



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Draft genome of the fungus-growing termite pathogenic fungus *Ophiocordyceps bispora* (Ophiocordycipitaceae, Hypocreales, Ascomycota)



Benjamin H. Conlon^{a,b,*}, Jannette Mitchell^c,
Z. Wilhelm de Beer^d, Christian Carøe^e, M. Thomas P. Gilbert^e,
Jørgen Eilenberg^f, Michael Poulsen^a, Henrik H. de Fine Licht^f

^a Centre for Social Evolution, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Denmark

^b Molecular Ecology, Institute of Biology, Martin-Luther-University Halle-Wittenberg, Hoher Weg 4, 06099 Halle an der Saale, Germany.

^c ARC-PPRI Rietondale, 600 Soutpansberg Road, Rietondale, Pretoria, Gauteng, South Africa

^d Department of Microbiology and Plant Pathology, Forestry and Agriculture Biotechnology Institute, University of Pretoria, Pretoria, Gauteng 0001, South Africa

^e Centre for GeoGenetics, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark

^f Section for Organismal Biology, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Denmark

ARTICLE INFO

Article history:

Received 15 December 2016

Received in revised form

21 January 2017

Accepted 28 February 2017

Available online 8 March 2017

ABSTRACT

This article documents the public availability of genome sequence data and assembled contigs representing the partial draft genome of *Ophiocordyceps bispora*. As one of the few known pathogens of fungus-farming termites, a draft genome of *O. bispora* represents the opportunity to further the understanding of disease and resistance in these complex termite societies. With the ongoing attempts to resolve the taxonomy of the Hypocrales family, more genetic data will also help to shed light on the phylogenetic relationship between sexual and asexual life stages. Next generation sequence data is available from the European Nucleotide Archive (ENA) under accession PRJEB13655; run numbers: ERR1368522, ERR1368523, and ERR1368524. Genome assembly available from ENA under accession

* Corresponding author.

E-mail addresses: benjamin.conlon@zoologie.uni-halle.de (B.H. Conlon), jannette.mitchell@bufo.co.za (J. Mitchell), wilhelm.debeer@fabi.up.ac.za (Z.W. de Beer), christiancaroe@snm.ku.dk (C. Carøe), tgilbert@snm.ku.dk (M.T.P. Gilbert), jei@plen.ku.dk (J. Eilenberg), MPoulsen@bio.ku.dk (M. Poulsen), hdefineicht@plen.ku.dk (H.H. de Fine Licht).

¹ Current address: Molecular Ecology, Institute of Biology, Martin-Luther-University Halle-Wittenberg, Hoher Weg 4, 06099 Halle an der Saale, Germany.

<http://dx.doi.org/10.1016/j.dib.2017.02.051>

2352-3409/© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

numbers: FKNF01000001–FKNF01000302. Gene prediction available as protein fasta, nucleotide fasta and GFF file from Mendeley Data with accession <http://dx.doi.org/10.17632/r99fd6g3s4.2> (<http://dx.doi.org/10.17632/r99fd6g3s4.2>).

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications Table

Subject area	Biology
More specific subject area	Mycology, Genomics
Type of data	Genomic sequence, gene prediction and phylogenetic placement of <i>Ophiocordyceps bispora</i>
How data was acquired	Shotgun whole-genome DNA sequencing using the Illumina HiSeq platform at The Danish National High-Throughput DNA Sequencing Centre
Data format	Raw sequencing reads, Draft genome assembly, gene prediction and phylogenetic analysis
Experimental factors	DNA was extracted from <i>Ophiocordyceps bispora</i> fruiting bodies protruding from the thorax of two <i>Macrotermes</i> spp. alates.
Experimental features	The Geneious ver. 8.0.5. read-mapping and assembly tool, which also takes read quality into account when mapping, were used to assemble the draft <i>O. bispora</i> genome. The genome of the closely related <i>Hirsutella thompsonii</i> MTCC 6686 [1] was used as reference genome. Sequences used for phylogeny construction were identified using blastn searches.
Data source location	Samples were collected in 1993 from two alates of <i>Macrotermes</i> sp. with <i>O. bispora</i> fruiting bodies growing out from the thorax. The specimens were found in a small cavity excavated underneath a rock in Isiolo and Kajiado, Kenya.
Data accessibility	European Nucleotide Archive (BioProject: PRJEB13655; Runs: ERR1368522, ERR1368523, and ERR1368524; Contigs: FKNF01000001–FKNF01000302) and Mendeley Data (http://dx.doi.org/10.17632/r99fd6g3s4.2 , http://dx.doi.org/10.17632/r99fd6g3s4.2).

