AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Genome Sequence of *Metarhizium rileyi*, a Microbial Control Agent for Lepidoptera

Eliseu Binneck,[a] Claudia Cristina López Lastra,[b] Daniel R. Sosa-Gómez[a]

[a]Embrapa Soja, Londrina, Paraná, Brazil
[b]Centro de Estudios Parasitológicos y de Vectores, La Plata, Buenos Aires, Argentina

ABSTRACT *Metarhizium rileyi* (formerly known as *Nomuraea rileyi*) is a potential agent for microbial control of many insect pests from the order Lepidoptera, the damages of which can cause considerable loss of productivity in agriculture. We report the genome sequence and annotation of *M. rileyi* strain Cep018-CH2/ARSEF 7053.

Metarhizium rileyi is a cosmopolitan species of entomopathogenic fungi of the family Clavicipitaceae (Hypocreales, Ascomycota) with extensive literature published under its synonym *Nomuraea rileyi* (1, 2). The main susceptible species of insects, which are key pests of crops such as cotton and soybean, belong to the lepidopteran families Noctuidae, Erebidae, and Nymphalidae (3–7). *Metarhizium rileyi* usually presents high genetic variability (8), which has been closely related to the host species from which it is isolated (4, 9, 10). Unlike other most common fungal entomopathogens with the greatest known epizootic potentials, such as *Metarhizium anisopliae* and *Beauveria bassiana*, *M. rileyi* has a narrow spectrum of hosts (10, 11). Because of its high selectivity and effective control under natural or agricultural conditions, *M. rileyi* is an attractive biocontrol agent with potential for development as a bioinsecticide (3, 5) or for prospecting potential biologically active compounds with many possible uses. The genome data of *M. rileyi* strain Cep018-CH2/ARSEF 7053 were obtained, with the aim of providing additional insights into fungal diversity and interactions with the host.

M. rileyi strain Cep018-CH2 was isolated from a velvetbean caterpillar (*Anticarsia gemmatalis*) on 4 April 2001, in Chivilcoy (Buenos Aires Province, Argentina), and it was deposited at the Centro de Estudios Parasitológicos y de Vectores (CEPAVE) collection, La Plata, Buenos Aires, Argentina, and also at ARSEF (12) under the accession number ARSEF 7053. A single-spore culture of Cep018-CH2/ARSEF 7053 was prepared on Sabouraud maltose agar with yeast extract (SMAY; 2.5 g of neopeptone, 10 g of maltose, 2.5 g of yeast extract, 3.75 g of agar, and 250 ml of water) at 26°C for 5 to 7 days. Conidia were inoculated in 50 ml of SMAY broth with shaking at 250 rpm at 26°C for 8 to 10 days. The fungal mycelia were collected by filtration and washed with sterile distilled water, and the DNA was extracted using a modified cetyltrimethylammonium bromide (CTAB) protocol (1). The preparation of libraries and sequencing were performed at Fasteris SA (Plan-les-Ouates, Switzerland). Whole-genome sequencing was carried out on an Illumina HiSeq platform using four libraries, two Fasteris *de novo* (Fasteris SA) paired-end libraries with insert sizes of 250 to 350 bp, and two Nextera (Illumina, Inc.) mate pair libraries with insert sizes of 3,000 and 5,000 bp, producing a total of 182,506,935 read pairs. The average read length was 125 bp for both sets of libraries. The quality of the sequencing raw reads was assessed with FastQC version 0.11.5 (13). Default parameters were used in all analyses, unless otherwise stated.

The reads were quality filtered and assembled into scaffolds using the ALLPATHS-LG pipeline version r52488 (14). The generated assembly was evaluated using QUAST-LG

Citation Binneck E, Lastra CCL, Sosa-Gómez DR. 2019. Genome sequence of *Metarhizium rileyi*, a microbial control agent for Lepidoptera. Microbiol Resour Announc 8:e00897-19. https://doi.org/10.1128/MRA.00897-19.

Editor Antonis Rokas, Vanderbilt University

Address correspondence to Eliseu Binneck, eliseu.binneck@embrapa.br.

