

## Complete mitochondrial genome of *Pleurocordyceps sinensis* (Hypocreales, Ascomycota), a species with uncertain family-level taxonomic assignment

Yi Li<sup>1,2†\*</sup>, Jie Lin<sup>3†</sup>, Yong-Hui Wang<sup>4</sup>, Ke Wang<sup>2</sup>, Ren-Lei Wang<sup>3</sup>, Xin-Chang Zhao<sup>1</sup>, Yi-Jian Yao<sup>2\*</sup>

<sup>1</sup>School of Food Science and Engineering, Yangzhou University, Yangzhou, China; <sup>2</sup>State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China; <sup>3</sup>Jiangsu Key Laboratory for Biofunctional Molecules, College of Life Science and Chemistry, Jiangsu Second Normal University, Nanjing, China; <sup>4</sup>Institute of Agriculture Resource and Environment, Shandong Academy of Agriculture Sciences, Jinan, China

†Yi Li and Jie Lin contributed equally to this manuscript.

\*Corresponding Authors: Yi Li, School of Food Science and Engineering, Yangzhou University, Yangzhou, China. Email: [liy062@yzu.edu.cn](mailto:liy062@yzu.edu.cn); Yi-Jian Yao, State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. Email: [yaoyj@im.ac.cn](mailto:yaoyj@im.ac.cn)

Received: 17 June 2022; Accepted: 12 August 2022; Published: 23 November 2022

© 2022 Codon Publications

OPEN ACCESS 

RESEARCH ARTICLE

### Abstract

The complete mitochondrial (mt) genome of the ex-type strain of *Pleurocordyceps sinensis*, a fungus originally isolated from *Ophiocordyceps sinensis*, was sequenced, and assembled as a single circular DNA of 31,841 bp. The mt genome encoded 15 conserved proteins (*rps3*, *cox1*, *cox2*, *cox3*, *cob*, *atp6*, *atp8*, *atp9*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6*), 2 rRNA (*rnl* and *rns*), and 25 tRNA, as well as 10 additional non-conserved open reading frames (ncORFs). Comparative analyses showed that mt genomes within the order Hypocreales encoded the same number and synteny of conserved protein coding genes despite an obvious size variation among this group of fungi. Phylogenetic analyses using 14 conserved protein sequences revealed that this fungus may not belong to the current designated family *Ophiocordycipitaceae* but is more closely related to the species of *Clavicipitaceae*. The mt genome presented herein would give valuable information on reconstructing the evolutionary history of clavicipitaceous fungi and also aid in resolving the family-level taxonomic assignment of *Polycephalomyces* s. l. species.

**Keywords:** *Cordyceps*; hyperparasitism; *Polycephalomyces sinensis*; taxonomy

### Introduction

*Pleurocordyceps sinensis* (Q.T. Chen *et al.*) Y.J. Yao *et al.* was first isolated from the sclerotium of *Ophiocordyceps sinensis* collected from Kangding, Sichuan, China in June 1980, and was described as a new species named *Paecilomyces sinensis* Q.T. Chen *et al.* (Chen *et al.*, 1984). The species gained broad scientific attention since 1980s in China. Large numbers of pharmacological studies have been carried out using the only authentic strain termed CN80-2. The species was reported to have various pharmacological activities such as anti-implantation (Lin *et al.*, 1988), anti-inflammatory (Li *et al.*, 1983),

anti-oxidant (Liu *et al.*, 1987, 1989, 1991), anti-tumor (Huang *et al.*, 1988; Wu *et al.*, 1986), fertility regulation (Li and Lin, 1991), immune modulation (Ge *et al.*, 1989; Lin *et al.*, 1987; Zhang *et al.*, 1998; Zheng *et al.*, 1983), and treating conditions of coronary arteriosclerotic heart disease (You *et al.*, 1986) and immunological liver injury (Cheng *et al.*, 2005; Zeng *et al.*, 2000). Because of the similarities in the pharmacological effects and chemical components, *P. sinensis* has once been recognized as the possible anamorph of *O. sinensis*. This viewpoint was widely accepted and cited when summarizing the anamorph of *O. sinensis* (Fang, 1991; Liang, 1991; Liu, 1990). The idea of using *Paecilomyces sinensis* as a substitute of

*O. sinensis* was thus proposed (e.g., Cheng *et al.*, 2005; Li *et al.*, 1983; Zeng *et al.*, 2000). However, several independent researches based on molecular evidences rejected the anamorph-teleomorph relationship between *P. sinensis* and *O. sinensis* (Chen *et al.*, 2001; Jiang and Yao, 2002, 2003; Li *et al.*, 2000; Zhao *et al.*, 1999). Wang *et al.* (2012) placed the species in *Polycephalomyces* based on morphological and molecular analyses and found *Polycephalomyces* species formed a new clade of clavicipitaceous fungi and stated that this new clade is distinct from the known families of Hypocreales.

The genus *Polycephalomyces* was identified by Kobayasi (1941) with *P. formosus* as the type. Only three species, that is, *P. paludosus*, *P. cylindrosporus*, and *P. tomentosus*, were described in the last century (Mains, 1948; Samson *et al.*, 1981; Seifert, 1986). Until recently, after the recombination of *Paecilomyces sinensis* (Wang *et al.*, 2012) and several species of *Cordyceps* s. l. into the genus (Kepler *et al.*, 2013), more and more new species have been discovered and described, especially from China and Southeast Asia (Crous *et al.*, 2017; Wang *et al.*, 2015a, 2015b; Xiao *et al.*, 2018; Yang *et al.*, 2020). A total of 24 species names are currently recorded by the Index Fungorum (4 April 2022, <http://www.indexfungorum.org/Names/Names.asp>), among which 4 were segregated to *Perennicordyceps* (Matočec *et al.*, 2014). Wang *et al.* (2021) recently proposed a new genus *Pleurocordyceps* for one of the subclades within the “*Polycephalomyces* clade” (*Polycephalomyces* sensu lato). Ten species were included in the new genus including *P. sinensis*. In multigene phylogenetic analyses, species of *Polycephalomyces* s. l. usually formed a distinct clade sister to *Ophiocordycipitaceae*, although this sister relationship did not receive much statistical confidence (Kepler *et al.*, 2013; Wang *et al.*, 2021). In other words, the family-level taxonomic position of *Polycephalomyces* s. l. was not fully resolved; species in this group were tentatively placed in *Ophiocordycipitaceae* in most researches (e.g., Kepler *et al.*, 2013; Xiao *et al.*, 2018), which was accepted by the Index Fungorum. *Polycephalomyces* s. l. may represent a new family that is different from the three existing families of clavicipitoid fungi (Wang *et al.*, 2021), that is, *Cordycipitaceae*, *Clavicipitaceae*, and *Ophiocordycipitaceae*. However, more evidence from morphology and molecular phylogenetics is required to support this hypothesis.

Mitochondria play various essential roles in eukaryotic cells, including respiratory metabolism, energy production, calcium homeostasis, and are also involved in cell death and aging (Basse, 2010). Mitochondrial (mt) genomes usually have a rapid rate of evolution compared with nuclear genomes, and thus are considered as powerful tools in evolutionary biology (Berbee and Taylor, 2001; Chris *et al.*, 1994). Previous studies revealed that the gene contents

and synteny of mt genomes of hypocrealean species were largely conserved, but in the meantime, the genome sizes expanded greatly in certain species such as *O. sinensis* (Li *et al.*, 2015). Complete mt genomes have been reported for a number of species of the three clavicipitaceous fungal families, that is, *Cordycipitaceae* (Fan *et al.*, 2019; Kouvelis *et al.*, 2004; Sung, 2015; Zhang *et al.*, 2021), *Ophiocordycipitaceae* (Abuduaini *et al.*, 2021; Li *et al.*, 2015; Zhang *et al.*, 2016; Zhang and Zhang, 2020), and *Clavicipitaceae* (Sun *et al.*, 2021; Winter *et al.*, 2018), while they have not been reported for *Polycephalomyces* s. l. species so far.

In this study, the complete mt genome of the type strain (CN 80-2) of the species *Pleurocordyceps sinensis* was sequenced, described, and compared with other hypocrealean species. Phylogenetic analyses using 14 conserved protein sequences were also performed to study the phylogenetic relationship between this species and other clavicipitaceous fungal groups.

