# Draft Genome Sequence of *Colletotrichum acutatum Sensu Lato* (*Colletotrichum fioriniae*)

Riccardo Baroncelli,[a,b] Surapareddy Sreenivasaprasad,[a]* Serenella A. Sukno,[b] Michael R. Thon,[b] Eric Holub[a]

University of Warwick, School of Life Sciences, Warwick Crop Centre, Wellesbourne, United Kingdom[a]; Centro Hispano-Luso de Investigaciones Agrarias (CIALE), Departamento de Microbiología y Genética, Universidad de Salamanca, Salamanca, Spain[b]

* Present address: Surapareddy Sreenivasaprasad, University of Bedfordshire, Department of Life Sciences, Institute of Biomedical and Environmental Science and Technology (iBEST), Luton, United Kingdom.

M.R.T. and E.H. contributed equally to this work.

This work is dedicated to Derek Barbara (University of Warwick, Warwick Crop Centre, Wellesbourne, United Kingdom), who sadly passed away in August 2012.

**In addition to its economic impact, *Colletotrichum acutatum sensu lato* is an interesting model for molecular investigations due to the diversity of host-determined specialization and reproductive lifestyles within the species complex. The pathogen *Colletotrichum fioriniae* forms part of this species complex and causes anthracnose in a wide range of crops and wild plants worldwide. Some members of this species have also been reported to be entomopathogenic. Here, we report the draft genome sequence of a heterothallic reference isolate of *C. fioriniae* (strain PJ7). This sequence provides a range of new resources that serve as a useful platform for further research in the field.**

Address correspondence to Riccardo Baroncelli, riccardobaroncelli@gmail.com.

Many species belonging to the genus *Colletotrichum* are causal agents of plant diseases, generally referred as anthracnose, in a wide range of hosts worldwide. Virtually every crop grown in the world is susceptible to one or more species of *Colletotrichum* (1). Many *Colletotrichum* species are characterized by a distinctive hemibiotrophic lifestyle. Members of the *Colletotrichum acutatum* species complex have a wide host range in both domesticated and wild plant species, and their capability to infect insects has also been described (2). Pathogenicity assays have shown that most isolates of the complex are not host specific (3–5). *C. fioriniae* (teleomorph: *Glomerella fioriniae*) strain PJ7 was isolated by Peter R. Johnston from infected strawberry (*Fragaria* x *ananassa*) fruit in the Auckland area, New Zealand, in 1988 (6, 7). The strain has been used as a reference strain for phylogenetic analyses of the *C. acutatum* species complex and for mating tests and pathogenicity assays (8, 9). The heterothallic mating capability of this strain has been demonstrated in laboratory experiments (8).

The genome sequence of *C. fioriniae* (*G. fioriniae*) strain PJ7 was obtained using Illumina mate-paired sequencing technology. Mate-paired reads of 50 bp and 70 bp (2.44 Gbp; average coverage, 49.7×) were assembled using Velvet (10). The contigs corresponding to the mitochondrial genome (mtDNA) and the rRNA-coding gene cluster were identified by BLASTn searches using Geneious R6. The mitochondrial genome was assembled into one scaffold using Geneious R6, with a total length of 29.868 Mbp and a G+C content of 30.10%. The mitochondrial DNA was inspected by tBLASTn searches to identify known conserved coding genes using *Colletotrichum graminicola* mtDNA orthologs as the query sequences, resulting in the identification of 16 protein-coding genes and 29 tRNA-coding genes.

The draft nuclear genome of *C. fioriniae* consists of 1,108 sequence scaffolds with a total length 49.01 Mbp ($N_{50}$, 137,254; $N_{90}$, 38,253), 52.50% G+C content, and a maximum scaffold size of 596,408 bp. The completeness of the assembly was assessed using CEGMA version 2.4 (11), which estimated the genome sequence to be 98.39% complete. The nuclear genome was annotated using the MAKER pipeline (12), and tRNAscan was used to predict tRNAs (13). Overall, 13,759 protein-coding gene models and 317 tRNA-coding gene models were predicted in the nuclear genome. Of the protein-coding gene models, 11,039 (80.2%) are supported by protein and/or mRNA sequence evidence.

Analysis with WoLF PSORT (14) revealed that 2,203 predicted proteins (16.01% of the proteome) are secreted. Among those, 90 (4.09% of the secretome and 0.65% of the proteome) do not have any sequence similarity to proteins in public databases, based on BLAST searches. Such characteristics are typical of fungal effectors, which are proteins that have important roles in disabling the host defense system (15).

