Journal of Plant Pathology
https://doi.org/10.1007/s42161-023-01439-5

GENOME RESOURCE PAPER (GRP)

Draft genome sequence of the keylime (*Citrus* × *aurantiifolia*) pathogen *Colletotrichum limetticola*

Andrea Menicucci[1] · Isis Tikami[1,2] · Tiziano Benocci[3] · Antonio Zapparata[4] · Nelson Sidnei Massola Júnior[2] · Natalia Aparecida Peres[5] · Lavern Wayne Timmer[5] · Antonio Prodi[1] · Riccardo Baroncelli[1,6]

Received: 26 May 2023 / Accepted: 22 June 2023

Keywords Acutatum · Fungus · Comparative genomics · Plant pathogen · *Glomerella*

Many species belonging to the genus *Colletotrichum* are causal agents of plant diseases, generally referred to as anthracnose, in a wide range of hosts worldwide. *Colletotrichum* spp. are responsible for impacting numerous economically important crops on a global scale. This genus comprises approximately 257 distinct species, which are further organized into at least 15 major phylogenetic lineages known as species complexes (Talhinhas and Baroncelli 2021). Virtually every crop grown in the world is susceptible to one or more species of *Colletotrichum* (Baroncelli et al. 2014). Among these, the *Colletotrichum acutatum* species complex stands out as a diverse group of closely related plant pathogenic fungi within the genus (Baroncelli et al. 2017). 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 (Damn et al. 2012, Marcelino et al. 2008). In this species complex, *Colletotrichum limetticola* (formerly known as *Gloeosporium limetticola*; Clausen 1912) was initially described in 2012 as a species predominantly associated with wither tip symptoms on sour lime (*Citrus aurantiifolia*) in Cuba and the USA during the 1910s (Damm et al. 2012). Later descriptions associated the disease with strains of *C. gloeosporioides* (Brown et al. 1996) or *C. acutatum* (Peres et al. 2008). Recent findings in Brazil have revealed the presence of *C. limetticola* causing Glomerella leaf spot on apples, although its prevalence remains low while displaying high virulence (Moreira et al. 2019). To the best of our knowledge, no further occurrences of *C. limetticola* have been documented, despite the presence of other known *Colletotrichum* species that infect citrus and apples (Talhinhas and Baroncelli 2021). This raises concerns regarding the conservation status of *C. limetticola* considering the scarcity of records on its original hosts and the occurrence of cross-infections.

In the present study, *Colletotrichum limetticola* strain KLA-Anderson was isolated from a leaf tissue of *Citrus* x *aurantiifolia* commonly known as the Key lime or Mexican lime in the Lake Alfred region (Florida, USA). *C. limetticola* genome was sequenced using the Illumina NovaSeq 6000 150 bp paired-end sequencing system. Illumina sequences were analyzed with FastQC (Babraham Bioinformatics) to assess the quality of the reads. Sequences adapters and low-quality reads were trimmed with TrimGalore! v0.6.10 (Krueger et al. 2021). Pair-end reads were merged with FLASH v1.2.11 (Magoc and Salzberg 2011). Merged and unmerged reads were then assembled using SPAdes v3.15.1 (Bankevich et al. 2012). Scaffolds with low coverage were removed as possible contaminations. Scaffolds corresponding to the mitochondrial DNA (mtDNA) and ribosomal DNA (rDNA) genome were identified by BLASTN v2.9.0 (Camacho et al. 2009) using queries of the closely related species *Colletotrichum lupini* (Baroncelli et al. 2021) which was

Andrea Menicucci and Isis Tikami contributed equally to this work.

✉ Riccardo Baroncelli
riccardo.baroncelli@unibo.it

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

[2] Escola Superior de Agricultura Luiz de Queiroz (ESALQ), University of São Paulo (USP), Piracicaba, São Paulo, Brazil

[3] Center for Health and Bioresources, AIT Austrian Institute of Technology, Tulln, Austria

[4] "Federigo Enriques" High School, Livorno, Italy

[5] Plant Pathology Department, Gulf Coast Research and Education Center, University of Florida, Wimauma, USA

[6] Center for Studies on Bioinspired Agro-Enviromental Technology, University of Naples Federico II, Portici, Italy

Published online: 07 July 2023

Table 1 Summary statistics of the *Colletotrichum limetticola* KLA-Anderson strain genome

