AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Evolution | Announcement

High-quality draft genome sequence of the tropical lichen-forming fungus *Lecanora helva* (Lecanoraceae, Ascomycota)

Yukun Sun,[1] Meredith M. Doellman,[1] Alejandrina Barcenas-Peña,[1] Vasun Poengsungnoen,[2] Sabine Huhndorf,[1] H. Thorsten Lumbsch,[1] Felix Grewe[1]

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT We present a high-quality genome assembly of the lichenized fungus *Lecanora helva* collected in a tropical mangrove in Thailand. Using PacBio HiFi sequencing and a novel multi-assembler approach, we generated a contiguous, nearly complete (30.07 Mb, 98.1% BUSCO completeness) genome sequence, enabling future studies on mangrove adaptation and climate resilience.

KEYWORDS Lecanora helva, lichenized fungi, whole genome assembly, telomere-to-telomere assembly

Lecanora is one of the largest genera of crustose lichens, with a worldwide distribution (1, 2). We sequenced the genome of the fungal component (mycobiont) of *Lecanora helva* Stizenb., a tropical lichen species potentially sensitive to climate change. The specimen was collected on 07 August 2021 from the bark of *Avicennia alba* Blume in a mangrove forest in Tambon Lat Khwang, Amphoe Ban Pho, Chachoengsao Province, Thailand (13°36′41.9″N, 101°02′21.4″E), and deposited in the Lichen Herbarium of Ramkhamhaeng University (accession: RAMK 036147; collection number: VPMG_37; herbarium database: http://www.lichen.ru.ac.th/images/data/STT/STT46_Lichen%20herbarium%20database%20and%20management.pdf). An axenic fungal culture was produced from a single ascospore and grown on malt-yeast extract agar under laboratory conditions (3).

High-molecular-weight DNA was extracted from the axenic fungal culture using a modified protocol in Wilken *et al.* (4) and sent to the UWBC DNA Sequencing Facility. DNA quality and quantity were measured using a NanoDrop One (ThermoFisher Scientific, Waltham, MA, USA), a Qubit dsDNA High Sensitivity kit, and a FemtoPulse System (Agilent, Santa Clara, CA, USA). A PacBio HiFi library was prepared using protocol PN 102-166-600 APR2022 (Pacific Biosciences, Menlo Park, CA, USA), with modifications including Covaris gTUBE shearing (Covaris, LLC, Woburn, MA, USA) without further size selection. Library quality was assessed using the FemtoPulse, and the library was quantified using the Qubit dsDNA High Sensitivity kit. Sequencing was performed on a Sequel II using the Sequel Polymerase Binding Kit 2.2. A total of 620,494 raw reads were generated (N_{50} =8148); adapter contamination ("AAGCAGTGGTATCAACGCAGAGTACT") was removed with lima v.2.7.1, and duplicates were marked with pbmarkdup v.1.0.3 (5).

The assembly was constructed with a novel consensus-based strategy combining outputs from multiple assemblers. PacBio HiFi reads were assembled into unphased contigs using six tools: Canu v.2.2 (6), NextDenovo v.2.5.2 (7), IPA (8), Flye v.2.9.5-b1801 (9), Peregrine-2021 v.0.4.13 (10), and RAFT-hifiasm v.0.19.9-r616 (11–14). The Peregrine-2021 output was selected as the primary assembly due to its lowest contig number, highest N_{50}, and greater number of identified telomeric ends, indicating superior genome contiguity. To obtain additional telomere-to-telomere contigs, contigs with telomeric sequences at one terminus from alternative assemblies were compared

Editor Jason E. Stajich, University of California Riverside, Riverside, California, USA

Address correspondence to Yukun Sun, ysun@fieldmuseum.org.

Yukun Sun and Meredith M. Doellman contributed equally to this article. The author order was determined by corresponding authorship.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 13 June 2025
Accepted 26 August 2025
Published 13 October 2025