## Materials and Methods

### Fungal isolation and cultivation

The ex-type strain (CN80-2) of *Pleurocordyceps sinensis* used in this study was isolated from a sclerotium of *O. sinensis* collected from Kangding, Sichuan, China in June 1980 (Chen *et al.*, 1984). Stock strain was maintained at 4°C on potato dextrose agar (PDA) slants. Seed cultures were grown in 250-mL Erlenmeyer flasks, containing 50 mL liquid potato-dextrose medium, by shaking at 100 rpm at 25°C for 10 days. Mycelia were harvested and washed with distilled water using vacuum filtration to remove extracellular polysaccharides, frozen with liquid nitrogen, and vacuum freeze-dried using a freeze dryer. Dried mycelia were then sent to the genome sequencing company.

### DNA extraction and genome sequencing

Genomic DNA was extracted by the sequencing company using TIANamp Yeast DNA Kit (TIANGEN Biotech Co., Ltd., Beijing, China) according to the manufacturer's instruction. The amount and quality of total DNA were visualized by 1% agarose gel electrophoresis and quantified with a Qubit2.0<sup>®</sup> Fluorometer (Life Technologies, New York, USA). A 20 K library was prepared from sheared genomic DNA (containing both mt and nuclear sequences) using a 20-Kb template library preparation workflow. Twelve single molecule real time (SMRT) sequencing cells were sequenced on PacBio RS II sequencing platform (Pacific Biosciences, Menlo Park, CA) with P6 polymerase and C4 sequencing chemistry at Tianjin Biochip Corporation (Tianjin, China).

## Mitochondrial genome assembly and annotation

Mt genome of the strain CN80-2 was assembled and annotated following a procedure described in Li *et al.* (2015). The adapter sequences, reads with length <50 bp, or average quality <0.75 (defined as low quality) were filtered before assembling. The mt sequences were extracted from filtered reads matching each read against the fungal mt genome database (<https://www.ncbi.nlm.nih.gov/genome/browse#!/organelles/>), pre-assembled and corrected using BLASR (Chaisson and Tesler, 2012). Corrected reads were retained and then re-assembled with the Celera Assembler program (Myers *et al.*, 2000). The assembly was further refined with Quiver (Chin *et al.*, 2013). A circular double-stranded DNA was finally obtained and proceeded to an online annotation tool MFannot using the Mold, Protozoan, and Coelenterate Mitochondrial Code (Beck and Lang, 2010). The annotated mt genome was submitted to GenBank under the accession number OK017430. The annotated mt genetic map was generated by Circos software (Krzywinski *et al.*, 2009) and modified with Adobe Illustrator® CS5 (Version 15.0.0, Adobe®, San Jose, CA).

## Phylogenetic analyses and comparative genomics

All the 82 complete mt genomes available from GenBank (accessed on 28 March, 2021) within the order Hypocreales were downloaded and used for phylogenetic and/or comparative genomic analysis. Among which *Ophiocordyceps camponoti-floridani* EC05 (CM022976) was used only for genome comparison but not included in phylogenetic analyses, as it seems to be incorrectly assembled; an unclassified mt genome (NC\_049089) was also excluded due to the large numbers of possible sequencing errors or possible assembly mistakes. A number of mt genomes were found to be incorrectly or incompletely annotated. For example, the *rps3* gene was not predicted in a number of species. Those genomes were re-annotated following the same procedure used in this study with missing genes replenished and wrongly predicted genes manually corrected. A phylogenetic tree was constructed using 14 conserved protein-coding genes (*cox1*, *cox2*, *cox3*, *cob*, *atp6*, *atp8*, *atp9*, *nad1*, *nad2*, *nad3*, *nad4*, *nad5*, *nad4L*, and *nad6*). Protein sequences were aligned with BioEdit version 7.0.9.0 (Hall, 1999) and refined manually. Maximum likelihood (ML) phylogenetic analyses were performed with RAxML v.7.2.659 (Stamatakis, 2006) using the LG substitution matrix and default parameters. Bootstrap values were calculated with 1000 re-sampling iterations using an approximate likelihood ratio test. Three mt genomes from two species of the order *Glomerellales*, that is, *Colletotrichum lindemuthianum* (NC\_023540) and *Verticillium dahliae* (NC\_008248 and CM019738), were used as outgroups (Supplementary Table S1).

The gene contents and synteny of mt genomes within the order Hypocreales and related outgroup species were compared and analyzed.

## Results

### Genome sequencing and assembly

A total of 23,757 reads (171,039,810 bp) were identified as mt among 601,168 reads (5,005,308,071 bp) of the raw sequencing output for the whole genome of *Pleurocordyceps sinensis*. The lengths of the putative mt reads ranged from 276 bp to 45,245 bp with an average length of 7200 bp, reaching a coverage depth of 5371× over the mt genome of the species. The mt reads were passed through the program BLASR and assembled with Celera Assembler program and Quiver, resulting in a circular DNA of 31,841 bp (Figure 1).

### Conserved protein genes and nonconserved open reading frames

The mt genome of *P. sinensis* had a low GC content of 25.46% and encoded 15 protein genes conserved within the order Hypocreales, including seven subunits of the electron transport complex I (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6*), cytochrome b (*cob*), three subunits of complex IV (*cox1*, *cox2* and *cox3*), three F0 subunits of the ATP-synthase complex (*atp6*, *atp8*, and *atp9*), and the *rps3* gene, which encodes 40S ribosomal protein S3 (Table 1, Figure 1). In addition to those genes, 10 non-conserved open reading frames (ncORFs) (7194 bp totally in length) were also predicted, among which two (ncORF3 and ncORF9) were found to encode homing endonucleases (HEs) with motif patterns GIY-YIG and LAGLIDADG, respectively (Table 1).

All conserved protein coding genes and ncORFs were found on the positive strand and oriented clockwise except for ncORF1 and ncORF2, which were on the negative strand and anticlockwise oriented. It was found that the *nad2/nad3* genes were joined and *nad4L/nad5* genes were fused, that is, the initial codon of the *nad3* gene (ATG) followed the terminal codon of the *nad2* gene (TAA), and the terminal codon of *nad4L* (TAA) uses the same nucleotide A with the initial codon (ATG) of *nad5* (Figure 1, Table 1). Other protein-coding genes and ncORFs were separated by either long or short intergenic regions (Figure 1).

The 15 protein-coding genes and 10 predicted ncORFs employed the standard fungal mt start codon ATG, except the *cox1* and ncORF10, which were initiated by

