’s doctoral studies supported by the Program Hubert Curien “Tassili”

between France and Algeria, BRASALT project N° 24MDU104 and by the University of Angers as part of a financial package for co-supervised PhDs. The authors are grateful to Muriel Bahut (UMR IRHS, ANAN platform) for her support with genome sequencing.

Data availability

This Whole Genome Shotgun project has been deposited in DDBJ/ENA/GenBank under the accession no. JCBLRW000000000 and JCBLRX000000000 for *A. japonica* CBS 118390 and *A. telliensis* CBS 145643, respectively. The versions described in this paper are the first versions.

Competing Interests

The authors have declared that no competing interest exists.

References

- Bessadat N, Hamon B, Bataille-Simoneau N, Mabrouk K, Simoneau P. *Alternaria telliensis* sp. nov., a new species isolated from Solanaceae in Algeria. *Phytotaxa*. 2020; 440: 89-100. doi:10.11646/phytotaxa.440.2.1
- Poursafar A, Ghosta Y, Azizi R. *Alternaria telliensis*, a new causal agent of cabbage leaf spot disease in Iran. *Mycologia Iranica*. 2021; 7(2): 231-9. doi:10.22092/MI.2021.123906
- Woudenberg JHC, Groenewald JZ, Binder M, Crous PW. *Alternaria* redefined. *Stud Mycol*. 2013; 75 (1), 171-212. doi:10.3114/sim0015.
- Möller EM, Bahnweg G, Sandermann H, Geiger HH. A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. *Nucl Acids Res*. 1992; 20(22): 6115-6. doi:10.1093/nar/20.22.6115
- Sereika M, Kirkegaard RH, Karst SM, Michaelsen TY, Sørensen EA, Wollenberg RD, Albertsen M. Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and

- metagenomes without short-read or reference polishing. *Nat Methods*. 2022; 19(7): 823-6. doi:10.1038/s41592-022-01539-7
6. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*. 2018; 34(15): 2666-9. doi:10.1093/bioinformatics/bty149
 7. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol*. 2019; 37(5): 540-6. doi:10.1038/s41587-019-0072-8
 8. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res*. 2017; 27(5): 737-46. doi:10.1101/gr.214270.116
 9. Li H. Minimap2: pairwise alignment for nucleotide sequences. Birol I, editor. *Bioinformatics*. 2018; 34(18): 3094-100. doi:10.1093/bioinformatics/bty191
 10. Mikheenko A, Saveliev V, Hirsch P, Gurevich A. WebQUAST: online evaluation of genome assemblies. *Nucl Acids Res*. 2023; 51(W1):W601-6. doi:10.1093/nar/gkad406
 11. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015; 31(19): 3210-2. doi:10.1093/bioinformatics/btv351
 12. Shen W, Sipos B, Zhao L. SeqKit2: A Swiss army knife for sequence and alignment processing. *iMeta*. 2024; 3(3): e191. doi:10.1002/imt2.191
 13. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019; 20(1): 238. doi:10.1186/s13059-019-1832-y
 14. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015; 12(1): 59-60. doi:10.1038/nmeth.3176
 15. Enright AJ. An efficient algorithm for large-scale detection of protein families. *Nucl Acids Res*. 2002; 30(7): 1575-84. doi:10.1093/nar/30.7.1575
 16. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol*. 2013; 30(4): 772-80. doi:10.1093/molbev/mst010
 17. Nguyen L, Schmidt HA, Von Haeseler A, Minh B. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol*. 2015; 32(1): 268-74. doi:10.1093/molbev/msu300
 18. Kalyaanamoorthy S, Minh B, Wong TF, Von Haeseler A, Jermiin L. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017; 14(6): 587-9. doi:10.1038/nmeth.4285
 19. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol*. 2010; 59(3): 307-21. doi:10.1093/sysbio/syq010
 20. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol*. 2018; 35(2): 518-22. doi:10.1093/molbev/msx281
 21. Paudel R, Muzhinji N, Pandey A, Panthee DR, Dean RA, Louws FJ, et al. Genome assembly and comparative analysis of *Alternaria linariae* reveal novel genes associated with host colonization and virulence. *BMC Genomics*. 2025; 26(1): 638. doi:10.1186/s12864-025-11819-z