Genome variables	Statistics
Assembly statistics	
Number of scaffolds	1750
Assembly length (Mb)	50.48
Maximum scaffold size (Mb)	0.38
N50	68638
N90	17501
L50	229
L90	752
Guanine-Cytosine content (%)	52.77
BUSCO completeness	
Complete (%)	97.7
Single copy (%)	97.6
Duplicate (%)	0.1
Fragmented (%)	1.1
Missing (%)	1.2
Protein encoding genes	
Number of predicted proteins	15248
Number of secreted proteins	1981
Number of predicted effector proteins	624
Number of Predicted cytoplasmic effectors	172
Number of Predicted apoplastic effectors	452

the closest complete genome to *C. limetticola*. The completeness of the assembly was assessed using BUSCO v5.4.7 (Simão et al. 2015) while statistics were evaluated with QUAST v5.2.0 (Gurevich et al. 2013). The total size of the nuclear genome assembly was 50,48 Mb, with an N50 contig length of 68638 kb and a L50 of 229. The nuclear genome assembly resulted in 1750 contigs with an average coverage of 90X and it was assessed to be 97.7% complete (Table 1). A total of 15248 protein-coding genes were predicted to be encoded using MAKER v3.01.02 pipeline (Holt and Yandell 2011) with both self-trained GeneMark-ES v4.10 (Borodovsky and Lomsadze 2011) and AUGUSTUS v3.3 prediction using the "*Colletotrichum*" model (Becerra et al. 2023). SignalP v5.0 (Almagro Armenteros et al. 2019) revealed that 1981 proteins in *C. limetticola* are secreted and among those 624 have been predicted by EffectorP v3.0 (Sperschneider and Dodds 2022) to be candidate effectors. A comparative analysis of the newly sequenced genome with those publicly available (Baroncelli et al. 2016, 2021, 2022; Goulin et al. 2023) revealed similar genomic features and gene content within closely related species (Fig. 1).

In this study we presented a draft genome sequence of *C. limetticola*, obtained using Illumina sequencing technology, providing a range of new resources that serve as a useful platform for further research in the field of comparative genomics of fungi. Further analysis of these genomes will enhance our understanding of the molecular mechanisms underlying the pathogenicity and virulence of *Colletotrichum* species facilitating the exploration of potential targeted and environmentally friendly strategies for its control.

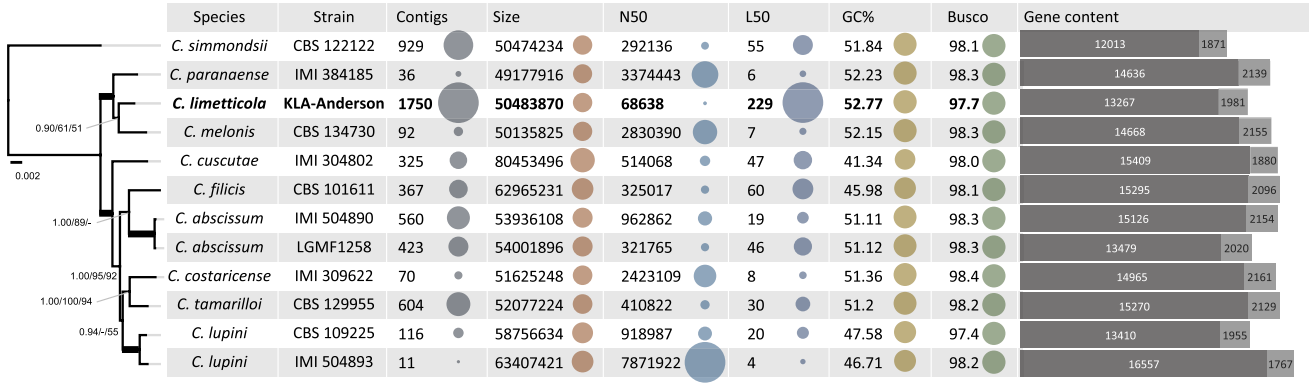

Fig. 1 Comparative analysis between the newly sequenced genome of *C. limetticola* and a selection of closely related species publicly available. The *C. limetticola* genome is highlighted in bold. On the left side, multilocus sequence typing (MLST) tree based on the concatenation of the partial sequences of following loci: actin [ACT], beta-tubulin 2 [TUB2], calmodulin [CAL], glyceraldehyde-3-phosphate dehydrogenase [GAPDH], chitin synthase [CHS-1], glutamine synthetase [GS], histone-3 [HIS3], superoxide dismutase 2 [SOD2] mating type 1–2 [MAT1-2] and the Apn2-Mat1-2 intergenic spacer [ApnMat]. Numbers next to the nodes represent in order: Bayesian posterior probability, FastTree and RAxML bootstrap support values. Bubble plots report on assembly fragmentation, genome size, N50 and L50, GC content and completeness. Bubble sizes have been scaled to each panel and are not comparable across panels. Horizontal histograms report on secreted and non-secreted protein coding gene content

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement. This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them and the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil, Grant/ Award Number: 88887.695424/2022–00.

