AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequences of Four *Aspergillus* Section *Fumigati* Clinical Strains

Renato Augusto Corrêa dos Santos,[a,b] Olga Rivero-Menendez,[c] Jacob L. Steenwyk,[b] Matthew E. Mead,[b] Gustavo Henrique Goldman,[a] Ana Alastruey-Izquierdo,[c] Antonis Rokas[b]

[a]Departmento de Ciências Farmacêuticas, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil
[b]Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, USA
[c]Medical Mycology Reference Laboratory, National Center for Microbiology, Instituto de Salud Carlos III, Madrid, Spain

ABSTRACT *Aspergillus* fungi in section *Fumigati* include important human pathogens. Here, we sequenced the genomes of two strains of *Aspergillus hiratsukae* and two strains of *Aspergillus felis*. The average genome sizes are 29.5 Mb for *A. hiratsukae* and 31.8 Mb for *A. felis*.

Aspergillus is a highly diverse genus of industrially and medically important fungi (1, 2). The genus is taxonomically divided into 27 sections (3). Section *Fumigati* contains the major human pathogen *Aspergillus fumigatus* (4) and several so-called cryptic species, such as *Aspergillus hiratsukae* and *Aspergillus felis* (5–7), which are morphologically similar but genetically distinct from *A. fumigatus*. Cryptic species account for over 10% of cases of *Aspergillus* infection (8). Here, we sequenced the genomes of two clinical strains of *A. hiratsukae*, CNM-CM5793 and CNM-CM6106, from nail and ear infections, respectively, both from Spain. We also sequenced two clinical strains of *A. felis*, strain CNM-CM7691 from an ear infection in Spain and strain CNM-CM5623 from Portugal. All four isolates were recovered from clinical samples following standard procedures and sent to the Medical Mycology Reference Laboratory (at the National Center for Microbiology, Instituto de Salud Carlos III, Madrid, Spain) for identification and susceptibility testing. Except for infection type, no clinical data were recorded. Therefore, the fungal isolates were judged to be exempt from informed consent of the patients and institutional review board approval.

Species assignment was based on a maximum-likelihood phylogenetic analysis (Fig. 1). For genome sequencing, we grew all strains in glucose-yeast extract-peptone (GYEP) liquid medium (0.3% yeast extract and 1% peptone; Difco, Soria Melguizo) with 2% glucose (Sigma-Aldrich, Spain) for 24 to 48 h at 30°C. The mycelium was mechanically disrupted by vortex mixing with glass beads and used to extract genomic DNA using the phenol-chloroform method (9). DNA was quantified using the QuantiFluor double-stranded DNA (dsDNA) system and the QuantiFluor ST fluorometer (Promega, Madison, WI, USA). DNA quality was checked with the Agilent 2100 bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). DNA libraries were prepared using the Nextera DNA library prep kit (Illumina, Inc., San Diego, CA, USA) according to the manufacturer's guidelines. Paired-end sequencing (2 × 150 bp) was performed using the NextSeq 500 platform following the manufacturer's protocols (Illumina, Inc.).

For all software, default parameters were used except where otherwise noted. The numbers of sequencing read pairs generated for strains CNM-CM5793, CNM-CM6106, CNM-CM7691, and CNM-CM5623 were 7,733,508, 5,237,901, 9,555,248, and 6,768,577, respectively. Quality control of the sequence reads was performed with FastQC v0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Raw reads were cleaned with Trimmomatic v0.38 (10) with the following parameters: NexteraPE-PE.fa:2:30:10:2:

Citation dos Santos RAC, Rivero-Menendez O, Steenwyk JL, Mead ME, Goldman GH, Alastruey-Izquierdo A, Rokas A. 2020. Draft genome sequences of four *Aspergillus* section *Fumigati* clinical strains. Microbiol Resour Announc 9:e00856-20. https://doi.org/10.1128/MRA.00856-20.

Editor Vincent Bruno, University of Maryland School of Medicine

Address correspondence to Gustavo Henrique Goldman, ggoldman@usp.br, or Antonis Rokas, antonis.rokas@vanderbilt.edu.