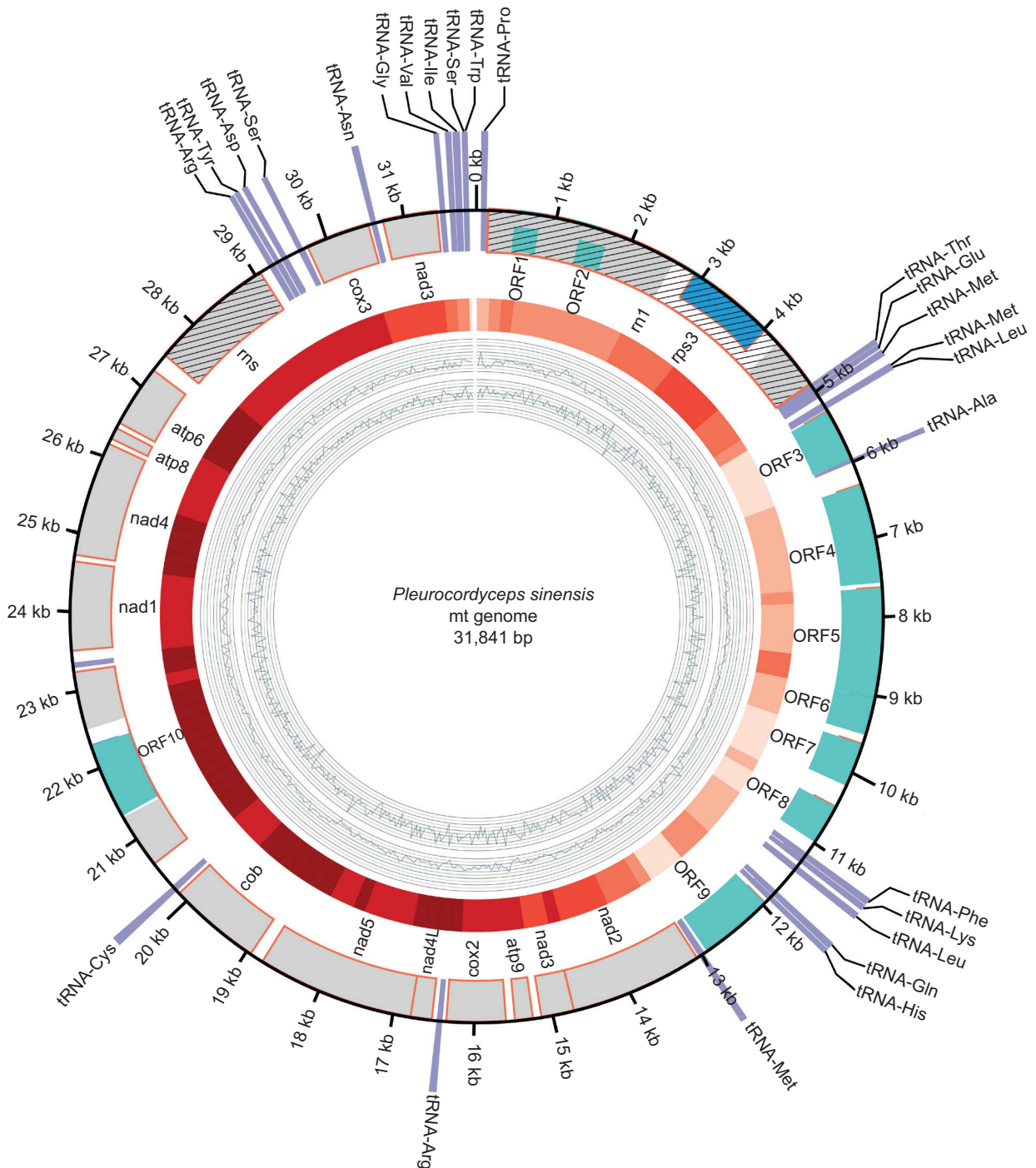


Figure 1. Genetic map of the mitochondrial genome of *Pleurocordyceps sinensis*. Shading blocks with orange frame indicate exons of predicted coding genes, with gene names labeled on the inner side; *rnl* and *rns* genes were marked with slashes; the *rps3* was in blue and nested within an intron of the *rnl* gene; two introns that located in *rnl* and *cox1*, respectively, were shown in blank. Thin flint lines protruding outside of the outer circle indicate tRNAs. Ten predicted ncORFs were shown as light blue blocks. The red circle represented sequence coverage with the highest at 5,200 $\times$  and the lowest at 4,300 $\times$  (average at 4,750 $\times$ ). The outer and inner diagram of curves represented GC content and GC skew, respectively.

**Table 1. Mitochondrial genome annotation of *Pleurocordyceps sinensis*.**

Genes	Strands	Positions	Lengths (bp)	Introns	Start/stop codons	Anticodons
tRNA-Pro [P]	+CW	53–125	73			TGG
<i>rnl</i>	+CW	155–4876	4722	IA (1643), 2583–4225		
ncORF1	–CW	818–519	300		ATG/TAA	
ncORF2	–CW	1581–1252	330		ATG/TAA	
<i>rps3</i>	+CW	2816–4132	1317		ATG/TAA	
tRNA-Thr [T]	+CW	4793–4863	71			TGT
tRNA-Glu [E]	+CW	4869–4941	73			TTC
tRNA-Met [M1]	+CW	4942–5012	71			CAT
tRNA-Met [M2]	+CW	5019–5091	73			CAT
tRNA-Leu [L]	+CW	5177–5258	82			TAA
ncORF3 (GIY-YIG)	+CW	5295–5942	648		ATG/TAA	
tRNA-Ala [A]	+CW	5933–6005	73			CGC
ncORF4	+CW	6307–7539	1233		ATG/TAA	
ncORF5	+CW	7629–8987	1359		ATG/TAA	
ncORF6	+CW	9018–9470	453		ATG/TAA	
ncORF7	+CW	9619–10,149	531		ATG/TAA	
ncORF8	+CW	10,473–10,958	486		ATG/TAG	
tRNA-Phe [F]	+CW	11,156–11,228	73			GAA
tRNA-Lys [K]	+CW	11,230–11,302	73			TTT
tRNA-Leu [L2]	+CW	11,354–11,435	82			TAG
tRNA-Gln [E2]	+CW	11,764–11,836	73			TTG
tRNA-His [H]	+CW	11,841–11,914	74			GTG
ncORF9 (LAGLIDADG)	+CW	11,968–12,888	921		ATG/TAA	
tRNA-Met [M3]	+CW	12,947–13,019	73			CAT
<i>nad2</i>	+CW	13,061–14,737	1677		ATG/TAA	
<i>nad3</i>	+CW	14,738–15,151	414		ATG/TAA	
<i>atp9</i>	+CW	15,260–15,484	225		ATG/TAA	
<i>cox2</i>	+CW	15,598–16,344	747		ATG/TAA	
tRNA-Arg [R1]	+CW	16,391–16,461	71			ACG
<i>nad4L</i>	+CW	16,526–16,795	270		ATG/TAA	
<i>nad5</i>	+CW	16,795–18,792	1998		ATG/TAA	
<i>cob</i>	+CW	18,951–20,120	1170		ATG/TAA	
tRNA-Cys [C]	+CW	20,176–20,247	72			GCA
<i>cox1</i>	+CW	20,597–23,225	2629	IB (1036), 21,336–22,372	ATA/TAA	
ncORF10	+CW	21,335–22,285	933		ATA/TAA	
tRNA-Arg [R2]	+CW	23,276–23,346	71			TCT
<i>nad1</i>	+CW	23,495–24,616	1122		ATG/TAA	
<i>nad4</i>	+CW	24,699–26,156	1458		ATG/TAA	
<i>atp8</i>	+CW	26,228–26,374	147		ATG/TAA	
<i>atp6</i>	+CW	26,450–27,235	786		ATG/TAA	
<i>rns</i>	+CW	27,513–29,036	1524			
tRNA-Tyr [Y]	+CW	29,189–29,272	84			GTA
tRNA-Asp [D]	+CW	29,284–29,357	74			GTC
tRNA-Ser [S1]	+CW	29,370–29,452	83			GCT
tRNA-Asn [N]	+CW	29,619–29,690	72			GTT
<i>cox3</i>	+CW	29,733–30,542	810		ATG/TAA	
tRNA-Gly [G]	+CW	30,578–30,648	71			TCC

(continues)

Table 1. Continued

Genes	Strands	Positions	Lengths (bp)	Introns	Start/stop codons	Anticodons
<i>nad6</i>	+CW	30,732–31,418	687		ATG/TAA	
tRNA-Val [V]	+CW	31,452–31,524	73			TAC
tRNA-Ile [I]	+CW	31,580–31,651	72			GAT
tRNA-Ser [S2]	+CW	31,656–31,742	87			TGA
tRNA-Trp [T]	+CW	31,755–31,826	72			TCA

Note: +, genes encoded on positive strain; –, genes encoded on negative strain; CW, genes were clockwise oriented.

ATA. In addition, 24 of those genes used TAA as the stop codon except the ncORF8, which used TAG (Table 1).

### Noncoding RNAs

In addition to the 15 protein-coding genes, a large and a small ribosomal RNA (*rnl* and *rns*, respectively) and 25 tRNA genes were also identified (Table 1). The tRNA genes ranged in size from 71 to 87 bp and could correspond to 20 amino acids. A majority of amino acids were coded by only one tRNA gene; however, Serine (Ser), Arginine (Arg), Methionine (Met), and Leucine (Leu) had 2, 2, 3, and 2 tRNA genes, respectively (Table 1). All non-coding RNAs (tRNA, rRNA) were found on the positive strand and oriented clockwise.

### Intronic and intergenic regions

Exons of protein-coding genes, rRNA and tRNA genes, had a total length of 20,873 bp accounting for 65.55% of the mt genome. Ten ncORFs (7,194 bp) accounted for 22.59% of the mt genome. Only two introns (group I) were predicted, including one further classified into subgroup IA (1,643 bp) in *rnl* and one classified into subgroup IB in *cox1* (1,036 bp), respectively, making up 8.5% of the entire mt genome. The intergenic sequences had a total length of 1,070 bp covering 3.4% of the genome.

