

De Novo Whole-Genome Sequence and Genome Annotation of *Lichtheimia ramosa*

Jörg Linde,^a Volker Schwartze,^b Ulrike Binder,^c Cornelia Lass-Flörl,^c Kerstin Voigt,^b Fabian Horn^a

Systems Biology/Bioinformatics, Hans-Knöll-Institut, Jena, Germany^a; Jena Microbial Resource Collection (JMRC), Hans-Knöll-Institut, Jena, Germany^b; Division of Hygiene and Medical Microbiology, Innsbruck Medical University, Innsbruck, Austria^c

We report the annotated draft genome sequence of *Lichtheimia ramosa* (JMRC FSU:6197). It has been reported to be a causative organism of mucormycosis, a rare but rapidly progressive infection in immunocompromised humans. The functionally annotated genomic sequence consists of 74 scaffolds with a total number of 11,510 genes.

Received 7 August 2014 Accepted 11 August 2014 Published 11 September 2014

Citation Linde J, Schwartze V, Binder U, Lass-Flörl C, Voigt K, Horn F. 2014. De novo whole-genome sequence and genome annotation of *Lichtheimia ramosa*. *Genome Announc.* 2(5):e00888-14. doi:10.1128/genomeA.00888-14.

Copyright © 2014 Linde et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Kerstin Voigt, kerstin.voigt@hki-jena.de, or Fabian Horn, fabian.horn@hki-jena.de.

Lichtheimia ramosa (formerly *Absidia idahoensis* var. *thermophila*, *L. hongkongensis*) belongs to the order of *Mucorales* (1). Besides *L. ramosa*, *L. ornata* and *L. corymbifera* are clinically relevant (1, 2, 3, 4). The virulence potential of this fungus is connected with thermotolerance (5), because other clinically nonrelevant *Lichtheimia* species possess a lower thermotolerance and stop growth at 42°C (1, 6). So far, missing genomic data has hindered the exploration of further virulence factors (7, 8). Here we present the full genome sequence of *L. ramosa*, whereas the mitogenome was announced recently (9).

DNA was obtained from mycelia cultured in liquid supplemented minimal medium (SUP medium) under shaken conditions for 3 days at 37°C (10). One library was prepared for 8-kb Roche/454PE GS FLX+ Titanium sequencing and a second library for Illumina HiSeq 2000 100-bp PE sequencing. Genome sequencing and assembly was generated by LGC Genomics (Berlin) using a hybrid approach. Illumina contigs, assembled by Velvet (11), and 454 scaffolds, assembled by Newbler 2.6 (454 Life Sciences), were merged using Minimus2 (12). The resulting scaffolds were finalized using SOAP GapCloser (13) and SEQuel (14). RNA-Seq data were obtained from a pooled sample cultured under five different conditions. Transcriptome sequencing was performed using Roche/454 GS FLX+ Titanium, and contigs were assembled using Newbler.

For gene prediction, the pipeline presented by Haas et al. (15) was customized, and tools incorporating *ab initio* models, transcriptome data, and protein alignments were applied. The parameter sets were trained using gene models that were predicted by TransDecoder (16) from aligned species-specific transcripts. All gene predictions were combined using EvidenceModeler. Untranscribed regions were added using PASA (17).

For *ab initio* gene prediction, GeneMark-ES (18), Augustus (19), SNAP (20), and Glimmer (21) were applied. Transcriptome data were incorporated into Augustus, FGENESH (22), and PASA. Protein alignments were obtained by mapping proteins from *L. hyalospora* (JGI), *Rhizopus delemar* (BROAD), *Rhizopus microsporus* var. *microsporus* (JGI), *Mucor circinelloides* (JGI), and

Phycomyces blakesleeanus (JGI) using Exonerate (23) and Scipio (24).

Genes were functionally annotated using Blast2GO (25) and InterProScan (26), including the TMHMM (27) option. Gene descriptions were obtained by blasting the predicted protein sequences against the fungal UniProt Knowledgebase (28). Secondary metabolite gene clusters were predicted using SMURF (29).