Received 25 August 2020
Accepted 9 September 2020
Published 1 October 2020

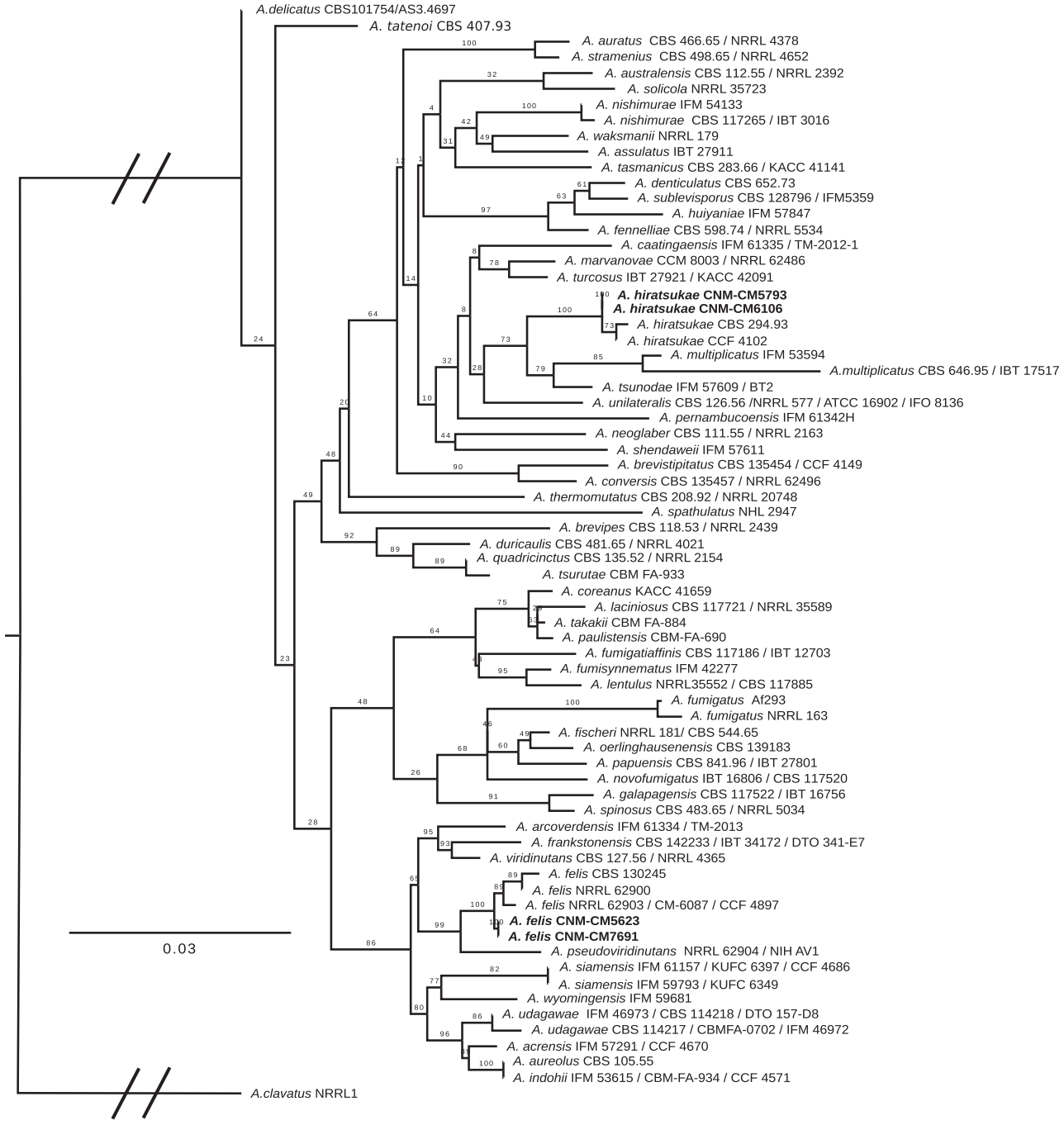

FIG 1 Maximum-likelihood phylogenetic tree of the four strains sequenced in this announcement (in bold) and related species in section *Fumigati*, based on the analysis of the four markers beta-tubulin gene (*benA*), calmodulin gene (*CaM*), actin gene (*act*), and RNA polymerase II second-largest subunit gene (*RPB2*), commonly used in *Aspergillus* taxonomy (18); sequences were obtained from reference 18 except for the sequences of the four newly sequenced strains, which were obtained by searching for markers of each strain in orthogroups generated by OrthoFinder v2.3.3 (19) using *A. fumigatus* Af293 (17) as the reference. Each marker was aligned with MAFFT v7.397 (20), and a supermatrix was generated with FASconCAT v1.11 (21). Tree inference was carried out on IQ-TREE v2.0.3 (22) with partitions with the option "MFP+MERGE," which employs ModelFinder to find the best partition scheme. The final tree was edited in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). Support values are based on 1,000 bootstrap replicates. *A. clavatus* (section *Clavati*) was used to root the tree. Note that the species *A. parafelis*, *A. pseudofelis*, and *A. felis* were merged (synonymized) into a single species, *A. felis* (18); thus, we infer that the two sequenced strains belong to *A. felis*.

