



# Draft Genome Sequence of the Globally Distributed Cockroach-Infesting Fungus *Herpomyces periplanetae* Strain D. Haelew. 1187d

 Danny Haelewaters,<sup>a,b\*</sup>  Alicja Okrańska,<sup>c</sup>  Michał Gorczak,<sup>c</sup>  Donald H. Pfister<sup>a</sup>

<sup>a</sup>Farlow Reference Library and Herbarium of Cryptogamic Botany, Harvard University, Cambridge, Massachusetts, USA

<sup>b</sup>Department of Biology, Research Group Mycology, Ghent University, Ghent, Belgium

<sup>c</sup>Department of Molecular Phylogenetics and Evolution, Faculty of Biology, University of Warsaw, Warsaw, Poland

**ABSTRACT** *Herpomyces periplanetae* is an obligate biotroph of *Periplaneta americana*, the American cockroach. Its nearly cosmopolitan distribution is shaped by its globally invasive host and the international pet trade. Here, we report the draft genome sequence of *H. periplanetae*, based on a thallus from *P. americana* collected in Cambridge, Massachusetts.

*Herpomyces* (Ascomycota, Pezizomycotina, Laboulbeniomycetes, Herpomycetales, Herpomycetaceae) is a genus of biotrophic fungi obligately associated with cockroaches (Arthropoda, Blattodea) (1–3). The distribution of at least some *Herpomyces* taxa has greatly expanded because of their globally invasive host species, which are used in the international pet trade (4). Recent molecular phylogenetic analyses have shown that *Herpomyces* species are distinct from members of the order Laboulbeniales, in agreement with developmental and morphological evidence (3, 5, 6). Species of both Herpomycetales and Laboulbeniales are unique among related fungi in that they do not form hyphae but instead produce multicellular units of hundreds to thousands of cells, referred to as “thalli,” through coordinated cell divisions from a single ascospore (3, 6–8). To date, these fungi have not been grown apart from their host; DNA needs to be isolated from individual thalli removed from the host (9–11).

Here, we sequenced and assembled the genome of *H. periplanetae*, isolate D. Haelew. 1187d, based on material removed from a *Periplaneta americana* specimen collected in Cambridge, Massachusetts (3). This genome is based on a single thallus (with three perithecia). Isolation and whole-genome amplification of genomic DNA (gDNA) were done using the REPLI-g single cell kit (Qiagen, Stanford, CA) (3). The single thallus was removed from the host and placed in a 0.2-ml PCR tube with 2.0  $\mu$ l of phosphate-buffered saline (PBS) solution. After adding 1.5  $\mu$ l of prepared denaturation (D2) solution, the tube was incubated at 65°C for 20 min, immediately followed by the addition of 1.5  $\mu$ l of neutralizing REPLI-g STOP solution. Whole-genome amplification by multiple displacement amplification (MDA) was done in a 25- $\mu$ l reaction mixture with 14.5  $\mu$ l of REPLI-g reaction buffer, 1.0  $\mu$ l of REPLI-g DNA polymerase, 4.5  $\mu$ l of double-distilled water (ddH<sub>2</sub>O), and 5.0  $\mu$ l of denatured DNA, under the following cycling conditions: incubation at 30°C for 6 h, followed by inactivation of the DNA polymerase at 65°C for 3 min. Following extraction and purification by ethanol precipitation, gDNA concentration and integrity were evaluated by agarose gel electrophoresis and using both a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and a Qubit double-stranded DNA (dsDNA) broad-range assay kit on a Qubit Fluorometer (Thermo Fisher Scientific). The NanoDrop values were within the required range for further

**Citation** Haelewaters D, Okrańska A, Gorczak M, Pfister DH. 2020. Draft genome sequence of the globally distributed cockroach-infesting fungus *Herpomyces periplanetae* strain D. Haelew. 1187d. *Microbiol Resour Announc* 9:e01458-19. <https://doi.org/10.1128/MRA.01458-19>.

**Editor** Antonis Rokas, Vanderbilt University

**Copyright** © 2020 Haelewaters et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Danny Haelewaters, [danny.haelewaters@gmail.com](mailto:danny.haelewaters@gmail.com).

\* Present address: Danny Haelewaters, Department of Botany and Plant Pathology, Purdue University, West Lafayette, Indiana, USA.

**Received** 9 December 2019

**Accepted** 18 January 2020

**Published** 6 February 2020

processing ( $A_{260}/A_{280} = 1.79$ ,  $A_{260}/A_{230} = 2.01$ ), and the gDNA concentration as measured by the Qubit was 4.62 ng/ $\mu$ l. Identity and DNA isolation were validated by the amplification and sequencing of the following three ribosomal DNA fragments (3, 12): the small subunit (SSU) (GenBank accession number [MG438331](#)), internal transcribed spacer (ITS) ([MG438309](#)), and large subunit (LSU) ([MG438359](#)). Library preparation was done using the PrepX DNA library kit on the Apollo NGS library prep system (TaKaRa Bio, Mountain View, CA) following the manufacturer's specifications and with 50  $\mu$ l of gDNA as the starting volume. Paired-end sequencing ( $2 \times 125$  bp) was performed on an Illumina HiSeq 2500 instrument at the Harvard University Bauer Core Facility, resulting in a total of 9,793,875 raw paired-end reads.