### Gene component and synteny

Although different numbers of ncORFs (hypothetical proteins) would be predicted for hypocrealean fungi, the content and synteny of 15 protein-coding genes remained largely conserved, except in a few cases. For instance, location of the *cox2* gene shifted in three species of *Acremonium chrysogenum*, *A. fuci*, and *Clonostachys rose* comparing to other Hypocreales; an additional copy of *rps3* and *atp9* gene was found in *Beauveria malawiensis* and *Fusarium solani* IISc-1; an extra copy of *nad1* and *nad4* was found from *E. oxysporum* UASWS AC1 (KR952337); the location of the two genes were

found to be reversed in *E. oxysporum* f. sp. *matthiolae* (CM019668); and the mt genome of *E. oxysporum* f. sp. *fragariae* GL1381 (CM029251) was found to lose *cox3* and *nad6* genes and possess an extra reversed copy of genes of *cob*, *cox1*, *nad1*, *nad4*, *atp8*, and *atp6*; an even extreme case was found in *Sarocladium implicatum* in which three genes (*cob*, *cox3*, and *nad6*) were lost and the *nad4* gene shifted its location from the *nad1-atp8* junction to a position between *rps3* and *nad2* (Supplementary Table S1).

### Phylogenetic analyses

Eighty-one complete mt genomes representing 63 distinct species from the order Hypocreales were included in phylogenetic analyses. After excluding the ambiguous aligned regions, a total of 4,345 amino acid sequences of 14 conserved proteins were retained. All species of Hypocreales formed a well-supported clade (BP = 100%) in ML analysis. Within the clade, four family-level subclades were recognized with strong supports (BP = 100%), that is, *Nectriaceae*, *Bionectriaceae*, *Hypocreaceae*, and *Clavicipitaceae* (Figure 2). Species in the family *Ophiocordycipitaceae* were clustered into two subclades, one subclade that consists of four *Tolyopcladium* species showed a sister group relationship with the *Clavicipitaceae* clade with low bootstrap support (BP = 75%), the other highly supported (BP = 100%) subclade comprised four *Hirsutella* species (*H. minnesotensis*, *H. rhossiliensis*, *H. thompsonii*, and *H. vermicola*) and *O. sinensis* (Figure 2). It is interesting to find that *P. sinensis* clustered with the *Clavicipitaceae* clade with 100% bootstrap support rather than grouped with either two subclades of *Ophiocordycipitaceae*.

### Discussion

The complete mt genome of the ex-type strain CN 80-2 of the species *P. sinensis* described here is the first reported case for the newly proposed genus *Pleurocordyceps* (Wang *et al.*, 2021) and *Polycephalomyces* s. l. It is rather compact compared with other Hypocreales species,

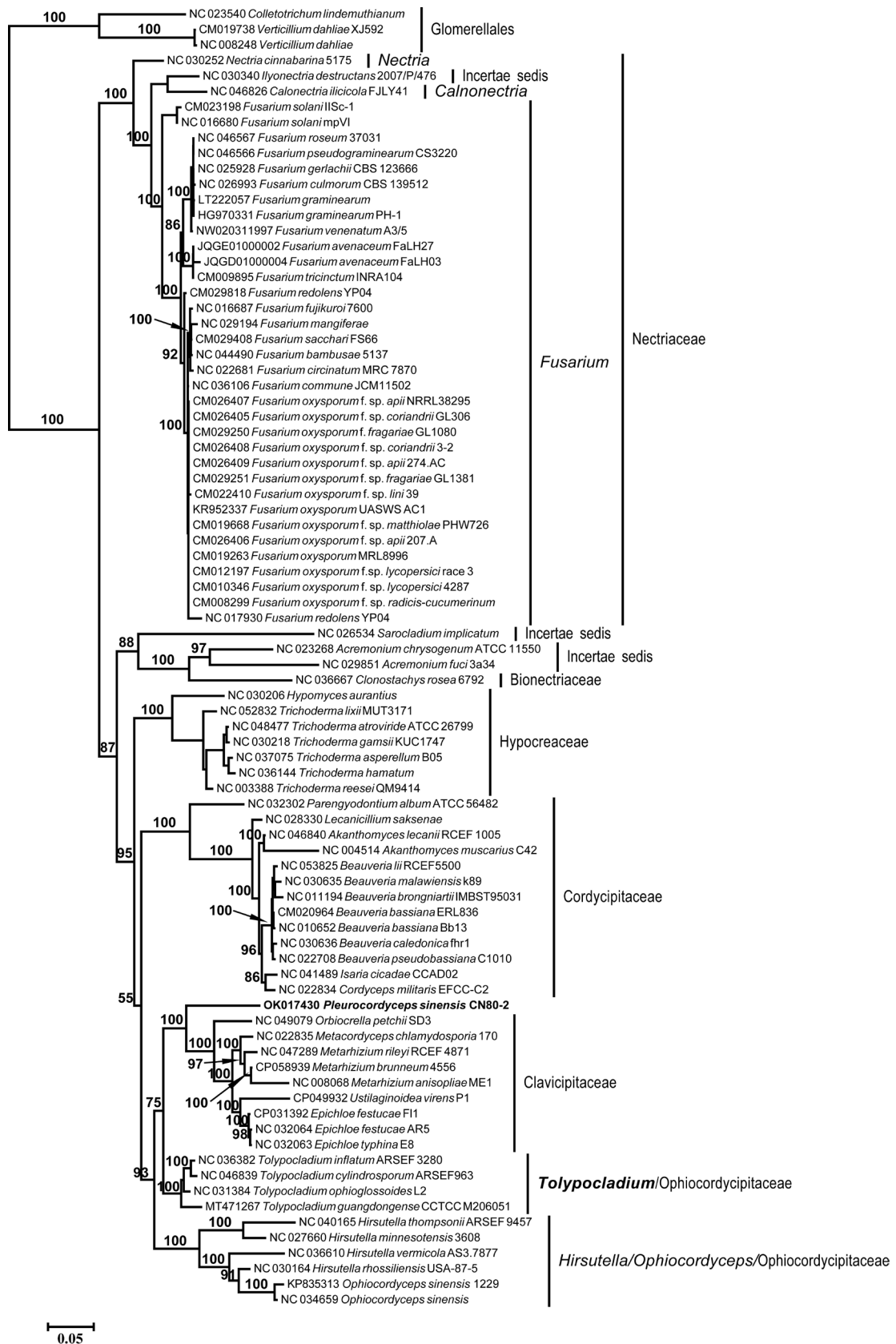


Figure 2. Phylogenetic relationships of Hypocreales inferred from 14 conserved protein sequences (*cox1*, *cox2*, *cox3*, *cob*, *atp6*, *atp8*, *atp9*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6*) using the maximum likelihood method. Bootstrap values were shown above the branches. Three species in the order *Glomerellales* were used as outgroups.

especially *O. sinensis*, from which the species was isolated. The genome sizes of the two sequenced isolates of *O. sinensis* were 157,510 bp (KP835313) and 157,539 bp (NC\_034659), respectively, almost five times larger than *P. sinensis*. The reason for this remarkable size variation was considered to be the presence of large numbers of repetitive regions, which mainly consisted of intronic mobile elements of HEs and reverse transcriptases (RTs) (Li *et al.*, 2015). In the expanded mt genome of *O. sinensis*, 32 HEs genes (21 LAGLIDADG and 11 GIY-YIG endonuclease) and 10 RTs genes were found in group I and group II introns, respectively (Li *et al.*, 2015), while only two HEs genes, that is, one GIY-YIG (ncORF3) and one LAGLIDADG endonuclease (ncORF9) were found in *P. sinensis*. After comparing mt genomes of three isolates of the same species, it was also found that larger genomes contained more introns and more intronic HEs genes in *Cordyceps militaris* (Zhang *et al.*, 2015).

As mt genome sizes varied greatly in hypocrealean fungi, ranging from 22,376 bp of *S. implicatum* (Yao *et al.*, 2016) to 272,497 bp of *Ophiocordyceps camponoti-floridani* (Will *et al.*, 2020), it will be interesting to see whether the mt genome size variation was related to expansion of mobile elements of HEs and RTs and how those elements evolved. Megarioti and Kouvelis (2020) recently proposed an “aenaon” model for the evolution of HEs genes and their host introns; thus, free-standing introns and HEs genes were the ancestral form and could invade intron-free coding genes together; HEs genes and their host introns coevolved through recombination, transposition, and horizontal gene transfer. As observed in this case, the two HEs genes found in *P. sinensis* were located in intergenic regions (free-standing or sole mobile in other words) of tRNA genes (Table 1) rather than invaded into intronic regions of coding genes (intron homing). While in *O. sinensis*, HEs and RTs genes were all found to be intronic, either in group I or group II introns (Li *et al.*, 2015). It may indicate that the species was earlier diverged than *O. sinensis* according to the “aenaon” model. Considering that the species was isolated from *O. sinensis* and might be a fungal parasite of the latter, and moreover, several other species in *Polycephalomyces* s. l. have often been found to associate with entomopathogenic *Cordyceps* s. l. (Kobayasi, 1941), it is reasonable to hypothesize that species of *Polycephalomyces* s. l. gained hyperparasitic ability to entomophagous fungi during the evolutionary process.