Value of the data

- *Ophiocordyceps bispora* is the first pathogen of fungus-farming termites (Macrotermitinae) to have its genome sequenced.
- *O. bispora* represents one of the few known parasites of fungus-farming social insects.
- The *O. bispora* draft genome assembly will be of value for future comparative and phylogenetic analyses.
- Further study of *O. bispora*'s relationship with *Hirsutella thompsonii* could potentially identify a teleomorph–anamorph connection bringing the two species in line with the 2011 recommendations for The International Code of Nomenclature for algae, fungi and plants [2].

1. Data

We present a partial draft genome assembly with gene prediction of the fungus-farming termite pathogenic fungus *Ophiocordyceps bispora* (Ophiocordycipitaceae, Hypocreales). *O. bispora* infects the

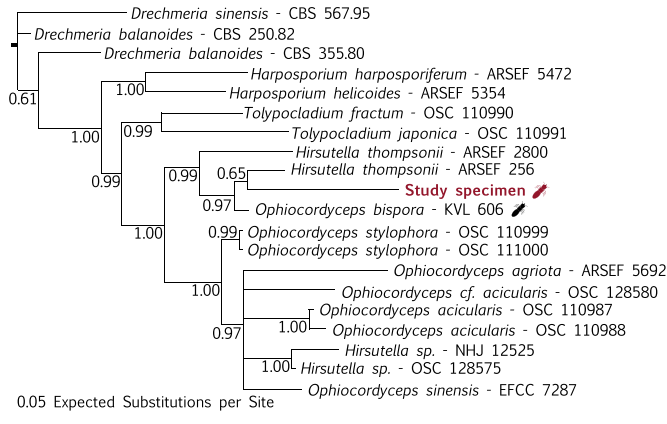


Fig. 1. Phylogenetic placement of our specimen and closely related fungi in a concatenated nrSSU and nrLSU Bayesian phylogeny with posterior probabilities given at nodes. The sequenced specimen is highlighted in red while all other specimens are labelled with their herbarium numbers. A termite indicates that the fungus was reported from termites.

reproductive alate caste of several termite genera, including fungus-growing termites in the genus *Macrotermes* in Africa [3–7]. It is one of the few known pathogens of fungus-farming social insects and its *Macrotermes* host apparently has no other known diseases [8,9]. We also provide a two-gene phylogenetic analysis (Fig. 1) showing that *O. bispora* is closely related to an asexual form of *Ophiocordyceps*: *Hirsutella thompsonii*. The stark difference between sexual and asexual life stages in the Hypocreales family imply the different life stages of many species are classified twice [10]. Since 2011, the International Code of Nomenclature for algae, fungi and plants has been changed so that one fungus can have only one name [2]. Efforts to reconcile the taxonomy of the Hypocreales with these changes are ongoing [7,11,12], and the draft genome of *O. bispora* presented here may prove valuable in this effort.