Read quality was assessed using FastQC version 0.11.5 (13). Adapters were removed using Trimmomatic version 0.32 (14). Low-quality sequences were not trimmed, because this resulted in a poorer assembly (with more contigs and a lower  $N_{50}$  value). The remaining 8.5 million Illumina reads were *de novo* assembled using SPAdes version 3.10.1 (15) with progressive k-mer numbers, ranging from 21 to 117. The final genome assembly included 968 contigs (959 scaffolds), of which 632 (628) were longer than 1 kbp. After discarding scaffolds shorter than 1 kbp, the size of the genome was estimated to be 15,006,634 bp, with the length of the largest contig being 347,628 bp, an  $N_{50}$  value of 95,243 kbp, an  $L_{50}$  value of 50, and a GC content of 40.23%. These statistics were calculated using QUAST version 4.5 (16). To assess the completeness of our genome, we used BUSCO version 3.0.2 (17) with the fungal\_odb9 database and a default species as the model. We retrieved 85% out of 290 orthologs present in the database (237 complete, 11 fragmented). This genome of *H. periplanetae* will enable studies of fungal genome evolution, host-parasite evolutionary genomics, patterns of speciation, and thallus development.

**Data availability.** The whole-genome shotgun sequencing project for the *Herpomyces periplanetae* isolate D. Haelew. 1187d has been deposited at DDBJ/ENA/GenBank under the accession number [WМКJ00000000](#). The version described in this paper is version WМКJ01000000. The Illumina raw reads have been deposited at the Sequence Read Archive under accession number [SRX7128323](#).

## ACKNOWLEDGMENTS

This research was supported by a Friends of the Farlow grant to D.H.

We gratefully acknowledge the help of Claire Hartmann (Reardon) of the Harvard University FAS Division of Science Bauer Core Facility and Tim Sackton of the Harvard University FAS Division of Science Informatics Group.

## REFERENCES

- Thaxter R. 1902. Preliminary diagnoses of new species of Laboulbeniaceae. *V. Proc Am Acad Arts Sci* 38:9–57. <https://doi.org/10.2307/20021736>.
- Wang TW, de Kesel A, Haelewaters D, Pfister DH. 2016. Farlow Herbarium cockroach hosts new record of Laboulbeniales for North America. *Rhodor* 118:26–31. <https://doi.org/10.3119/15-20>.
- Haelewaters D, Pfliegler WP, Gorczak M, Pfister DH. 2019. Birth of an order: comprehensive molecular phylogenetic study excludes *Herpomyces* (Fungi, Laboulbeniomycetes) from Laboulbeniales. *Mol Phylogenet Evol* 133:286–301. <https://doi.org/10.1016/j.ympev.2019.01.007>.
- Pfliegler WP, Báthori F, Wang TW, Tartally A, Haelewaters D. 2018. *Herpomyces* ectoparasitic fungi (Ascomycota, Laboulbeniales) are globally distributed by their invasive cockroach hosts and through the pet trade industry. *Mycologia* 110:39–46. <https://doi.org/10.1080/00275514.2017.1418567>.
- Tavares II. 1980. Notes on perithecial development in the Euceratomycetaceae fam. nov. (Laboulbeniales, Laboulbeniineae) and *Herpomyces* (Herpomycetinae). *Mycotaxon* 11:485–492.
- Tavares II. 1985. Laboulbeniales (Fungi, Ascomycetes). *Mycol Mem* 9:1–627.
- Thaxter R. 1896. Contribution towards a monograph of the Laboulbeniaceae. *Mem Am Acad Arts Sci* 12:187–429.
- Benjamin RK. 1971. Introduction and supplement to Roland Thaxter's contribution towards a monograph of the Laboulbeniaceae. *Biblioth Mycol* 30:1–155.
- Weir A, Blackwell M. 2001. Extraction and PCR amplification of DNA from minute ectoparasitic fungi. *Mycologia* 93:802–806. <https://doi.org/10.1080/00275514.2001.12063212>.
- Haelewaters D, Gorczak M, Pfliegler WP, Tartally A, Tischer M, Wrzosek M, Pfister DH. 2015. Bringing Laboulbeniales into the 21st century: enhanced techniques for extraction and PCR amplification of DNA from minute ectoparasitic fungi. *IMA Fungus* 6:363–372. <https://doi.org/10.5598/ima fungus.2015.06.02.08>.
- Sundberg H, Ekman S, Krusys Å. 2018. A crush on small fungi: an efficient and quick method for obtaining DNA from minute ascomycetes. *Methods Ecol Evol* 9:148–158. <https://doi.org/10.1111/2041-210X.12850>.
- Haelewaters D, De Kesel A, Pfister DH. 2018. Integrative taxonomy reveals hidden species within a common fungal parasite of ladybirds. *Sci Rep* 8:15966. <https://doi.org/10.1038/s41598-018-34319-5>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 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>.

15. 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>.
16. 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>.
17. 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>.