

Genome Sequence and Annotation of *Trichoderma parareesei*, the Ancestor of the Cellulase Producer *Trichoderma reesei*

Dongqing Yang,^{a,b} Kyle Pomraning,^c Alexey Kopchinskiy,^a Razieh Karimi Aghcheh,^a Lea Atanasova,^a Komal Chenthamara,^a Scott E. Baker,^c Ruifu Zhang,^b Qirong Shen,^b Michael Freitag,^d Christian P. Kubicek,^a Irina S. Druzhinina^a

Research Area Biotechnology and Microbiology, Institute of Chemical Engineering, TU Wien, Vienna, Austria^a; Jiangsu Key Lab for Organic Waste Utilization and National Engineering Research Centre for Organic-Based Fertilizers, Nanjing Agricultural University, Nanjing, China^b; Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington, USA^c; Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon, USA^d

D.Y. and K.P. contributed equally to this work.

The filamentous fungus *Trichoderma parareesei* is the asexually reproducing ancestor of *Trichoderma reesei*, the holomorphic industrial producer of cellulase and hemicellulase. Here, we present the genome sequence of the *T. parareesei* type strain CBS 125925, which contains genes for 9,318 proteins.

Received 6 July 2015 Accepted 10 July 2015 Published 13 August 2015

Citation Yang D, Pomraning K, Kopchinskiy A, Karimi Aghcheh R, Atanasova L, Chenthamara K, Baker SE, Zhang R, Shen Q, Freitag M, Kubicek CP, Druzhinina IS. 2015. Genome sequence and annotation of *Trichoderma parareesei*, the ancestor of the cellulase producer *Trichoderma reesei*. *Genome Announc* 3(4):e00885-15. doi:10.1128/genomeA.00885-15.

Copyright © 2015 Yang 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/4.0/).

Address correspondence to Irina S. Druzhinina, irina.druzhinina@tuwien.ac.at.

The filamentous ascomycete *Trichoderma reesei* is used in the biotechnological industry for the production of cellulolytic and hemicellulolytic enzymes and recombinant proteins (1). In nature, *T. reesei* almost exclusively occurs in its sexual form on dead wood (2). Earlier isolations of putative *T. reesei* anamorphs from soil (3, 4) have been shown to be a sympatric sister species that is now named *Trichoderma parareesei* (5). Compared to *T. reesei*, the later taxon has an entirely clonal lifestyle, is considerably more versatile in substrate utilization, and has enhanced mycoparasitic vigor (2, 5).

We sequenced the genome of the type strain of *T. parareesei* CBS 125925 in an Illumina-based whole-genome shotgun sequencing approach delivering 366,865,176 paired reads with an approximate insert size of 350 bp. The acquired sequence reads were assembled into 1,123 contigs using Velvet v1.0.12 (6) with a k-mer length of 75 nucleotides (nt). The resulting genome sequence has an estimated size of 32.0 Mb (N_{50} , 68,608 bp; N_{Max} , 286,763 bp; median coverage, 250.5) with a G+C content of 53.8%.

The genome assembly was repeat masked by RepeatMasker (<http://www.repeatmasker.org>), and the protein-coding genes were predicted by combining ab initio and homology-based approaches. The training set combined self-training GeneMark-ES v2.3f predictions and homology protein alignment using Exonerate. This set was then used to train AUGUSTUS v.27 and SNAP and finally combined by EVIDENCEModeler (EVM) to yield 9,318 consensus gene models (7), 8,651 (93%) of which had orthologs in *T. reesei*.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [LFM10000000](https://www.ncbi.nlm.nih.gov/nuccore/LFM10000000). The version described in this paper is version LFM101000000.

ACKNOWLEDGMENTS

Work in the Vienna Laboratory was supported by grants from the Austrian Science Fund (I-1249 and P25613-B20 to C.P.K. and I.S.D., respectively).

REFERENCES

- Singh A, Taylor LE, II, Vander Wall TA, Linger J, Himmel ME, Podkaminer K, Adney WS, Decker SR. 2015. Heterologous protein expression in *Hypocrea jecorina*: a historical perspective and new developments. *Biotechnol Adv* 33:142–154. <http://dx.doi.org/10.1016/j.biotechadv.2014.11.009>.
- Druzhinina IS, Komoń-Zelazowska M, Atanasova L, Seidl V, Kubicek CP. 2010. Evolution and ecophysiology of the industrial producer *Hypocrea jecorina* (anamorph *Trichoderma reesei*) and a new sympatric agamospecies related to it. *PLoS One* 5:e9191. <http://dx.doi.org/10.1371/journal.pone.0009191>.
- Kubicek CP, Bissett J, Druzhinina I, Kullnig-Gradinger C, Szakacs G. 2003. Genetic and metabolic diversity of *Trichoderma*: a case study on South-East Asian isolates. *Fungal Genet Biol* 38:310–317. [http://dx.doi.org/10.1016/S1087-1845\(02\)00583-2](http://dx.doi.org/10.1016/S1087-1845(02)00583-2).
- Druzhinina IS, Koptchinskiy AG, Komoń M, Bissett J, Szakacs G, Kubicek CP. 2005. An oligonucleotide barcode for species identification in *Trichoderma* and *Hypocrea*. *Fungal Genet Biol* 42:813–828. <http://dx.doi.org/10.1016/j.fgb.2005.06.007>.
- Atanasova L, Jaklitsch WM, Komoń-Zelazowska M, Kubicek CP, Druzhinina IS. 2010. Clonal species *Trichoderma parareesei* sp. nov. likely resembles the ancestor of the cellulase producer *Hypocrea jecorina*/*T. reesei*. *Appl Environ Microbiol* 76:7259–7267. <http://dx.doi.org/10.1128/AEM.01184-10>.
- 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>.
- 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>.