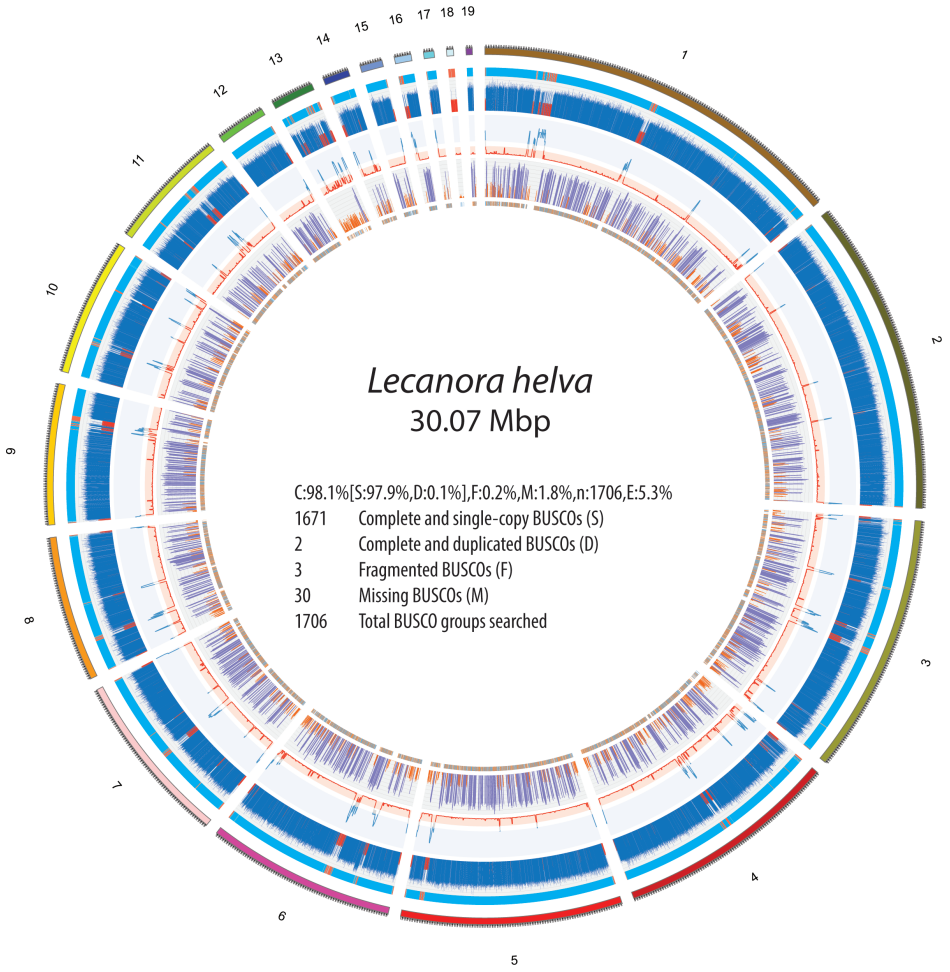

FIG 1 Circular representation of *L. helva* genomic features. The plot comprises concentric tracks, organized from inner to outer rings: gene locations, with forward strand genes in orange and reverse strand genes in blue; gene density (1/10 k base pair), highlighting regions with high density (>0.7) in purple and low density (<0.3) in orange; repeat density (1/10 k base pair), illustrating areas with high repeat content (>0.7) in blue and low repeat content (<0.3) in red; GC content, marking regions with high GC content (>0.5) in blue and low GC content (<0.3) in red; and GC regions, identifying AT-rich regions (with GC content 0–37.9%) in red and higher GC content (37.9–100%) in cyan. In the center of the plot, the BUSCO completeness assessment of the assembled genome using the Ascomycota lineage data set (ascomycota_odb10) is displayed. Scores indicate the percentages of complete (single-copy and duplicated), fragmented, and missing orthologs.

and merged with QuickMerge v.0.3 (15). Resulting contigs were manually curated; Redundans v2.0.1 (16) was used for redundancy reduction. Genome contiguity and completeness were assessed using QUAST v.5.3.0 and BUSCO v.5.7.0 with the ascomycota_odb10 data set (17, 18). Repeats were identified and masked before annotation using RepeatModeler 2.0.1 (19), followed by RepeatMasker 4.1.2-p1 to mask repetitive elements (20). Annotation was performed using Funannotate v.1.8.17 and InterProScan v.5.72-103.0 (21, 22). Default parameters were used except where otherwise noted.

Our final *L. helva* genome assembly is 30.07 Mb in total length, distributed across 19 contigs, with an N_{50} of 2.68 Mb (Fig. 1). The estimated coverage is 140.99 ×, and the GC content is 46.87%. BUSCO analysis indicated 98.1%. A total of 9,745 protein-coding genes were predicted. Our assembly method, which leverages the strengths of multiple assemblers, will support further work on the adaptation of *L. helva* to mangrove habitats and the conservation of tropical mangrove forest diversity in the face of climate change.

ACKNOWLEDGMENTS

The authors utilized the University of Wisconsin–Madison Biotechnology Center's DNA Sequencing Facility (Research Resource Identifier – RRID: SCR_017759) to generate and sequence PacBio libraries. We thank S. Phokaeo and M. Molsil for collecting the specimen with one of us.

AUTHOR AFFILIATIONS

[1]Grainger Bioinformatics Center, Collections, Conservation and Research Division, The Field Museum, Chicago, Illinois, USA
[2]Lichen Research Unit, Department of Biology, Faculty of Science, Ramkhamhaeng University, Bangkapi, Bangkok, Thailand

AUTHOR ORCIDs

Yukun Sun http://orcid.org/0000-0002-1493-5558
Meredith M. Doellman http://orcid.org/0000-0002-0903-6590
Alejandrina Barcenas-Peña http://orcid.org/0000-0003-1674-1164
Vasun Poengsungnoen http://orcid.org/0000-0001-8905-1098
H. Thorsten Lumbsch http://orcid.org/0000-0003-1512-835X
Felix Grewe http://orcid.org/0000-0002-2805-5930

FUNDING

Funder	Grant(s)	Author(s)
Grainger Foundation		Yukun Sun

AUTHOR CONTRIBUTIONS

Yukun Sun, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – review and editing | Meredith M. Doellman, Data curation, Writing – original draft, Writing – review and editing | Alejandrina Barcenas-Peña, Data curation, Validation, Writing – review and editing | Vasun Poengsungnoen, Resources, Writing – review and editing | Sabine Huhndorf, Methodology, Resources | H. Thorsten Lumbsch, Funding acquisition, Investigation, Writing – review and editing | Felix Grewe, Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

This Whole Genome Shotgun project has been deposited in DDBJ/ENA/GenBank under the BioProject number PRJNA1209903. The genome accession number for the version described in this announcement is JBKOTY000000000. PacBio reads can be found at the Sequence Read Archive accession number SRR33792350. This study involved sequencing fungal material only and did not require ethical approval by an institutional ethics committee.