454 DNA sequencing resulted in 1,345,023 reads (760 Mbp; estimated genome coverage, 24.3-fold). Illumina DNA sequencing resulted in 426,388,592 raw reads, where 45,982,894 reads passed stringent quality filters (4.10 Gbp; estimated genome coverage, 130-fold) and have been used to create the final assembly. The assembly consists of 74 scaffolds and 30.71 Mbp (N_{50} , 1.22 Mbp; N_{90} , 338 kbp). The G+C content of the assembly is 41.2%. RNA sequencing and transcriptome assembly led to 12,134 transcripts (11.19 Mbp; estimated transcriptome coverage, 0.5-fold). The final gene prediction consists of 11,510 genes and 11,546 transcripts, and 452 (98.7%) eukaryotic core proteins were identified using CEGMA (30). The coding density of the genome is 52%. Functional names were assigned to 980 transcripts, gene ontology categories to 6,899 transcripts, and protein domains to 9,664 translated transcripts; 2,645 transcripts were predicted to contain transmembrane domains, and 38 transcripts have been assigned to three secondary metabolite gene clusters.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession numbers [LK023313](https://www.ncbi.nlm.nih.gov/nuclink/LK023313) to [LK023386](https://www.ncbi.nlm.nih.gov/nuclink/LK023386). The version described in this paper is the first version. Genome data and additional information are also available at the HKI (Hans-Knöll-Institute) Genome Resource (<http://www.genome-resource.de>).

ACKNOWLEDGMENTS

Funding was provided by the Era-Net PathoGenoMics Project “OXY-stress: Human fungal pathogens under oxygen stress—adaptive mechanisms to hypoxia and reactive oxygen species and their consequences for host interaction and therapy,” funded by the Austrian FWF, I661-B09 to C.L.F. J.L. was supported by the Deutsche Forschungsgemeinschaft

(DFG) CRC/Transregio 124 project “Pathogenic fungi and their human host: Networks of interaction,” subproject INF.

