

# Genome Sequence and Annotation of *Acremonium chrysogenum*, Producer of the $\beta$ -Lactam Antibiotic Cephalosporin C

Dominik Terfehr,<sup>a</sup> Tim A. Dahlmann,<sup>a</sup> Thomas Specht,<sup>b</sup> Ivo Zadra,<sup>b</sup> Hubert Kürnsteiner,<sup>b</sup> Ulrich Kück<sup>a</sup>

Christian Doppler Laboratory for "Fungal Biotechnology," Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-Universität Bochum, Bochum, Germany<sup>a</sup>; Sandoz GmbH, Kundl, Austria<sup>b</sup>

**The filamentous fungus *Acremonium chrysogenum* is the industrial producer of the  $\beta$ -lactam antibiotic cephalosporin C. Here, we present the genome sequence of strain ATCC 11550, which contains genes for 8,901 proteins, 127 tRNAs, and 22 rRNAs. Genome annotation led to the prediction of 42 gene clusters for secondary metabolites.**

Received 19 August 2014 Accepted 20 August 2014 Published 25 September 2014

**Citation** Terfehr D, Dahlmann TA, Specht T, Zadra I, Kürnsteiner H, Kück U. 2014. Genome sequence and annotation of *Acremonium chrysogenum*, producer of the  $\beta$ -lactam antibiotic cephalosporin C. *Genome Announc.* 2(5):e00948-14. doi:10.1128/genomeA.00948-14.

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

Address correspondence to Ulrich Kück, ulrich.kueck@rub.de.

*Acremonium chrysogenum* is an imperfect ascomycete which was first described in 1954 by Guiseppe Brotzu who isolated this fungus from Sardinian coastal seawater. Later the antibiotic potential of its extracts was analyzed and described (1, 2). In contrast to the commonly used  $\beta$ -lactam antibiotic penicillin, cephalosporin C is active against both Gram-negative and Gram-positive bacteria. The ongoing demand for antibiotic chemotherapy has generated a world market where  $\beta$ -lactam antibiotics are the most commonly used drugs with an estimated annual turnover of US\$ 22 billion and an estimated market share of 50% by cephalosporin C derivatives (3).

Previous studies on the karyotype from the ATCC 11550 strain reveal an estimated genome size of 32.7 Mb on 8 chromosomes (4) and subsequent progress in developing molecular tools allowed the genetic manipulation of this industrial fungus (5).

In this study, we sequenced the genome of the *A. chrysogenum* ATCC 11550 strain in a whole-genome shotgun sequencing approach which delivered ~23.4 million paired reads with a median insert size of 266 bp and ~12.6 million paired reads with a median insert size of 7,704 bp. The acquired sequence reads were assembled into 2,799 contigs using Velvet v1.2.10 with a k-mer length of 63 nt (6). These were subsequently assembled into 541 scaffolds using SSPACE v2.0 (7). The resulting genome sequence has an estimated size of 28.6 Mb ( $N_{50}$  166,906 bp,  $N_{Max}$  878,651 bp, mean coverage 137.1) with 1,189 gaps and a G+C content of 54.6%. Within these sequences scaffold 543 contains the complete mitochondrial genome.

The protein coding genes were predicted using the MAKER annotation pipeline v2.31.5 in collaboration with Augustus v2.5.5 and SNAP resulting in 8,901 protein coding sequences with minimal protein lengths of 25 amino acids (8–10). A total of 5,433 (61%) from these putative protein coding genes were successfully annotated via BLASTp similarity searches against the Swiss-Prot database. Furthermore, 127 tRNA and 22 rRNA genes were predicted using tRNAscan-SE v1.3.1 and RNAmmer v1.2 (11, 12).

To evaluate the potential to produce secondary metabolites the genome sequence was also used for secondary metabolite cluster

prediction with antiSMASH v2.0 (13). Overall 42 secondary metabolite clusters were predicted which were subcategorized as 14 type 1 polyketide synthetase clusters, 10 terpene synthase clusters, 7 nonribosomal peptide synthetase clusters, 8 hybrid clusters, and 3 not further specified secondary metabolite clusters.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. [JPKY000000000](https://www.ncbi.nlm.nih.gov/nuccore/JPKY000000000). The version described in this paper is the first version, JPKY01000000.

## ACKNOWLEDGMENTS

We thank Minou Nowrousian (Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-University Bochum, Germany) for her advice in bioinformatics.

We also thank the Christian Doppler Society (Vienna, Austria) for financial support. D.T. is grateful for a Ph.D. scholarship from the Studienstiftung des Deutschen Volkes (German National Academic Foundation). D.T. and T.A.D. were supported as members of the RUB research school.

## REFERENCES

- Abraham EP, Newton GG, Olson BH, Schuurmans DM, Schenck JR, Hargie MP, Fisher MW, Fusari SA. 1955. Identity of cephalosporin N and synnematin B. *Nature* 176: 551. <http://dx.doi.org/10.1038/176551a0>.
- Newton GG, Abraham EP. 1955. Cephalosporin C, a new antibiotic containing sulphur and D- $\alpha$ -aminoadipic acid. *Nature* 175:548. <http://dx.doi.org/10.1038/175548a0>.
- Demain AL. 2009. Antibiotics: natural products essential to human health. *Med. Res. Rev.* 29:821–842. <http://dx.doi.org/10.1002/med.20154>.
- Walz M, Kück U. 1991. Polymorphic karyotypes in related *Acremonium* strains. *Curr. Genet.* 19:73–76. <http://dx.doi.org/10.1007/BF00326285>.
- Bloemendal S, Löper D, Terfehr D, Kopke K, Kluge J, Teichert I, Kück U. 2014. Tools for advanced and targeted genetic manipulation of the  $\beta$ -lactam antibiotic producer *Acremonium chrysogenum*. *J. Biotechnol.* 169:51–62. <http://dx.doi.org/10.1016/j.jbiotec.2013.10.036>.
- 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>.
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
- Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-

- database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491. <http://dx.doi.org/10.1186/1471-2105-12-491>.
9. Stanke M, Morgenstern B. 2005. Augustus: a Web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 33:W465–W467. <http://dx.doi.org/10.1093/nar/gki458>.
  10. Korf I. 2004. Gene finding in novel genomes. *BMC Bioinformatics* 5:59. <http://dx.doi.org/10.1186/1471-2105-5-59>.
  11. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108. <http://dx.doi.org/10.1093/nar/gkm160>.
  12. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS Web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33:W686–W689. <http://dx.doi.org/10.1093/nar/gki366>.
  13. Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T. 2013. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* 41:W204–W212. <http://dx.doi.org/10.1093/nar/gkt449>.