

Draft Genome Sequences of Symbiotic and Nonsymbiotic *Rhizopus microsporus* Strains CBS 344.29 and ATCC 62417

Fabian Horn,^a Zerrin Üzümlü,^b Nadine Möbius,^b Reinhard Guthke,^a Jörg Linde,^a Christian Hertweck^b

Systems Biology/Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute, Jena, Germany^a; Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute, Jena, Germany^b

Specific *Rhizopus microsporus* pathovars harbor bacterial endosymbionts (*Burkholderia rhizoxinica*) for the production of a phytotoxin. Here, we present the draft genome sequences of two *R. microsporus* strains, one symbiotic (ATCC 62417), and one endosymbiont-free (CBS 344.29). The gene predictions were supported by RNA sequencing (RNA-seq) data. The functional annotation sets the basis for comparative analyses.

Received 18 November 2014 Accepted 9 December 2014 Published 22 January 2015

Citation Horn F, Üzümlü Z, Möbius N, Guthke R, Linde J, Hertweck C. 2015. Draft genome sequences of symbiotic and nonsymbiotic *Rhizopus microsporus* strains CBS 344.29 and ATCC 62417. *Genome Announc.* 3(1):e01370-14. doi:10.1128/genomeA.01370-14.

Copyright © 2015 Horn 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 Jörg Linde, joerg.linde@hki-jena.de, or Christian Hertweck, christian.hertweck@hki-jena.de.

The zygomycete *Rhizopus microsporus* is a terrestrial filamentous fungus that is used in food fermentation but also is known to cause rice seedling blight, which leads to substantial losses in agriculture, as well as mucormycoses in immunocompromised patients (1, 2). Phytopathogens, such as *R. microsporus* strain ATCC 62417, harbor endosymbiotic bacteria of the genus *Burkholderia* (3, 4) producing the antimetabolic agent rhizoxin (5–7). Surprisingly, these symbiotic fungi lose their ability to produce mature sporangia and vegetative spores in the absence of the endobacteria (8). An analysis of the genome sequence of the endosymbiont *Burkholderia rhizoxinica* has revealed bacterial symbiosis factors (9–12). Besides information gained from genetic barcode studies (13), nothing is known about the genomic differences between symbiotic and symbiont-free *R. microsporus* strains.

For each strain, multiple libraries (paired-end, 2-kb mate-pair, 3 × 5-kb mate-pair) were sequenced with Illumina HiSeq 2000. *R. microsporus* CBS 344.29 was additionally sequenced using Illumina MiSeq V2. The reads were quality trimmed, error corrected (14), digitally normalized (15), and finally assembled with AllPaths-LG (16) (LGC Genomics, Berlin, Germany). The assemblies were postprocessed using SOAP GapCloser (17). Transcriptome sequencing was performed using HiSeq 2000 100 bp. CBS 344.29 was cultured on potato dextrose agar (PDA) plates for 3 days at 30°C. ATCC 62417 was cultured under three different conditions, namely, with and without the endosymbiont (*B. rhizoxinica*) on PDA plates, and with bacterium in Vorkultur medium at 30°C for 3 days. The culture with Vorkultur medium was sequenced using the 454 FLX Titanium platform, and the contigs were assembled using Newbler (454 Life Sciences). rRNA was detected and removed with riboPicker (18).

Structural and functional gene annotation were performed as described previously (19, 20). Protein sequences from *Lichtheimia hyalospora* (Joint Genome Institute [JGI]), *Mucor circinelloides* (Broad Institute), *Phycomyces blakesleeianus* (JGI), *Rhizopus delemar* (Broad Institute), and *R. microsporus* var. *microsporus* (JGI) were mapped to the reference genomes. The transcriptome assem-

blies of Newbler, Cufflinks (21), and genome-independent and genome-guided Trinity (22) were combined using Evidential-Gene (23). The resulting assembly was checked for contamination with the help of BLAST, the NCBI nonredundant database, and the UniProt fungal knowledge base.