Despite the size variation of mt genome, the gene contents and synteny (gene order) are largely conserved within the order Hypocreales, generally encoding 15 known proteins and 2 rRNAs (*rnl* and *rns*) (Li *et al.*, 2015). The genome size variation observed in hypocrealean fungi was probably not associated with taxonomic classification since notable variation was also observed within the same genus or

even within the same species. As shown in Supplementary Table S1, mt genome sizes varied from 30,629 bp to 110,525 bp in the genus *Fusarium*, and from 34,477 bp to 52,424 bp within the species of *F. oxysporum*. This variation is largely due to the presence of various introns and the lengths of intergenic regions (Burger *et al.*, 2003).

Although remarkable variation in terms of gene order, genome size, composition of intergenic regions, and presence of repeats, introns, and associated ncORFs have been observed between the major fungal phyla (Aguileta *et al.*, 2014), this variance may not occur within the same fungal groups (order or below). As observed in this study, the genome size, composition of intergenic regions, and presence of repeats, introns, and associated ncORFs varied within the order Hypocreales; however, the gene content and synteny remained highly conserved even though a few exceptional cases were observed (listed in Supplementary Table S1). A part of these exceptions were probably due to the incorrect assembly. It would be interesting to know the mt genome evolutionary process, that is, gene gain and loss events happened during the evolutionary history of different major fungal groups.

Most protein-coding genes and ncORFs used standard mt initial and terminal codons (ATG and TAA, respectively) in *P. sinensis*, except the ncORF8, *cox1*, and ncORF10. ncORF10 and *cox1* were initiated by ATA, and ncORF8 was terminated by TAG. It is noteworthy that *cox1* is usually found to use nonstandard start codons such as TCG, ACC, CGA, CTA, CCG, and AAA in insect mt genomes (Fenn *et al.*, 2007; Wei *et al.*, 2010), and ATA has been recorded to be used as the initial codon in organisms like *Pseudocohnilembus persalinus* (Gao *et al.*, 2018), *Wellcomia siamensis* (Park *et al.*, 2011), and *Calanus sinicus* (Wang *et al.*, 2011). Although most hypocrealean species used ATG as the initial codon of *cox1* gene, exceptional cases were also reported in *Hirsutella rhossiliensis* (NC\_030164) and *Calonectria ilicicola* (Gai *et al.*, 2020), in which TTG were used.

The mt genome released in this study provided additional evidence that *P. sinensis* is not the anamorph of *O. sinensis* but represents another fungus, and moreover, *P. sinensis* was found to cluster with *Clavicipitaceae* rather than *Ophiocordycipitaceae* species with very strong supports (BP = 100%) in ML phylogenetic analyses (Figure 2). It is also noteworthy that the family *Ophiocordycipitaceae* was paraphyletic, which contradicts previous studies applying multi-gene phylogeny (Sung *et al.*, 2007a) although the paraphyly was not well supported (Figure 2). It should be clarified whether those contradictions were due to the incongruence of phylogenies revealed by different molecular markers since mt DNA may tell different evolutionary stories than nuclear genes (Burger *et al.*, 2003), or were just caused by the insufficient taxon

sampling or analytical difference. Sung *et al.* (2007b) conducted multi-gene phylogenetic analyses of clavicipitaceous fungi and compared the performance of seven loci including the nuclear ribosomal small and large subunit DNA (*nrSSU* and *nrLSU*),  $\beta$ -*tubulin*, elongation factor 1 $\alpha$  (*EF-1 $\alpha$* ), the largest and second largest subunits of RNA polymerase II (*RPB1* and *RPB2*), and one mt protein-coding gene ATP Synthase subunit 6 (*mtATP6*), and found that seven genes gave incongruent topologies in higher-level relationships from each other and also from the combined dataset. It also showed that the only mt fragment (*mtATP6*) used in the study possessed localized incongruence and simultaneously provided an increased level of support for certain nodes. Phylogenetic incongruence revealed by different markers, especially those from mt and nuclear fragments, respectively, has been frequently reported and compared in different organisms (e.g. Kimball *et al.*, 2021; Mikula *et al.*, 2021; Zhang *et al.*, 2021). It still remains unclear as to whether the nuclear genome sequences (fragments or whole genome data) or the mt genome sequences (fragments or complete data) could provide a better resolution of fungal phylogeny.

As in the case of clavicipitaceous fungi, almost all the later publications on taxonomy and phylogenetic studies accepted the backbone phylogeny created by Sung *et al.* (2007a, 2007b), and continued to use the five gene dataset (e.g. Kepler *et al.*, 2013; Wang *et al.*, 2021; Xiao *et al.*, 2018). While the above two studies (Sung *et al.*, 2007a, 2007b) failed to include species of *Polycephalomyces* s. l. Kepler *et al.* (2013) then included several species of this group and found those species represented a clade distinct from other clavicipitoid genera, and treated them as incertae sedis of Hypocreales. Further studies are needed to reconstruct a reliable phylogenetic relationship of clavicipitaceous fungi, especially the assignment of species of *Polycephalomyces* s. l. Since an increasing number of whole mt genomes have recently been sequenced and released for hypocrealean species, and even more are being proceeded, the plentiful phylogenetically informative sites from the conserved protein coding genes of the whole mt genome would provide valuable information for phylogenetic reconstruction of this important fungal group. While publically released data should be carefully treated since they could probably include assembly and annotation errors as observed in *Ophiocordyceps camponoti-floridani* EC05 (CM022976) and *Ophiocordycipitaceae* sp. (NC\_049089), correct annotation and characterization are always necessary (Kortsinoglou *et al.*, 2019; Megarioti and Kouvelis, 2020).

## Acknowledgments

The authors would like to thank Sheng-Rong Xiao for his generous donation of the ex-type strain CN80-2. This

work was supported by the National Science Foundation of China (31400018, 31170017, 31700009, 32170001) and the Natural Science Fund for Colleges and Universities in Jiangsu Province (17KJB350003).

## Conflict of Interest

The authors declare that they have no competing interests.

## References

- Abuduaini, A., Wang, Y.B., Zhou, H.Y., Kang, R.P., Ding, M.L., Jiang, Y., et al. 2021. The complete mitochondrial genome of *Ophiocordyceps gracilis* and its comparison with related species. *IMA Fungus* 12: 1–14. <https://doi.org/10.1186/s43008-021-00081-z>
- Aguilera, G., de Vienne, D.M., Ross, O.N., Hood, M.E., Giraud, T., Petit, E., et al. 2014. High variability of mitochondrial gene order among fungi. *Genome Biology and Evolution* 6: 451–465. <https://doi.org/10.1093/gbe/evu028>
- Basse, C.W., 2010. Mitochondrial inheritance in fungi. *Current Opinion in Microbiology* 13: 712–719. <https://doi.org/10.1016/j.mib.2010.09.003>
- Beck, N. and Lang, B., 2010. MFannot, organelle genome annotation webserver. Université de Montréal, Montréal, QC.
- Berbee, M.L. and Taylor, J.W., 2001. Fungal molecular evolution: gene trees and geologic time. In: McLaughlin, D.J., McLaughlin, E.G. and Lemke, P.A. (eds.) *Systematics and evolution. The Mycota (A comprehensive treatise on fungi as experimental systems for basic and applied research)*, vol 7B: 229–243. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-662-10189-6\\_10](https://doi.org/10.1007/978-3-662-10189-6_10)
- Burger, G., Gray, M.W. and Lang, B.F., 2003. Mitochondrial genomes: anything goes. *Trends in Genetics* 19: 709–716. <https://doi.org/10.1016/j.tig.2003.10.012>
- Chaisson, M.J. and Tesler, G., 2012. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* 13: 238. <https://doi.org/10.1186/1471-2105-13-238>
- Chen, Q.T., Xiao, S.R. and Shi, Z.Y., 1984. *Paecilomyces sinensis* sp. nov. and its connection with *Cordyceps sinensis*. *Acta Mycologica Sinica* 3: 24–28 (in Chinese with English abstract).
- Chen, Y.Q., Wang, N., Qu, L.H., Li, T.H. and Zhang, W.M., 2001. Determination of the anamorph of *Cordyceps sinensis* inferred from the analyses of the ribosomal DNA internal transcribed spacers and 5.8S rDNA. *Biochemical Systematics and Ecology* 29: 597–607. [https://doi.org/10.1016/S0305-1978\(00\)00100-9](https://doi.org/10.1016/S0305-1978(00)00100-9)
- Cheng, L., Xu, P.X. and Tang, Y., 2005. Protective effects of CN80-2 on immunological liver injury in mice. *Chinese Journal of Clinical Pharmacology and Therapeutics* 10: 318–320 (in Chinese with English abstract).
- Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., et al. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods* 10: 563–569. <https://doi.org/10.1038/nmeth.2474>