2. Experimental design, materials and methods

2.1. Library

- Strategy: Shotgun whole-genome DNA sequencing.
- Taxon: *Ophiocordyceps bispora*.
- Sample details: Two alates of *Macrotermes* sp. with *O. bispora* fruiting bodies growing out from the thorax were collected from a small cavity excavated underneath a rock.
- Tissue: Fruiting body structures growing out from the dead, infected termite hosts.
- Location: Isiolo and Kajiado, Kenya.
- Sample handling: The sample was collected in 1993 by J. Eilenberg/G. Ochiel/H. Evans and preserved in ethanol at the Department of Plant and Environmental Sciences, University of Copenhagen.
- Selection: None.
- Layout: Single-end 100 bp reads.

2.2. Library construction protocol

Total fungal DNA was extracted from the fruiting bodies protruding from the termites by crushing and dissolving the sample in extraction buffer (10 mM Tris, 10 mM sodium chloride (NaCl), 5 mM calcium chloride (CaCl), 2.5 mM ethylenediaminetetraacetic acid (EDTA), 1% sodiumdodecyl sulfate (SDS), 10%

proteinase K and 6.17 mg/mL dithiothreitol (DTT)) overnight at 56 °C, followed by organic extraction of the DNA using one volume of chloroform. The DNA-containing supernatant was purified by mixing with 10 × volumes of modified PB buffer [13] and using a MinElute spin column (Qiagen) as demonstrated elsewhere [14]. The DNA was eluted in 60 µL EB buffer after a ten-minute incubation of the spin column with buffer at 37 °C and subsequently fragmented using a Diagenode Bioruptor with a program of 30 s on, 90 s off for 6 cycles. Illumina compatible shotgun sequencing library was generated using reagents from New England Biolabs kit for the 454, E#6070 L as described elsewhere [15]. The finished library was indexed and amplified for sequencing using 10 µL template and a mastermix consisting of 1 × AmpliTaq Gold buffer (Invitrogen), 2.5 mM MgCl₂ (Invitrogen), 0.8 µg/µL bovine serum albumin (BSA), 0.25 mM dNTP (Thermo Fischer Scientific, 25 mM stock), 0.2 µM forward and reverse indexed primer (10 µM stock), 0.2 U/µL AmpliTaq Gold enzyme (Invitrogen) and molecular grade water to a final reaction volume of 100 µL. The library was amplified in an Applied Biosystems 2720 Thermal Cycler, using the following conditions: 95 °C for 1 min, followed by a number of cycles of 95 °C for 30 sec, 60 °C for 30 s and 72 °C for 1 min for 5 cycles and finished by 7 min at 72 °C. Finally, the library was purified using a QiaQuick spin column (Qiagen), following manufacturer's instructions. The library was pooled with other samples and sequenced on 10% of a lane on an Illumina HiSeq 2500 instrument at The Danish National High-Throughput DNA Sequencing Centre, Copenhagen, Denmark.

2.3. Processing

Pipeline: DNA of the insect pathogenic fungus *O. bispora* was shotgun sequenced using Illumina HiSeq. Raw sequencing reads are available as three fastq files with accession nos. ERR1368522, ERR1368523, and ERR1368524.

- Run data file type: Fastq
- File Names (Runs: ERR1368522, ERR1368523, and ERR1368524, respectively):
 - TOG_QEHU_Term_fung_03_15_ACATAC_L004_R1_001.fastq.gz
 - TOG_QEHU_Term_fung_03_15_ACATAC_L004_R1_002.fastq.gz
 - TOG_QEHU_Term_fung_03_15_ACATAC_L004_R1_003.fastq.gz

Table 1

The fungal strains included in the phylogeny including known host taxa and GenBank accession numbers for nrLSU and nrSSU sequences.