Received 8 August 2019
Accepted 14 August 2019
Published 5 September 2019

version 5.0.2 (15). The final assembly had a total length of 31,808,756 bp, with 1,044 contigs joined into 249 scaffolds, 240 of which were larger than 1,000 bp. The lengths of the longest and N_{50} scaffolds were 2,535,063 bp and 815,204 bp, respectively, and the L_{50} value was 10. The overall G+C content was 51.30%.

Gene prediction and annotation using Funannotate version 1.5.0-b99af2c (16), with the gene predictors AUGUSTUS version 3.3.1 (17), GeneMark-ES Suite version 4.35 (18), and tRNAscan-SE version 2.0.0 (19), resulted in 8,945 protein-coding and 102 tRNA genes. This annotation is comparable with that of *M. rileyi* RCEF 4871 (GenBank accession number AZHC00000000), which has a total assembly length of 32,013,981 bp and 8,764 protein-coding genes (20). Secondary metabolite analysis was performed using antiSMASH fungal version 4.2.0 (21), identifying 30 gene clusters involved in the biosynthesis of specialized metabolites, 480 biosynthetic enzymes, and 155 secondary metabolism Clusters of Orthologous Groups (smCOGs). Functional annotation of the predicted proteins, by pattern matching with the Pfam (22), UniProtKB/Swiss-Prot (23), eggNOG (24), CAZy (25), MEROPS (26), InterPro (27), and antiSMASH (28) databases, as well as comparison to other entomopathogenic fungal genomes, revealed key genes coding for peptidases, carbohydrate-active enzymes, secreted proteins, and transcription factors.

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number SBHS00000000. The version described in this paper is version SBHS01000000. The raw reads were deposited in the NCBI Sequence Read Archive under the BioProject accession number PRJNA503201.

ACKNOWLEDGMENT

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant 305398/2018-0.