Data availability The data generated in this study are publicly available from the NCBI GenBank database at Bioproject ID PRJNA952538 and Biosample ID SAMN34075281. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JARUPT000000000. The version described in this paper is version JARUPT010000000.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

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 37:420–423. https://doi.org/10.1038/s41587-019-0036-z

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021

Baroncelli R, Amby DB, Zapparata A, Sarrocco S, Vannacci G, Le Floch G, Harrison RJ, Holub E, Sukno SA, Sreenivasaprasad S, Thon MR (2016) Gene family expansions and contractions are associated with host range in plant pathogens of the genus Colletotrichum. BMC Genomics 17:555. https://doi.org/10.1186/s12864-016-2917-6

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, Sukno SA, 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

Baroncelli R, Pensec F, Da Lio D, Boufleur T, Vicente I, Sarrocco S, Picot A, Baraldi E, Sukno S, Thon M, Le Floch G (2021) Complete Genome Sequence of the Plant-Pathogenic Fungus Colletotrichum lupini. MPMI 34:1461–1464. https://doi.org/10.1094/MPMI-07-21-0173-A

Baroncelli R, Sreenivasaprasad S, Sukno SA, Thon MR, Holub E (2014) Draft Genome Sequence of *Colletotrichum acutatum* Sensu Lato (*Colletotrichum fioriniae*). Genome Announc 2:e00112-e114. https://doi.org/10.1128/genomeA.00112-14

Baroncelli R, Talhinhas P, Pensec F, Sukno SA, Le Floch G, Thon MR (2017) The *Colletotrichum acutatum* Species Complex as a Model System to Study Evolution and Host Specialization in Plant Pathogens. Front Microbiol 8:2001. https://doi.org/10.3389/fmicb.2017.02001

Becerra S, Baroncelli R, Boufleur TR, Sukno SA, Thon MR (2023) Chromosome-level analysis of the *Colletotrichum graminicola* genome reveals the unique characteristics of core and minichromosomes. Front Microbiol 14:760. https://doi.org/10.3389/fmicb.2023.1129319

Borodovsky M, Lomsadze A (2011) Eukaryotic Gene Prediction Using GeneMark.hmm-E and GeneMark-ES. Curr Protoc Bioinformatics 35:1. https://doi.org/10.1002/0471250953.bi0406s35

Brown AE, Sreenivasaprasad S, Timmer LW (1996) Molecular characterization of slow-growing orange and key lime anthracnose strains of Colletotrichum from citrus as *C. acutatum*. Phytopathology 86:523–527. https://doi.org/10.1094/Phyto-86-523

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

Clausen RE (1912) A new fungus concerned in wither tip of varieties of *Citrus medica*. Phytopathology 2:217–234

Damm U, Cannon PF, Woudenberg JHC, Crous PW (2012) The *Colletotrichum acutatum* species complex. Stud Mycol 73:37–113. https://doi.org/10.3114/sim0010

Goulin E, Boufleur TR, Negrini F, Carneiro GA, Baraldi E, Machado MA, Floch GL. Baroncelli R (2023) Genome Sequence Resources of *Colletotrichum abscissum*, the Causal Agent of Citrus Post-Bloom Fruit Drop, and the Closely Related Species *C. fiticis*. Phytopathology® 113:104–107. https://doi.org/10.1094/PHYTO-05-22-0176-A

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

Holt C, Yandell M (2011) MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12:491. https://doi.org/10.1186/1471-2105-12-491

Krueger F, James F, Ewels P, Afyounian E, Schuster-Boeckler B (2021) FelixKrueger/TrimGalore. Zenodo. https://doi.org/10.5281/zenodo.5127898

Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963. https://doi.org/10.1093/bioinformatics/btr507

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. https://doi.org/10.3852/07-174R

Moreira RR, Peres NA, May De Mio LL (2019) *Colletotrichum acutatum* and *C. gloeosporioides* Species Complexes Associated with

Apple in Brazil. Plant Dis 103:268–275. https://doi.org/10.1094/PDIS-07-18-1187-RE

Peres NA, MacKenzie SJ, Peever TL, Timmer LW (2008) Postbloom fruit drop of citrus and Key lime anthracnose are caused by distinct phylogenetic lineages of *Colletotrichum acutatum*. Phytopathology 98:345–352. https://doi.org/10.1094/PHYTO-98-3-0345

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Sperschneider J, Dodds PN (2022) EffectorP 3.0: Prediction of Apoplastic and Cytoplasmic Effectors in Fungi and Oomycetes. MPMI 35:146–156. https://doi.org/10.1094/MPMI-08-21-0201-R

Talhinhas P, Baroncelli R (2021) Colletotrichum species and complexes: geographic distribution, host range and conservation status. Fungal Divers 110:109–198. https://doi.org/10.1007/s13225-021-00491-9