Genus	Species	Strain	Host Taxa (where known)	nrLSU	nrSSU
<i>Drechmeria</i>	<i>balanoides</i>	CBS 250.82	Nematode	AF339539	AF339588
<i>Drechmeria</i>	<i>balanoides</i>	CBS 335.80	Nematode	AF339540	AF339589
<i>Drechmeria</i>	<i>sinensis</i>	CBS 567.95	Nematode	AF339545	AF339594
<i>Harposporium</i>	<i>helicoides</i>	ARSEF 5354	Nematode	AF339527	AF339577
<i>Harposporium</i>	<i>harposporiferum</i>	ARSEF 5472	Arthropod	AF339519	AF339569
<i>Hirsutella</i>	<i>sp.</i>	NHJ 12525	Hemiptera	EF469078	EF469125
<i>Hirsutella</i>	<i>sp.</i>	OSC 128575	Hemiptera	EF469079	EF469126
<i>Hirsutella</i>	<i>thompsonii</i>	ARSEF 256	Acari	KM652135	KM652090
<i>Hirsutella</i>	<i>thompsonii</i>	ARSEF 2800	Acari	KM652142	KM652095
<i>Ophiocordyceps</i>	<i>acicularis</i>	OSC 110987	Coleopteran larvae	EF468805	EF468950
<i>Ophiocordyceps</i>	<i>acicularis</i>	OSC 110988	Coleopteran larvae	EF468804	EF468951
<i>Ophiocordyceps</i>	<i>agriota</i>	ARSEF 5692	Coleoptera	DQ518754	DQ522540
<i>Ophiocordyceps</i>	<i>bispora</i>	KVL 606	Termite (Isoptera)	AF009654	AH006986
<i>Ophiocordyceps</i>	<i>cf.acicularis</i>	OSC 128580	Coleoptera	DQ518757	DQ588543
<i>Ophiocordyceps</i>	<i>sinensis</i>	EFCC 7287	Lepidopteran pupae	EF468827	EF468971
<i>Ophiocordyceps</i>	<i>stylophora</i>	OSC 110999	Coleopteran larvae	EF468837	EF468982
<i>Ophiocordyceps</i>	<i>stylophora</i>	OSC 111000	Coleopteran larvae	DQ518766	DQ522522
<i>Tolyposcladium</i>	<i>fractum</i>	OSC 110990	Euteriomycete	DQ518759	DQ522545
<i>Tolyposcladium</i>	<i>japonica</i>	OSC 110991	Euteriomycete	DQ518761	DQ588547

The Geneious ver. 8.0.5. read-mapping and assembly tool, which also takes read quality into account when mapping, were used to assemble the draft *O. bispora* genome. The genome of the closely related *Hirsutella thompsonii* MTCC 6686 [1] was used as reference genome. The final assembled *O. bispora* draft genome consisted of 302 contigs and was 6,359,382 bp long. The maximum contig length, with a mean coverage of 3.1, was 158,451 bp with an N50 of 41,819 bp and a GC of 58.5%. Assembled contigs are available with accession nos. FKNF01000001–FKNF01000302. Protein coding sequences were predicted on both strands in the *O. bispora* draft genome using Augustus ver. 3.2.1 [16] using default parameters and found 3324 Open Reading Frames (ORFs).

The common fungal barcoding genes nuclear-ribosomal Large Sub-Unit (nrLSU) and nuclear ribosomal Small Sub-Unit (nrSSU) were located in the draft *O. bispora* genome assembly using blastn searches with Geneious 4.8.5. The identified *O. bispora* nrLSU and nrSSU sequences were combined with other Hypocreales sequences from GenBank (Table 1), and a concatenated alignment of nrSSU and nrLSU was produced using MUSCLE [17]. A Bayesian phylogenetic analysis was performed using Topali v2 (Runs: 2; Generations: 1,500,000; Burn in: 50%) [18].

Acknowledgements

This work was supported by Docent, Dr. Scient. Lauritz Olsons Rejsefond to BHC, the Carlsberg Foundation to HHdFL (2012_01_0599; 2013_01_0737), the DST-NRF Centre of Excellence in Tree Health Biotechnology (CTHB) to WdB, the Lundbeck Foundation to MTPG (R52-5062) and Villum Kann Rasmussen Young Investigator grants to MP (VKR10101) and HHdFL (VKR10122).

Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2017.02.051>.

References

- [1] Y. Agrawal, I. Khatri, S. Subramanian, et al., Genome sequence, comparative analysis, and evolutionary insights into chitinases of entomopathogenic fungus *Hirsutella thompsonii*, *Genome Biol. Evol.* 7 (3) (2015) 916–930.
- [2] D.L. Hawksworth, P.W. Crous, S.A. Redhead, et al., The Amsterdam declaration on fungal nomenclature, *IMA Fungus* 2 (1) (2011) 105–112.
- [3] C.B. Stifler, A new genus of Hypocreales, *Mycologia* 33 (1) (1941) 82–86.
- [4] M. Blackwell, R.L. Gilbertson, New information on *Cordycepioideus bispora* and *Cordycepioideus octosporus*, *Mycologia* 76 (4) (1984) 763–765.
- [5] G.S. Ochiel, H.C. Evans, J. Eilenberg, *Cordycepioideus*, a pathogen of termites in Kenya, *Mycologist* 11 (1) (1997) 7–9.
- [6] G.H. Sung, N.L. Hywel-Jones, J.M. Sung, et al., Phylogenetic classification of *Cordyceps* and the clavicipitaceous fungi, *Studies Mycol.* 57 (2007) 5–59.
- [7] C.A. Quandt, R.M. Kepler, W. Gams, et al., Phylogenetic-based nomenclatural proposals for Ophiocordycipitaceae (Hypocreales) with new combinations in *Tolyptocladium*, *IMA Fungus* 5 (1) (2014) 121–134.
- [8] D.P. Hughes, H.C. Evans, N. Hywel-Jones, et al., Novel fungal disease in complex leaf-cutting ant societies, *Ecol. Entomol.* 34 (2009) 214–220.
- [9] J.J. Boomsma, A.B. Jensen, N.V. Meyling, et al., Evolutionary interaction networks of insect pathogenic fungi, *Annu. Rev. Entomol.* 59 (2014) 467–485.
- [10] A. Chioza, S. Ohga, A review on fungal isolates reported as anamorphs of *Ophiocordyceps sinensis*, *J. Mycol.* (2014) (Article ID 913917).
- [11] S. Ban, Y. Azuma, H. Sato, et al., *Isaria takamizusanensis* is the anamorph of *Cordyceps ryogamimontana*, warranting a new combination, *Purpureocillium takamizusanense* comb. nov., *Int. J. Syst. Evol. Microbiol.* 65 (8) (2015) 2459–2465.
- [12] J.W. Spatafora, C.A. Quandt, R.M. Kepler, et al., New 1F1N species combinations in Ophiocordycipitaceae (Hypocreales), *IMA Fungus* 6 (2015) 357–362.
- [13] M.E. Allentoft, M. Sikora, K.G. Sjögren, et al., Population genomics of Bronze Age Eurasia, *Nature* 522 (7555) (2015) 167–172.

- [14] J. Dabney, M. Knapp, I. Glocke, et al., Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments, *Proc. Natl. Acad. Sci.* 110 (39) (2013) 15758–15763.
- [15] N. Wales, C. Carøe, M. Sandoval-Velasco, et al., New insights on single-stranded versus double-stranded DNA library preparation for ancient DNA, *BioTechniques* 59 (2015) 368–371.
- [16] M. Stanke, S. Waack, Gene prediction with a hidden Markov model and a new intron model, *Bioinformatics* 19 (Suppl. 2) (2003) (ii215–ii225).
- [17] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.* 32 (5) (2004) 1792–1797.
- [18] I. Milne, F. Wright, G. Rowe, et al., TOPALI: software for automatic identification of recombinant sequences within DNA multiple alignments, *Bioinformatics* 20 (11) (2004) 1806–1807.