The genome assembly is based on sequencing data amounting to 15.9 Gbp (CBS 344.29) and 15.7 Gbp (ATCC 62417), which represent a 310-fold estimated genome coverage. The assembly for CBS 344.29 consists of 1,554 scaffolds, with a total size of 49.2 Mbp (N_{50} , 53 kbp), whereas the assembly of ATCC 62417 consists of 1,386 scaffolds, with a total size of 49.6 Mbp (N_{50} , 198 kbp). The exceptionally low G+C content of the assemblies are 34.5% and 33.4%, respectively. Using CEGMA (24), we identified 431 and 451 core proteins within the respective genomes. During the genome annotation, we utilized RNA sequencing (RNA-seq) data amounting to 5.5 Mbp (CBS 344.29) and 5.1 Mbp (ATCC 62417), which represent an estimated 100-fold transcriptome coverage. For CBS 344.29, we predicted 19,564 genes and 20,209 transcripts, of which 10,698 have GO categories, 14,642 have InterPro domains, and 3,879 contain transmembrane domains. For ATCC 62417, we predicted 18,869 genes and 23,603 transcripts, of which 9,704 have GO categories, 17,264 have InterPro domains, and 4,807 contain transmembrane domains. The GO annotations of the fungus have been made available for enrichment analysis with FungiFun2 (25).

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession numbers CCYT01000001 to CCYT01001386 (strain ATCC 62417) and CDGI01000001 to CDGI01001554 (strain CBS 344.29). The versions described in this paper are the first versions. The genome data and additional information are also available at the HKI Genome Resource (<http://www.genome-resource.de/>).

ACKNOWLEDGMENTS

This work was financially supported by the Deutsche Forschungsgemeinschaft (DFG), in the CRC/Transregio 124 “Pathogenic fungi and their

human host: Networks of interaction,” subproject INF (to J.L.) and in the Jena School for Microbial Communication (JSMC) (to Z.Ü. and N.M.).