REFERENCES

- Alastruey-Izquierdo A, Hoffmann K, de Hoog GS, Rodriguez-Tudela JL, Voigt K, Bibashi E, Walther G. 2010. Species recognition and clinical relevance of the Zygomycetous genus *Lichtheimia* (syn. *Absidia pro parte*, *Mycocladas*). *J Clin Microbiol.* 48:2154–2170. <http://dx.doi.org/10.1128/JCM.01744-09>.
- Binder U, Maurer E, Lass-Flörl C. 2014. Mucormycosis—from the pathogens to the disease. *Clin. Microbiol. Infect.* 20:60–66. <http://dx.doi.org/10.1111/1469-0691.12566>.
- García-Hermoso D, Hoinard D, Gantier JC, Grenouillet F, Dromer F, Dannaoui E. 2009. Molecular and phenotypic evaluation of *Lichtheimia corymbifera* (formerly *Absidia corymbifera*) complex isolates associated with human mucormycosis: rehabilitation of *L. ramosa*. *J. Clin. Microbiol.* 47:3862–3870. <http://dx.doi.org/10.1128/JCM.02094-08>.
- Kutlu M, Ergin C, Bir F, Hilmioğlu-Polat S, Gümrall R, Necan C, Koçyigit A, Sayin-Kutlu S. 2014. Pulmonary mucormycosis due to *Lichtheimia ramosa* in a patient with HIV infection. *Mycopathologia* 178: 111–115. <http://dx.doi.org/10.1007/s11046-014-9761-5>.
- Schwartz VU, Hoffmann K, Nyilasi I, Papp T, Vágvölgyi C, de Hoog S, Voigt K, Jacobsen ID. 2012. *Lichtheimia* species exhibit differences in virulence potential. *PLoS One* 7:e40908. <http://dx.doi.org/10.1371/journal.pone.0040908>.
- Hoffmann K, Voigt K. 2011. *Lichtheimia* (*Absidia*-like fungi), p 735–748. In Liu D (ed), *Molecular detection of human fungal pathogens*. CRC Press, Boca Raton, FL.
- Mendoza L, Vilela R, Voelz K, Ibrahim AS, Voigt K, Lee SC. Chapter 27. Human fungal pathogens of Mucorales and Entomophthorales. 27. In Casadevall A, Mitchell AP, Berman J, Kwon-Chung KJ, Perfect JR, Heitman T (ed). *Human fungal pathogens*, in press. Cold Spring Harbor Press, Cold Spring Harbor, NY.
- Shelest E, Voigt K. 2014. Genomics to study basal lineage fungal biology: phylogenomics suggests a common origin, p 31–60. In Nowrousian M (ed), *The mycota*, vol XIII, 2nd ed., vol XIII. Springer, Berlin.
- Leung S-Y, Huang Y, Lau SKP, Woo PCY. 2014. Complete mitochondrial genome sequence of *Lichtheimia ramosa* (syn. *Lichtheimia hongkongensis*). *Genome Announc.* 2(4):e00644-14. <http://dx.doi.org/10.1128/genomeA.00644-14>.
- Wöstemeyer J. 1985. Strain-dependent variation in ribosomal DNA arrangement in *Absidia glauca*. *Eur. J. Biochem.* 146:443–448. <http://dx.doi.org/10.1111/j.1432-1033.1985.tb08671.x>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
- Sommer DD, Delcher AL, Salzberg SL, Pop M. 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics* 8:64. <http://dx.doi.org/10.1186/1471-2105-8-64>.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience* 1:18. <http://dx.doi.org/10.1186/2047-217X-1-18>.
- Ronen R, Boucher C, Chitsaz H, Pevzner P. 2012. SEQuel: improving the accuracy of genome assemblies. *Bioinformatics* 28:i188–i196. <http://dx.doi.org/10.1093/bioinformatics/bts219>.
- Haas BJ, Zeng Q, Pearson MD, Cuomo CA, Wortman JR. 2011. Approaches to fungal genome annotation. *Mycology* 2:118–141.
- Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the program to assemble spliced alignments. *Genome Biol.* 9:R7. <http://dx.doi.org/10.1186/gb-2008-9-1-r7>.
- Borodovsky M, Lomsadze A. 2011. Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES. *Curr. Protoc. Bioinformatics* 35(4.6):4.6.1–4.6.10. <http://dx.doi.org/10.1002/0471250953.bi0406s35>.
- Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntentically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics* 24:637–644. <http://dx.doi.org/10.1093/bioinformatics/btn013>.
- Korf I. 2004. Gene finding in novel genomes. *BMC Bioinformatics* 5:59. <http://dx.doi.org/10.1186/1471-2105-5-59>.
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* 27: 4636–4641. <http://dx.doi.org/10.1093/nar/27.23.4636>.
- Salamov AA, Solovyev VV. 2000. Ab initio gene finding in *Drosophila* genomic DNA. *Genome Res.* 10:516–522. <http://dx.doi.org/10.1101/gr.10.4.516>.
- Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD, Salzberg SL, White O. 2003. Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* 31:5654–5666. <http://dx.doi.org/10.1093/nar/gkg770>.
- Slater GS, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6:31. <http://dx.doi.org/10.1186/1471-2105-6-31>.
- Keller O, Odronitz F, Stanke M, Kollmar M, Waack S. 2008. Scipio: using protein sequences to determine the precise exon/intron structures of genes and their orthologs in closely related species. *BMC Bioinformatics* 9:278. <http://dx.doi.org/10.1186/1471-2105-9-278>.
- Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676. <http://dx.doi.org/10.1093/bioinformatics/bti610>.
- Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. *Nucleic Acids Res.* 33:W116–W120. <http://dx.doi.org/10.1093/nar/gki442>.
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305:567–580. <http://dx.doi.org/10.1006/jmbi.2000.4315>.
- Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS. 2004. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* 32:D115–D119. <http://dx.doi.org/10.1093/nar/gkh131>.
- Khalidi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, Fedorova ND. 2010. SMURF: genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* 47:736–741. <http://dx.doi.org/10.1016/j.fgb.2010.06.003>.
- Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23:1061–1067. <http://dx.doi.org/10.1093/bioinformatics/btm071>.