- Chris, S., Francesco, F., Andrew, B., Bernie, C., Hong, L. and Paul, F., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* 87: 651–701. <https://doi.org/10.1093/aesa/87.6.651>
- Crous, P.W., Wingfield, M.J., Burgess, T.I., Hardy, G.E.S.J., Barber, P.A., Alvarado, P., et al. 2017. Fungal planet description sheets: 558–624. *Persoonia* 38: 240–384. <https://doi.org/10.3767/003158517X698941>
- Fan, W.W., Zhang, S. and Zhang, Y.J., 2019. The complete mitochondrial genome of the Chan-hua fungus *Isaria cicadae*: a tale of intron evolution in Cordycipitaceae. *Environmental Microbiology* 21: 864–879. <https://doi.org/10.1111/1462-2920.14522>
- Fang, H.M., 1991. Notes on anamorph determination of the genus *Cordyceps*. In: Li, Y.W., Li, Z.Z., Wu, Z.K., Chen, Z.A., Wu, J.W., Liang, Z.Q. and Fan, M.Z., editors. Study and application of entomogenous fungi in China. Vol. 2, pp. 67–68. China Agricultural Sciencetech Press, Beijing.
- Fenn, J.D., Cameron, S.L. and Whiting, M.F., 2007. The complete mitochondrial genome sequence of the *Mormon cricket* (*Anabrus simplex*: Tettigoniidae: Orthoptera) and an analysis of control region variability. *Insect Molecular Biology* 16: 239–252. <https://doi.org/10.1111/j.1365-2583.2006.00721.x>
- Gai, Y.P., Pan, R.Q. and Peng, X.J., 2020. A phylogenomic tree of fungi: evolutionary relationships among *Calonectria ilicicola* and 586 fungal mitochondrial genomes. *Mitochondrial DNA B* 5: 1709–1711. <https://doi.org/10.1080/23802359.2020.1749163>
- Gao, Y.Q., Jin, S.B., Dang, H.F., Ye, S.G. and Li, R.J., 2018. Mitochondrial genome sequencing of notorious scuticociliates (*Pseudocohnilembus persalinus*) isolated from Turbot (*Scophthalmus maximus* L.). *Mitochondrial DNA B* 3: 1077–1078. <https://doi.org/10.1080/23802359.2018.1508388>
- Ge, Z.H., Wang, R.Y. and Lin, Z.Q., 1989. Effect of *Paecilomyces sinensis* on IgM antibody forming cells in mice lymph gland. *Chinese Journal of Immunology* 5: 117 (in Chinese).
- Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98. <https://doi.org/10.1021/bk-1999-0734.ch008>
- Huang, Z.Q., Li, C.C., Wu, D.F. and Lin, J.Z., 1988. Antitumor activity and toxicity of *Paecilomyces sinensis* sp. nov. (CN80-2). *Cancer Research and Treatment* 15: 124–126 (in Chinese with English abstract).
- Jiang, Y. and Yao, Y.J., 2002. Names related to *Cordyceps sinensis* anamorph. *Mycotaxon* 84: 245–254.
- Jiang, Y. and Yao, Y.J., 2003. Anamorphic fungi related to *Cordyceps sinensis*. *Mycosystema* 22: 161–176.
- Kepler, R., Ban, S., Nakagiri, A., Bischoff, J., Hywel-Jones, N., Owensby, C.A. and Spatafora, J.W., 2013. The phylogenetic placement of hypocrealean insect pathogens in the genus *Polycephalomyces*: an application of one fungus one name. *Fungal Biology* 117: 611–622. <https://doi.org/10.1016/j.funbio.2013.06.002>
- Kimball, R.T., Guido, M., Hosner, P.A. and Braun, E.L., 2021. When good mitochondria go bad: cyto-nuclear discordance in landfowl (Aves: Galliformes). *Gene* 801: 145841. <https://doi.org/10.1016/j.gene.2021.145841>
- Kobayasi, Y., 1941. The genus *Cordyceps* and its allies. Report of the Tokyo Bunrika Daigaku Section B 5: 53–260.
- Kortsinoglou, A.M., Korovesi, A.G., Theelen, B., Hagen, F., Boekhout, T. and Kouvelis, V.N., 2019. The mitochondrial intergenic regions *nad1-cob* and *cob-rps3* as molecular identification tools for pathogenic members of the genus *Cryptococcus*. *FEMS Yeast Research* 19: foz077. <https://doi.org/10.1093/femsyr/foz077>
- Kouvelis, V.N., Ghikas, D.V. and Typas, M.A., 2004. The analysis of the complete mitochondrial genome of *Lecanicillium muscarium* (synonym *Verticillium lecanii*) suggests a minimum common gene organization in mtDNAs of Sordariomycetes: phylogenetic implications. *Fungal Genetics and Biology* 41: 930–940. <https://doi.org/10.1016/j.fgb.2004.07.003>
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. 2009. Circos: an information aesthetic for comparative genomics. *Genome Research* 19: 1639–1645. <https://doi.org/10.1101/gr.092759.109>
- Li, C.C., Huang, Z.Q., Guo, X.B., Lin, J.Z. and Xue, W.J., 1983. Pharmacological study on *Cordyceps sinensis* and *Paecilomyces sinensis*. *Fujian Medical Journal* 5: 51–54 (in Chinese).
- Li, C.C. and Lin, Q.Q., 1991. Pharmacological study of *Paecilomyces sinensis* sp. nov. (CN80-2). *Edible Fungi of China* 10: 16–17 (in Chinese with English abstract).
- Li, Y., Hu, X.D., Yang, R.H., Hsiang, T., Wang, K., Liang, D.Q., Liang, F., et al. 2015. Complete mitochondrial genome of the medicinal fungus *Ophiocordyceps sinensis*. *Scientific Reports* 5: 13892. <https://doi.org/10.1038/srep13892>
- Li, Z.Z., Huang, B., Li, C.R. and Fan, M.Z., 2000. Molecular evidence for anamorph determination of *Cordyceps Sinensis* (Berk.) Sacc. I. Relation between *Hirsutella sinensis* and *C. sinensis*. *Mycosystema* 19: 60–64.
- Liang, P.Q., 1991. Current status of studies on *Cordyceps* spp. in China. In: Li, Y.W., Li, Z.Z., Wu, Z.K., Chen, Z.A., Wu, J.W., Liang, Z.Q. and Fan, M.Z., editors. Study and application of entomogenous fungi in China. Vol. 2, pp. 55–57. China Agricultural Sciencetech Press, Beijing. (in Chinese with English abstract).
- Lin, Q.X., Qiu, S.Y., Li, C.C. and Liu, B.X., 1988. Effects of anti-implantation in mice by *Paecilomyces sinensis* sp. nov. (CN80-2). *Journal of Fujian Medical University* 22: 210–212 (in Chinese with English abstract).
- Lin, S.W., Liu, Y.S., Lin, Y.Y., Lin, M.F., Wang, Y.X. and Zhu, Z., 1987. Regulation of *Cordyceps sinensis* and *Paecilomyces sinensis* on cellular immune function. *Chinese Traditional Patent Medicine* 12: 22–23 (in Chinese).
- Liu, J.L., 1990. Anamorph of *Cordyceps* and artificial cultivation of its fruiting body. *Journal of Guizhou Agriculture Science* 1: 43–48 (in Chinese).
- Liu, Y.Y., Wu, C.Z. and Li, C.C., 1991. Anti-oxidation of *Paecilomyces sinensis* sp. nov. *Journal of Fujian Medical University* 16: 240–242 (in Chinese with English abstract).
- Liu, Y.Y., Wu, C.Z., Li, C.C. and Huang, D.H., 1989. Experimental on antioxidant activity of *Paecilomyces sinensis*. *Journal of Fujian Medical University* 11: 33–35 (in Chinese with English abstract).