REFERENCES

1. Sosa-Gómez DR, Humber RA, Hodge KT, Binneck E, da Silva-Brandão KL. 2009. Variability of the mitochondrial SSU rDNA of *Nomuraea* species and other entomopathogenic fungi from hypocreales. Mycopathologia 167:145–154. https://doi.org/10.1007/s11046-008-9157-5.
2. Kepler RM, Humber RA, Bischoff JF, Rehner SA. 2014. Clarification of generic and species boundaries for *Metarhizium* and related fungi through multigene phylogenetics. Mycologia 106:811–829. https://doi.org/10.3852/13-319.
3. Ignoffo CM. 1981. The fungus *Nomuraea rileyi* as a microbial insecticide, p 513–538. *In* Burges HD (ed), Microbial control of pests and plant diseases 1970–1980. Academic Press, London, United Kingdom.
4. Boucias DG, Schoborg EA, Allen GE. 1982. The relative susceptibility of six noctuid species to infection by *Nomuraea rileyi* isolated from *Anticarsa gemmatalis*. J Invertebr Pathol 39:238–240. https://doi.org/10.1016/0022-2011(82)90017-9.
5. Devi PSV, Prasad YG. 2001. Nomuraea rileyi—a potential mycoinsecticide, p 23–38. *In* Upadhyay RK, Mukerji KG, Chamola BP (ed), Biocontrol potential and its exploitation in sustainable agriculture. Springer, Boston, MA.
6. Fuxa JR. 1984. Dispersion and spread of the entomopathogenic fungus *Nomuraea rileyi* (Moniliales: Moniliaceae) in a soybean field. Environ Entomol 13:252–258. https://doi.org/10.1093/ee/13.1.252.
7. Sosa-Gómez DR. 2017. Chapter 13–microbial control of soybean pest insects and mites, p 199–208. *In* Lacey LA (ed), Microbial control of insect and mite pests. Academic Press, London, United Kingdom.
8. Devi UK, Reineke A, Rao UCM, Reddy NRN, Khan A. 2007. AFLP and single-strand conformation polymorphism studies of recombination in the entomopathogenic fungus *Nomuraea rileyi*. Mycol Res 111:716–725. https://doi.org/10.1016/j.mycres.2007.03.003.
9. Moscardi F, Kastelic JG, Sosa-Gómez DR. 1992. Suscetibilidade de três espécies de lepidópteros associados à soja a três isolados do fungo *Nomuraea rileyi* (Farlow) Samson. An Soc Entomol Bras 21:93–100.
10. Suwannakut S, Boucias DG, Wiwat C. 2005. Genotypic analysis of *Nomuraea rileyi* collected from various noctuid hosts. J Invertebr Pathol 90:169–176. https://doi.org/10.1016/j.jip.2005.08.010.
11. Sosa-Gómez DR, López Lastra CC, Humber RA. 2010. An overview of arthropod-associated fungi from Argentina and Brazil. Mycopathologia 170:61–76. https://doi.org/10.1007/s11046-010-9288-3.
12. U.S. Department of Agriculture, Agricultural Research Service. 2016. ARS collection of entomopathogenic fungal cultures. U.S. Department of Agriculture, Agricultural Research Service, Washington, DC.
13. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
14. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci U S A 108:1513–1518. https://doi.org/10.1073/pnas.1017351108.
15. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34:i142–i150. https://doi.org/10.1093/bioinformatics/bty266.
16. Love J, Palmer J, Stajich J, Esser T, Kastman E, Winter D. 2018. Nextgenusfs/Funannotate: Funannotate v1.5.0. https://zenodo.org/record/1342272#.XVtaGSNKi1s.
17. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res 34:W435–W439. https://doi.org/10.1093/nar/gkl200.
18. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res 18:1979–1990. https://doi.org/10.1101/gr.081612.108.
19. Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol 1962:1–14. https://doi.org/10.1007/978-1-4939-9173-0_1.
20. Shang Y, Xiao G, Zheng P, Cen K, Zhan S, Wang C. 2016. Divergent and convergent evolution of fungal pathogenicity. Genome Biol Evol 8:1374–1387. https://doi.org/10.1093/gbe/evw082.
21. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, de Los Santos ELC, Kim HU, Nave M, Dickschat JS, Mitchell DA, Shelest E, Breitling R, Takano E, Lee SY, Weber T, Medema MH. 2017.

antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res 45:W36–W41. https://doi.org/10.1093/nar/gkx319.
22. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD. 2019. The Pfam protein families database in 2019. Nucleic Acids Res 47:D427–D432. https://doi.org/10.1093/nar/gky995.
23. UniProt Consortium. 2019. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 47:D506–D515. https://doi.org/10.1093/nar/gky1049.
24. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44:D286–D293. https://doi.org/10.1093/nar/gkv1248.
25. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The Carbohydrate-Active enZymes database (CAZy) in 2013. Nucleic Acids Res 42:D490–D495. https://doi.org/10.1093/nar/gkt1178.
26. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res 46:D624–D632. https://doi.org/10.1093/nar/gkx1134.
27. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang H-Y, El-Gebali S, Fraser MI, Gough J, Haft DR, Huang H, Letunic I, Lopez R, Luciani A, Madeira F, Marchler-Bauer A, Mi H, Natale DA, Necci M, Nuka G, Orengo C, Pandurangan AP, Paysan-Lafosse T, Pesseat S, Potter SC, Qureshi MA, Rawlings ND, Redaschi N, Richardson LJ, Rivoire C, Salazar GA, Sangrador-Vegas A, Sigrist CJA, Sillitoe I, Sutton GG, Thanki N, Thomas PD, Tosatto SCE, Yong S-Y, Finn RD. 2019. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res 47:D351–D360. https://doi.org/10.1093/nar/gky1100.
28. Blin K, Pascal Andreu V, de los Santos ELC, Del Carratore F, Lee SY, Medema MH, Weber T. 2019. The antiSMASH database version 2: a comprehensive resource on secondary metabolite biosynthetic gene clusters. Nucleic Acids Res 47:D625–D630. https://doi.org/10.1093/nar/gky1060.