keepBothReads, SLIDINGWINDOW:4:15, LEADING:3, TRAILING:3, and MINLEN:90. The genome sequences were assembled with SPAdes v3.14.0 (11) employing multiple k-mers (31, 41, 51, 61, 71, 81, and 91) and the --careful parameter. The genomic reads were mapped to the assembly with Bowtie v2.3.4.1 (12), followed by a single iteration

TABLE 1 Overall genome assembly, completeness, and annotation statistics

Strain	Assembly size[a] (bp)	No. of contigs >1,000 bp	Avg genome coverage (×)	GC content (%)	N_{50} (bp)	No. of genes	No. (%) of complete and single-copy BUSCOs	No. (%) of fragmented BUSCOs
A. hiratsukae CNM-CM5793	29,562,918	745	63	50.38	100,935	9,685	3,487 (98.33)	31 (0.87)
A. hiratsukae CNM-CM6106	29,374,270	922	39	50.37	71,695	9,663	3,466 (97.74)	38 (1.07)
A. felis CNM-CM5623	31,643,783	663	47	49.93	112,776	10,161	3,494 (98.53)	26 (0.73)
A. felis CNM-CM7691	31,957,614	559	70	49.93	138,232	10,243	3,503 (98.78)	18 (0.51)

[a] Based on contigs with more than 1,000 bp.

of Pilon v1.22 (13) used in base correction to fix misassemblies and for gap filling. Overall statistics of the final assemblies were assessed with QUAST v4.6.3 (14). Genome assembly completeness was assessed by examining for the presence of 3,546 universal single-copy orthologs in Eurotiomycetes (eurotiomycetes_odb10) with BUSCO v4.0.4 (15). We used AUGUSTUS v3.3.1 (16) for prediction of protein-coding genes using the *A. fumigatus* Af293 gene models as a reference (17).

Assembly sizes, numbers of contigs, GC content, N_{50} contig values, gene numbers, percentages of complete and single-copy BUSCOs, and percentages of fragmented BUSCOs for all four strains are reported in Table 1. We summarized the genome statistics based on contigs of greater than 1,000 bp but included all contigs of greater than 200 bp in the assemblies submitted to GenBank. Genomic information of *Aspergillus* strains that are closely related to major human pathogens is important for understanding the origin and evolution of opportunistic human pathogenicity in the genus *Aspergillus*.

Data availability. The draft genome sequences of *A. felis* strains CNM-CM5623 and CNM-CM7691 and *A. hiratsukae* strains CNM-CM5793 and CNM-CM6106 are deposited in GenBank under the accession numbers JACBAE000000000, JACBAG000000000, JACBAD000000000, and JACBAF000000000, respectively. The raw reads of *A. felis* strains CNM-CM5623 and CNM-CM7691 and *A. hiratsukae* strains CNM-CM5793 and CNM-CM6106 are deposited in the NCBI Sequence Read Archive (SRA) under accession numbers SRR11804853, SRR11804830, SRR11802685, and SRR11802449, respectively. Genome assemblies and raw data are associated with BioProject number PRJNA633131.

ACKNOWLEDGMENTS

This research was funded by Brazilian São Paulo Research Foundation (FAPESP) grant 2016/07870-9 (to G.H.G.) and scholarships 2017/21983-3 and 2019/07526-4 (to R.A.C.D.S.) and by research projects from the Fondo de Investigación Sanitaria (PI13/02145 and PI16CIII/00035) of the Instituto de Salud Carlos III. J.L.S. and A.R. were funded by the Howard Hughes Medical Institute through the James H. Gilliam Fellowships for Advanced Study program. A.R. also received funding from the National Institutes of Health/National Institute of Allergy and Infectious Diseases (1R56AI146096-01A1).

Computational infrastructure was provided by the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University.