In this study, we generated the draft genome sequence from a member of the *C. acutatum* species complex. A number of distinct genetic groups within *C. acutatum sensu lato* were previously described (16), leading to recent disaggregation of the complex into 31 species (7). The sequence represents a new resource that will be useful for further research into the biology, ecology, and evolution of these key pathogens.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited in GenBank under the accession no. JARH00000000 (BioProject PRJNA233987). The version described in this paper is JARH00000000.1.

## ACKNOWLEDGMENTS

## REFERENCES

1. **Dean R, Al Van Kan J, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, Foster GD.** 2012. Top 10 fungal pathogens in molecular plant pathology. Mol. Plant Pathol. **13:** 414–430. http://dx.doi.org/10.1111/j.1364-3703.2011.00783.x.

2. **Marcelino J, Giordano R, Gouli S, Gouli V, Parker BL, Skinner M, TeBeest D, Cesnik R.** 2008. *Colletotrichum acutatum* var. *fioriniae* (teleomorph: *Glomerella acutata* var. *fioriniae* var. nov.) infection of a scale insect. Mycologia **100:**353–374. http://dx.doi.org/10.3852/07-174R.

3. **Alahakoon PW, Brown AE, Sreenivasaprasad S.** 1994. Cross infection potential of genetic groups of *Colletotrichum gloeosporioides* on tropical fruits. Physiol. Mol. Plant Pathol. **44:**93–103. http://dx.doi.org/10.1016/S0885-5765(05)80104-3.

4. **Freeman S, Katan T, Shabi E.** 1998. Characterization of *Colletotrichum* species responsible for anthracnose diseases of various fruits. Plant Dis. **82:**596–605. http://dx.doi.org/10.1094/PDIS.1998.82.6.596.

5. **Freeman S, Shabi E.** 1996. Cross-infection of subtropical and temperate fruits by *Colletotrichum* species from various hosts. Physiol. Mol. Plant Pathol. **49:**395–404. http://dx.doi.org/10.1006/pmpp.1996.0062.

6. **Lardner R, Johnston PR, Plummer KM, Pearson MN.** 1999. Morphological and molecular analysis of *Colletotrichum acutatum sensu lato*. Mycol. Res. **103:**275–285. http://dx.doi.org/10.1017/S0953756298007023.

7. **Damm U, Cannon PF, Woudenberg JH, Crous PW.** 2012. The *Colletotrichum acutatum* species complex. Stud. Mycol. **73:**37–113. http://dx.doi.org/10.3114/sim0010.

8. **Guerber JC, Liu B, Correll JC, Johnston PR.** 2003. Characterization of diversity in *Colletotrichum acutatum sensu lato* by sequence analysis of two gene introns, mtDNA and intron RFLPs, and mating compatibility. Mycologia **95:**872–895. http://dx.doi.org/10.2307/3762016.

9. **Baroncelli R.** 2012. *Colletotrichum acutatum sensu lato*: from diversity study to genome analysis. Ph.D. thesis. University of Warwick, Coventry, United Kingdom.

10. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. **18:**821–829. http://dx.doi.org/10.1101/gr.074492.107.

11. **Parra G, Bradnam K, Korf I.** 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics **23:**1061–1067. http://dx.doi.org/10.1093/bioinformatics/btm071.

12. **Cantarel BL, Korf I, Robb SM, 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:**188–196. http://dx.doi.org/10.1101/gr.6743907.

13. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:** 955–964. http://dx.doi.org/10.1093/nar/25.5.0955.

14. **Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K.** 2007. WoLF PSORT: protein localization predictor. Nucleic Acids Res. **35:**W585–W587. http://dx.doi.org/10.1093/nar/gkm259.

15. **O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, Torres MF, Damm U, Buiate EA, Epstein L, Alkan N, Altmüller J, Alvarado-Balderrama L, Bauser CA, Becker C, Birren BW, Chen Z, Choi J, Crouch JA, Duvick JP, Farman MA, Gan P, Heiman D, Henrissat B, Howard RJ, Kabbage M, Koch C, Kracher B, Kubo Y, Law AD, Lebrun MH, Lee YH, Miyara I, Moore N, Neumann U, Nordström K, Panaccione DG, Panstruga R, Place M, Proctor RH, Prusky D, Rech G, Reinhardt R, Rollins JA, Rounsley S, Schardl CL, Schwartz DC, Shenoy N, Shirasu K, Sikhakolli UR, Stüber K, et al.** 2012. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. Nat. Genet. **44:**1060–1065. http://dx.doi.org/10.1038/ng.2372.

16. **Sreenivasaprasad S, Talhinhas P.** 2005. Genotypic and phenotypic diversity in *Colletotrichum acutatum*, a cosmopolitan pathogen causing anthracnose on a wide range of hosts. Mol. Plant Pathol. **6:**361–378. http://dx.doi.org/10.1111/j.1364-3703.2005.00291.x.