- Liu, Y.Y., Wu, C.Z., Xu, Y.C. and Li, C.C., 1987. The effect of *Paecilomyces sinensis* on the level lipid peroxide of mice. Journal of Fujian Medical University 21: 86–88 (in Chinese with English abstract).
- Mains, E.B., 1948. Entomogenous fungi. Mycologia 40: 402–416. <https://doi.org/10.1080/00275514.1944.12017718>
- Matočec, N., Kušan, I. and Ozimec, R., 2014. The genus *Polycephalomyces* (Hypocreales) in the frame of monitoring Veternica cave (Croatia) with a new segregate genus *Perennicordyceps*. Ascomycete.org 6: 125–133.
- Megarioti, A.H. and Kouvelis, V.N., 2020. The coevolution of fungal mitochondrial introns and their homing endonucleases (GIY-YIG and LAGLIDADG). Genome Biology and Evolution 12: 1337–1354. <https://doi.org/10.1093/gbe/evaa126>
- Mikula, O., Nicolas, V., Šumbera, R., Konečný, A., Denys, C., Verheyen, E., et al. 2021. Nuclear phylogenomics, but not mitogenomics, resolves the most successful Late Miocene radiation of African mammals (Rodentia: Muridae: Arvicanthini). Molecular Phylogenetics and Evolution 157: 107069. <https://doi.org/10.1016/j.ympev.2021.107069>
- Myers, E.W., Sutton, G.G., Delcher, A.L., Dew, I.M., Fasulo, D.P., Flanigan, M.J., et al. 2000. A whole-genome assembly of *Drosophila*. Science 287: 2196–2204. <https://doi.org/10.1126/science.287.5461.2196>
- Park, J.K., Sultana, T., Lee, S.H., Kang, S., Kim, H.K., Min, G.S., et al. 2011. Monophyly of clade III nematodes is not supported by phylogenetic analysis of complete mitochondrial genome sequences. BMC Genomics 12: 392–407. <https://doi.org/10.1186/1471-2164-12-392>
- Samson, R.A., Evans, H.C. and Van, D.K.G., 1981. Notes on entomogenous fungi from Ghana. V. The genera *Stilbella* and *Polycephalomyces*. Proceedings. Series C. Biological and Medical Sciences 84: 289–301.
- Seifert, K.A., 1986. A monograph of *Stilbella* and some allied Hyphomycetes. Studies in Mycology 78: 980–986. <https://doi.org/10.2307/3807446>
- Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690. <https://doi.org/10.1093/bioinformatics/btl446>
- Sun, H.H., Zhang, Y.J. and Zhang, S., 2021. Complete mitogenome of the entomopathogenic fungus *Metarhizium album* and phylogenetic analysis of Hypocreales. Mitochondrial DNA B 6: 1689–1690. <https://doi.org/10.1080/23802359.2021.1914229>
- Sung, G.H., 2015. Complete mitochondrial DNA genome of the medicinal mushroom *Cordyceps militaris* (Ascomycota, Cordycipitaceae). Mitochondrial DNA 26: 789–790. <https://doi.org/10.3109/19401736.2013.855754>
- Sung, G.H., Hywel-Jones, N.L., Sung, J.M., Luangsa-Ard, J.J., Shrestha, B. and Spatafora, J.W., 2007a. Phylogenetic classification of *Cordyceps* and the clavicipitaceous fungi. Studies in Mycology 57: 5–59. <https://doi.org/10.3114/sim.2007.57.01>
- Sung, G.H., Sung, J.M., Hywel-Jones, N.L. and Spatafora, J.W., 2007b. A multi-gene phylogeny of Clavicipitaceae (Ascomycota, Fungi): identification of localized incongruence using a combinational bootstrap approach. Molecular Phylogenetics and Evolution 44: 1204–1223. <https://doi.org/10.1016/j.ympev.2007.03.011>
- Wang, L., Li, H.H., Chen, Y.Q., Zhang, W.M. and Qu, L.H., 2015a. *Polycephalomyces lianzhouensis* sp. nov., a new species, co-occurs with *Ophiocordyceps crinalis*. Mycological Progress 13: 1089–1096. <https://doi.org/10.1007/s11557-014-0996-9>
- Wang, M.X., Sun, S., Li, C.L. and Shen, X., 2011. Distinctive mitochondrial genome of Calanoid copepod *Calanus sinicus* with multiple large non-coding regions and reshuffled gene order: useful molecular markers for phylogenetic and population studies. BMC Genomics 12: 73–93. <https://doi.org/10.1186/1471-2164-12-73>
- Wang, W.J., Wang, X.L., Li, Y., Xiao, S.R., Kepler, R.M. and Yao, Y.J., 2012. Molecular and morphological studies of *Paecilomyces sinensis* reveal a new clade in clavicipitaceous fungi and its new systematic position. Systematics and Biodiversity 10: 221–232. <https://doi.org/10.1080/14772000.2012.690784>
- Wang, Y.B., Yu, H., Dai, Y.D., Wu, C.K., Zeng, W.B., Yuan, F., et al. 2015b. *Polycephalomyces agaricus*, a new hyperparasite of *Ophiocordyceps* sp. infecting melolonthid larvae in southwestern China. Mycological Progress 14: 70–79. <https://doi.org/10.1007/s11557-015-1090-7>
- Wang, Y.H., Ban, S., Wang, W.J., Li, Y., Wang, K., Kirk, P.M., et al. 2021. *Pleurocordyceps* gen. nov. for a clade of fungi previously included in *Polycephalomyces* based on molecular phylogeny and morphology. Journal of Systematics and Evolution 59: 1065–1080. <https://doi.org/10.1111/jse.12705>
- Wei, S.J., Pu, T., Zheng, L.H., Min, S. and Chen, X.X., 2010. The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Evaniidae) has low A+T content and a long intergenic spacer between *atp8* and *atp6*. Molecular Biology Reports 37: 1931–1942. <https://doi.org/10.1007/s11033-009-9640-1>
- Will, I., Das, B., Trinh, T., Brachmann, A., Ohm, R.A. and de Bekker, C., 2020. Genetic underpinnings of host manipulation by *Ophiocordyceps* as revealed by comparative transcriptomics. G3-Genes Genomes Genetics 10: 2275–2296. <https://doi.org/10.1534/g3.120.401290>
- Winter, D.J., Ganley, A.R.D. and Young, C.A., 2018. Repeat elements organise 3D genome structure and mediate transcription in the filamentous fungus *Epichloë festucae*. PLoS Genetics 14: 1007467–1007499. <https://doi.org/10.1371/journal.pgen.1007467>
- Wu, D.F., Zheng, Z.X., Zhang, Y., Fang, C. and Li, C.C., 1986. Inhibition of human uterus cancer cell line by Cordycepin and *Paecilomyces sinensis* in vitro. Chinese Journal of Cancer 5: 337–340 (in Chinese with English abstract).
- Xiao, Y.P., Wen, T.C., Hongsanan, S., Jeewon, R., Luangsa-ard, J.J., Brooks, S., et al. 2018. Multigene phylogenetics of *Polycephalomyces* (Ophiocordycipitaceae, Hypocreales), with two new species from Thailand. Scientific Reports 8: 18087–18098. <https://doi.org/10.1038/s41598-018-36792-4>
- Yang, J.I., Stadler, M., Chuang, W.Y., Wu, S. and Ariyawansa, H.A., 2020. In vitro inferred interactions of selected entomopathogenic fungi from Taiwan and eggs of *Meloidogyne graminicola*. Mycological Progress 19: 97–109. <https://doi.org/10.1007/s11557-019-01546-7>