REFERENCES

1. Papong K, Lumbsch HT. 2011. A taxonomic survey of Lecanora sensu stricto in Thailand (Lecanoraceae; Ascomycota). The Lichenologist 43:299–320. https://doi.org/10.1017/S0024282911000247
2. Papong K, Boonpragob K, Parnmen S, Lumbsch HT. 2013. Molecular phylogenetic studies on tropical species of Lecanora sensu stricto (Lecanoraceae, Ascomycota). nova_hedwigia 96:1–13. https://doi.org/10.1127/0029-5035/2012/0072
3. Rosabal D, Pino-Bodas R. 2024. A review of laboratory requirements to culture lichen mycobiont species. J Fungi (Basel) 10:621. https://doi.org/10.3390/jof10090621
4. Wilken PM, Aylward J, Chand R, Grewe F, Lane FA, Sinha S, Ametrano C, Distefano I, Divakar PK, Duong TA, Huhndorf S, Kharwar RN, Lumbsch HT, Navathe S, Pérez CA, Ramírez-Berrutti N, Sharma R, Sun Y, Wingfield BD, Wingfield MJ. 2020. IMA Genome - F13. IMA Fungus 11:19. https://doi.org/10.1186/s43008-020-00039-7
5. Armintoepfer A. 2020. pbmarkdup: Mark duplicate reads from PacBio sequencing of an amplified library. GitHub. Available from: https://github.com/PacificBiosciences/pbmarkdup. Retrieved 20 Jul 2022.
6. Nurk S, Walenz BP, Rhie A, Vollger MR, Logsdon GA, Grothe R, Miga KH, Eichler EE, Phillippy AM, Koren S. 2020. HiCanu: accurate assembly of

segmental duplications, satellites, and allelic variants from high-fidelity long reads. Genome Res 30:1291–1305. https://doi.org/10.1101/gr.263566.120
7. Hu J, Wang Z, Sun Z, Hu B, Ayoola AO, Liang F, Li J, Sandoval JR, Cooper DN, Ye K, Ruan J, Xiao C-L, Wang D, Wu D-D, Wang S. 2024. NextDenovo: an efficient error correction and accurate assembly tool for noisy long reads. Genome Biol 25:107. https://doi.org/10.1186/s13059-024-03252-4
8. PacificBiosciences. 2022. pbipa: Improved Phased Assembler (v1.8.0). GitHub. Available from: https://github.com/PacificBiosciences/pbipa. Retrieved 10 Jun 2024.
9. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8
10. Chin C-S, Khalak A. 2019. Human genome assembly in 100 minutes. bioRxiv. https://doi.org/10.1101/705616
11. Kamath SS, Bindra M, Pal D, Jain C. 2024. Telomere-to-telomere assembly by preserving contained reads. Genome Res 34:1908–1918. https://doi.org/10.1101/gr.279311.124
12. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. Nat Methods 18:170–175. https://doi.org/10.1038/s41592-020-01056-5
13. Cheng H, Jarvis ED, Fedrigo O, Koepfli K-P, Urban L, Gemmell NJ, Li H. 2022. Haplotype-resolved assembly of diploid genomes without parental data. Nat Biotechnol 40:1332–1335. https://doi.org/10.1038/s41587-022-01261-x
14. Cheng H, Asri M, Lucas J, Koren S, Li H. 2024. Scalable telomere-to-telomere assembly for diploid and polyploid genomes with double graph. Nat Methods 21:967–970. https://doi.org/10.1038/s41592-024-02269-8
15. Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. 2016. Contiguous and accurate *de novo* assembly of metazoan genomes with modest long read coverage. Nucleic Acids Res 44:e147. https://doi.org/10.1093/nar/gkw654
16. Pryszcz LP, Gabaldón T. 2016. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Res 44:e113. https://doi.org/10.1093/nar/gkw294
17. 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
18. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol 38:4647–4654. https://doi.org/10.1093/molbev/msab199
19. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci USA 117:9451–9457. https://doi.org/10.1073/pnas.1921046117
20. Smit AFA, Hubley R, Green P. 2013–2015. RepeatMasker Open-4.0. RepeatMasker. Available from: http://www.repeatmasker.org. Retrieved 31 Jul 2025.
21. Palmer JM, Stajich JE. 2020. Funannotate v1.8.1: Eukaryotic Genome Annotation (v1.8). Zenodo. Available from: https://doi.org/10.5281/zenodo.1134477
22. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240. https://doi.org/10.1093/bioinformatics/btu031