REFERENCES

1. Gibbons JG, Rokas A. 2013. The function and evolution of the Aspergillus genome. Trends Microbiol 21:14–22. https://doi.org/10.1016/j.tim.2012.09.005.
2. de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, Uchima CA, Anderluh G, Asadollahi M, Askin M, Barry K, Battaglia E, Bayram Ö, Benocci T, Braus-Stromeyer SA, Caldana C, Cánovas D, Cerqueira GC, Chen F, Chen W, Choi C, Clum A, dos Santos RAC, de Lima Damásio AR, Diallinas G, Emri T, Fekete E, Flipphi M, Freyberg S, Gallo A, Gournas C, Habgood R, Hainaut M, Harispe ML, Henrissat B, Hildén KS, Hope R, Hossain A, Karabika E, Karaffa L, Karányi Z, Kraševec N, Kuo A, Kusch H, LaButti K, Lagendijk EL, Lapidus A, Levasseur A, Lindquist E, Lipzen A, Logrieco AF, MacCabe A, Mäkelä MR, et al. 2017. Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus Aspergillus. Genome Biol 18:28. https://doi.org/10.1186/s13059-017-1151-0.
3. Houbraken J, Kocsubé S, Visagie CM, Yilmaz N, Wang X-C, Meijer M, Kraak B, Hubka V, Samson RA, Frisvad JC. 2020. Classification of Aspergillus, Penicillium, Talaromyces and related genera (Eurotiales): an overview of families, genera, subgenera, sections, series and species. Stud Mycol 95:5–169. https://doi.org/10.1016/j.simyco.2020.05.002.
4. Rokas A, Mead ME, Steenwyk JL, Oberlies NH, Goldman GH. 2020. Evolving moldy murderers: Aspergillus section Fumigati as a model for studying the repeated evolution of fungal pathogenicity. PLoS Pathog 16:e1008315. https://doi.org/10.1371/journal.ppat.1008315.
5. Guarro J, Kallas EG, Godoy P, Karenina A, Gené J, Stchigel A, Colombo AL. 2002. Cerebral aspergillosis caused by Neosartorya hiratsukae, Brazil. Emerg Infect Dis 8:989–991. https://doi.org/10.3201/eid0809.020073.
6. Koutroutsos K, Arabatzis M, Bougatsos G, Xanthaki A, Toutouza M, Velegraki A. 2010. Neosartorya hiratsukae peritonitis through continuous

ambulatory peritoneal dialysis. J Med Microbiol 59:862–865. https://doi.org/10.1099/jmm.0.019133-0.
7. Barrs VR, van Doorn TM, Houbraken J, Kidd SE, Martin P, Pinheiro MD, Richardson M, Varga J, Samson RA. 2013. Aspergillus felis sp. nov., an emerging agent of invasive aspergillosis in humans, cats, and dogs. PLoS One 8:e64871. https://doi.org/10.1371/journal.pone.0064871.
8. Alastruey-Izquierdo A, Alcazar-Fuoli L, Cuenca-Estrella M. 2014. Antifungal susceptibility profile of cryptic species of Aspergillus. Mycopathologia 178:427–433. https://doi.org/10.1007/s11046-014-9775-z.
9. Holden DW. 1994. DNA mini prep method for "*Aspergillus fumigatus* (and other filamentous fungi), p 3–4. *In* Maresca B, Kobayashi GS (ed), Molecular biology of pathogenic fungi, a laboratory manual. Telos Press, New York, NY.
10. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
11. 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.
12. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
13. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
14. 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.
15. 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.
16. Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a Web server for gene finding in eukaryotes. Nucleic Acids Res 32:W309–W312. https://doi.org/10.1093/nar/gkh379.
17. Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulsen R, Davies R, Dyer PS, Farman M, Fedorova N, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, García JL, García MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, Gwilliam R, Haas B, Haas H, Harris D, Horiuchi H, Huang J, Humphray S, Jiménez J, Keller N, Khouri H, Kitamoto K, Kobayashi T, Konzack S, Kulkarni R, Kumagai T, Lafton A, Latgé JP, Li W, Lord A, Lu C, et al. 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus. Nature 438:1151–1156. https://doi.org/10.1038/nature04332.
18. Hubka V, Barrs V, Dudová Z, Sklenář F, Kubátová A, Matsuzawa T, Yaguchi T, Horie Y, Nováková A, Frisvad JC, Talbot JJ, Kolařík M. 2018. Unravelling species boundaries in the aspergillus viridinutans complex (Section fumigati): opportunistic human and animal pathogens capable of interspecific hybridization. Persoonia 41:142–174. https://doi.org/10.3767/persoonia.2018.41.08.
19. Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol 20:238. https://doi.org/10.1186/s13059-019-1832-y.
20. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/molbev/mst010.
21. Kück P, Meusemann K. 2010. FASconCAT: convenient handling of data matrices. Mol Phylogenet Evol 56:1115–1118. https://doi.org/10.1016/j.ympev.2010.04.024.
22. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R, Teeling E. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534. https://doi.org/10.1093/molbev/msaa015.