- Yao, Y.R., Lin, R.M., Tian, X.L., Shen, B.M., Mao, Z.C. and Xie, B.Y., 2016. The complete mitochondrial genome of the nematophagous fungus *Acremonium implicatum*. *Mitochondrial DNA A* 27: 3246–3247. <https://doi.org/10.3109/19401736.2015.1007367>
- You, J.G., Chen, B.W., You, J.C., Lin, B.H., Ye, Y., Li, Y.J., et al. 1986. Clinical observation on 33 cases of coronary heart disease treated with *Cordyceps sinensis* granules (*Paecilomyces sinensis*). *Fujian Medical Journal* 5: 24–25 (in Chinese).
- Zeng, X.K., Tang, Y. and Yuan, S.R., 2000. Effect of CS and CN80-2 on T-lymphocyte subsets and natural killer cell activities. *Pharmacy and Clinics of Chinese Materia Medica* 16: 21–23 (in Chinese with English abstract).
- Zhang, C.K., Yuan, S.R. and Liu, J.X., 1998. Effect of *Cordyceps sinensis* (CS) and *Paecilomyces sinensis* (PS) on immune function in mice. *Pharmacy and Clinics of Chinese Materia Medica* 14: 21–23 (in Chinese with English abstract).
- Zhang, S. and Zhang, Y.J., 2020. Complete mitogenome of the entomopathogenic fungus *Tolypocladium cylindrosporium*. *Mitochondrial DNA B* 5: 680–682. <https://doi.org/10.1080/23802359.2020.1714495>
- Zhang, S.L., Pu, S.C., Lin, A.T. and Luan, F.G., 2021. The complete mitochondrial genome of *Beauveria lii* (Hypocreales: Cordycipitaceae). *Mitochondrial DNA B* 6: 586–588. <https://doi.org/10.1080/23802359.2021.1875917>
- Zhang, Y.J., Zhang, S. and Liu, X.Z., 2016. The complete mitochondrial genome of the nematode endoparasitic fungus *Hirsutella minnesotensis*. *Mitochondrial DNA A* 27: 2693–2694. <https://doi.org/10.1007/s00253-017-8257-x>
- Zhang, Y.J., Zhang, S., Zhang, G.Z., Liu, X.Z., Wang, C.S. and Xu, J.P., 2015. Comparison of mitochondrial genomes provides insights into intron dynamics and evolution in the caterpillar fungus *Cordyceps militaris*. *Fungal Genetics and Biology* 77: 95–107. <https://doi.org/10.1016/j.fgb.2015.04.009>
- Zhao, J., Wang, N., Chen, Y.Q., Li, T.H. and Qu, L.H., 1999. Molecular identification for the asexual stage of *Cordyceps sinensis*. *Acta Scientiarum Naturalium Universitatis Sunyatseni* 38: 121–123 (in Chinese with English abstract).
- Zheng, Y.L., Ye, J.R., Lin, D.J., Xu, Y. and Chen, W.X., 1983. Effects of *Cordyceps sinensis* and *Paecilomyces sinensis* on immune function. *Fujian Medical Journal* 5: 55–57 (in Chinese).



<i>Sarcodinium implicatum</i>	NC_0206534	Hypocreales	Incertae sedis	22,376	26.12	rps3	nad4	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Colonectria lilicola</i> FJLY41	NC_046826	Hypocreales	Nectriaceae	39,891	28.48	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium graminearum</i>	LT222057	Hypocreales	Nectriaceae	100,131	32.38	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium culmorum</i> CBS_139512	NC_020893	Hypocreales	Nectriaceae	103,844	31.68	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium pseudograminearum</i> CS3220	NC_046566	Hypocreales	Nectriaceae	110,525	31.64	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium mangiferae</i>	NC_020194	Hypocreales	Nectriaceae	30,629	31.25	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> F11	NC_017930	Hypocreales	Nectriaceae	34,477	30.98	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. lini 39	CM022410	Hypocreales	Nectriaceae	38,745	31.02	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. fragariae GL1381	CM029251	Hypocreales	Nectriaceae	40,945	31.32	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. fragariae GL1080	CM029250	Hypocreales	Nectriaceae	45,629	32.08	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. api 274.AC	CM026409	Hypocreales	Nectriaceae	45,699	32.13	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. api NRRL38295	CM026407	Hypocreales	Nectriaceae	45,699	32.13	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. coriandrii 3-2	CM026408	Hypocreales	Nectriaceae	45,699	32.13	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. coriandrii GL306	CM026405	Hypocreales	Nectriaceae	45,699	32.13	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium commune</i> JCM11502	NC_036106	Hypocreales	Nectriaceae	47,526	32.42	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. radialis-cucumerinum ForcD16	CM008299	Hypocreales	Nectriaceae	47,541	32.23	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. api 207.A	CM026406	Hypocreales	Nectriaceae	47,671	32.30	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium tricinatum</i> INRA104	CM009895	Hypocreales	Nectriaceae	48,506	33.05	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> MRL8996	CM019263	Hypocreales	Nectriaceae	48,644	32.00	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium avenaceum</i> FalH27	JQGE01000002	Hypocreales	Nectriaceae	49,396	33.06	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium avenaceum</i> FalH03	JQGD01000004	Hypocreales	Nectriaceae	49,402	33.09	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium redolens</i> YP04	CM029818	Hypocreales	Nectriaceae	49,602	32.05	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> UASWS AC1	KR952337	Hypocreales	Nectriaceae	51,536	31.91	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. lycopersici race 3	CM012197	Hypocreales	Nectriaceae	52,353	31.46	rps3		nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad1	nad4	atp8	atp6	cox3	nad6

(continues)

**Table S1. Continued**

Organism current Name	Accession	Order	Family	Size (Kb)	GC%	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. <i>mathiobolae</i> PHW726	CM019668	Hypocreales	Nectriaceae	52.365	31.54	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium oxysporum</i> f. sp. <i>lyopersici</i> 4287	CM010346	Hypocreales	Nectriaceae	52.424	31.47	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium fujikuroi</i> 7600	NC_016687	Hypocreales	Nectriaceae	53.753	32.61	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium solani</i> HSc-1	CM023198	Hypocreales	Nectriaceae	59.514	28.78	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	atp9	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium sacchari</i> FS66	CM029408	Hypocreales	Nectriaceae	59.755	32.35	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium solani</i> mpVI	NC_016680	Hypocreales	Nectriaceae	62.978	28.88	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium bambusae</i> 5137	NC_044490	Hypocreales	Nectriaceae	63.593	31.92	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium circinatum</i> MRC 7870	NC_022681	Hypocreales	Nectriaceae	67.109	31.45	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium venenatum</i> A35	NW_020311997	Hypocreales	Nectriaceae	78.612	31.71	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium roseum</i> 37031	NC_046557	Hypocreales	Nectriaceae	93.160	31.74	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium gielachii</i> CBS 123666	NC_025928	Hypocreales	Nectriaceae	93.428	31.91	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Fusarium graminearum</i> PH-1	HG970331	Hypocreales	Nectriaceae	95.638	31.83	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Nectria cinnabarina</i> 5175	NC_030252	Hypocreales	Nectriaceae	69.895	28.71	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Hirsutella minnesotensis</i> 3608	NC_027660	Hypocreales	Ophiocordycipitaceae	52.245	28.42	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Hirsutella rhossiliensis</i> USA-87-5	NC_030164	Hypocreales	Ophiocordycipitaceae	62.483	28.21	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Hirsutella thompsonii</i> ARSEF 9457	NC_040165	Hypocreales	Ophiocordycipitaceae	62.509	29.82	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Hirsutella vermicola</i> AS3.7877	NC_038610	Hypocreales	Ophiocordycipitaceae	53.793	25.27	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Ophiocordyceps camponoti-floridani</i> EC05a	CM022976	Hypocreales	Ophiocordycipitaceae	272.497	27.56	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1(3)	nad1(3)	nad4(2)	atp8	atp6	cox3	nad6
<i>Ophiocordyceps sinensis</i>	NC_034659	Hypocreales	Ophiocordycipitaceae	157.539	30.20	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Ophiocordyceps sinensis</i> 1229	KP835313	Hypocreales	Ophiocordycipitaceae	157.510	30.19	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Ophiocordycipitaceae</i> sp.a	NC_049089	Hypocreales	Ophiocordycipitaceae	66.785	30.58	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Pleurocordyceps sinensis</i> CN80-2	OK017430	Hypocreales	Ophiocordycipitaceae	31.841	25.46	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Tolypocladium cylindrosporum</i> ARSEF963	NC_046839	Hypocreales	Ophiocordycipitaceae	34.698	26.97	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Tolypocladium inflatum</i> ARSEF 3280	NC_035392	Hypocreales	Ophiocordycipitaceae	253.28	27.79	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Tolypocladium ophioglossoides</i> L2	NC_031384	Hypocreales	Ophiocordycipitaceae	351.59	27.53	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6
<i>Tolypocladium guangdongense</i> GD15	MT471267	Hypocreales	Ophiocordycipitaceae	46.102	26.10	rps3	nad2	nad3	atp9	cox2	nad4L	nad5	cob	cox1	nad4	nad1	atp8	atp6	cox3	nad6

a Taxa used in comparative analysis while excluded from phylogenetic reconstruction due to the possible error occurred during sequencing and assembly;

b Originally annotated as rps5 but really represent rps3;

c Exceptionals from the majority were highlighted.