REFERENCES

- Scherlach K, Graupner K, Hertweck C. 2013. Molecular bacteria-fungi interactions: effects on environment, food, and medicine. *Annu Rev Microbiol* 67:375–397. <http://dx.doi.org/10.1146/annurev-micro-092412-155702>.
- Lackner G, Hertweck C. 2011. Impact of endofungal bacteria on infection biology, food safety, and drug development. *PLoS Pathog* 7:e1002096. <http://dx.doi.org/10.1371/journal.ppat.1002096>.
- Partida-Martinez LP, Hertweck C. 2005. Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* 437:884–888. <http://dx.doi.org/10.1038/nature03997>.
- Partida-Martinez LP, Groth I, Roth M, Schmitt I, Buder K, Hertweck C. 2007. *Burkholderia rhizoxinica* and *Burkholderia endofungorum*, bacterial endosymbionts of the rice pathogenic fungus *Rhizopus microsporus*. *Int J Syst Evol Microbiol* 57:2583–2590. <http://dx.doi.org/10.1099/ijs.0.64660-0>.
- Scherlach K, Partida-Martinez LP, Dahse HM, Hertweck C. 2006. Antimitotic rhizoxin derivatives from a cultured bacterial endosymbiont of the rice pathogenic fungus *Rhizopus microsporus*. *J Am Chem Soc* 128:11529–11536. <http://dx.doi.org/10.1021/ja062953o>.
- Scherlach K, Busch B, Lackner G, Paszkowski U, Hertweck C. 2012. Symbiotic cooperation in the biosynthesis of a phytotoxin. *Angew Chem Int Ed Engl* 51:9615–9618. <http://dx.doi.org/10.1002/anie.201204540>.
- Partida-Martinez LP, Hertweck C. 2007. A gene cluster encoding rhizoxin biosynthesis in *Burkholderia rhizoxinica*, the bacterial endosymbiont of the fungus *Rhizopus microsporus*. *ChemBioChem* 8:41–45. <http://dx.doi.org/10.1002/cbic.200600393>.
- Partida-Martinez LP, Monajembashi S, Greulich KO, Hertweck C. 2007. Endosymbiont-dependent host reproduction maintains bacterial-fungal mutualism. *Curr Biol* 17:773–777. <http://dx.doi.org/10.1016/j.cub.2007.03.039>.
- Lackner G, Moebius N, Partida-Martinez LP, Boland S, Hertweck C. 2011. Evolution of an endofungal lifestyle: deductions from the *Burkholderia rhizoxinica* genome. *BMC Genomics* 12:210. <http://dx.doi.org/10.1186/1471-2164-12-210>.
- Leone MR, Lackner G, Silipo A, Lanzetta R, Molinaro A, Hertweck C. 2010. An unusual galactofuranose lipopolysaccharide warrants intracellular survival of toxin-producing bacteria in their fungal host. *Angew Chem Int Ed Engl* 49:7476–7480. <http://dx.doi.org/10.1002/anie.201003301>.
- Lackner G, Moebius N, Hertweck C. 2011. Endofungal bacterium controls its host by an hrp type III secretion system. *ISME J* 5:252–261. <http://dx.doi.org/10.1038/ismej.2010.126>.
- Moebius N, Üzüüm Z, Dijksterhuis J, Lackner G, Hertweck C. 2014. Active invasion of bacteria into living fungal cells. *Elife* 3:e03007. <http://dx.doi.org/10.7554/eLife.03007>.
- Lackner G, Möbius N, Scherlach K, Partida-Martinez LP, Winkler R, Schmitt I, Hertweck C. 2009. Global distribution and evolution of a toxinogenic *Burkholderia-Rhizopus* symbiosis. *Appl Environ Microbiol* 75:2982–2986. <http://dx.doi.org/10.1128/AEM.01765-08>.
- Liu Y, Schröder J, Schmidt B. 2013. Musket: a multistage *k*-mer spectrum-based error corrector for Illumina sequence data. *Bioinformatics* 29:308–315. <http://dx.doi.org/10.1093/bioinformatics/bts690>.
- Crusoe MR, Edverson G, Fish J, Howe A, McDonald E, Nahum J, Nanlohy K, Ortiz-Zuazaga H, Pell J, Simpson J, Scott C, Srinivasan RR, Zhang Q, Brown CT. 2014. The Khmer software package: enabling efficient sequence analysis. *figshare* <http://dx.doi.org/10.6084/m9.figshare.979190>.
- 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. <http://dx.doi.org/10.1073/pnas.1017351108>.
- 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* 27:18. <http://dx.doi.org/10.1186/2047-217X-1-18>.
- Schmieder R, Lim YW, Edwards R. 2012. Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* 28:433–435. <http://dx.doi.org/10.1093/bioinformatics/btr669>.
- 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:e00888-14. <http://dx.doi.org/10.1128/genomeA.00888-14>.
- Linde J, Duggan S, Weber M, Horn F, Sieber P, Hellwig D, Riege K, Marz M, Martin R, Guthke R, Kurzai O. Defining the transcriptomic landscape of *Candida glabrata* by RNA-Seq. *Nucleic Acids Res* [Epub ahead of print.] <http://dx.doi.org/10.1093/nar/gku1357>.
- Trapnell C, Williams BA, Pertea G, Mortazavi AM, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28:511–515. <http://dx.doi.org/10.1038/nbt.1621>.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol* 29:644–652. <http://dx.doi.org/10.1038/nbt.1883>.
- Gilbert D. 2013. Gene-omes built from mRNA-seq not genome DNA: simple, quick, accurate, less cost, more complete gene sets. Indiana University, Bloomington, IN. <http://arthropods.eugenes.org/EvidentialGene/about/EvigeneRNA2013poster.pdf>.
- 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>.
- Priebe S, Kreisel C, Horn F, Guthke R, Linde J. 2014. FungiFun2: a comprehensive online resource for systematic analysis of gene lists from fungal species. *Bioinformatics* [Epub ahead of print.] <http://dx.doi.org/10.1093/bioinformatics/btu627>.