Comparative analysis of the complete mitochondrial genomes of four cordyceps fungi

can zhong[1], jian jin[1], rongrong zhou[1], hao liu [1], jing xie[1], dan wan[1], shengen xiao[2], and shuihan zhang[1]

[1]Hunan Academy of Chinese Medicine
[2]Hunan Agricultural University

November 23, 2021

Abstract

Cordyceps is a large group of entomogenous, medicinally important fungi. In this study, we sequenced, assembled, and annotated the entire mitochondrial genome of O. xuefengensis, in addition to comparing it against three other complete cordyceps mitogenomes that were previously published. Comparative analysis indicated that the four complete mitogenomes are all composed of circular DNA molecules, although their sizes significantly differ due to high variability in intron and intergenic region sizes in the O. sinensis and O. xuefengensis mitogenomes. All mitogenomes contain 14 conserved genes and two ribosomal RNA genes, but varying numbers of tRNA introns. The Ka/Ks ratios for all 14 PCGs and rps3 were all less than 1, indicating that these genes have been subject to purifying selection. Phylogenetic analysis was conducted using concatenated amino acid and nucleotide sequences of the 14 PCGs and rps3 using two different methods (Maximum Likelihood and Bayesian analysis), revealing highly supported relationships between O. xuefengensis and other Ophiocordyceps species, in addition to a close relationship with O. sinensis. Further, the analyses indicated that cox1 and rps3 play important roles in population differentiation. These mitogenomes will allow further study of the population genetics, taxonomy, and evolutionary biology of medicinally important cordyceps species.

Article

Comparative analysis of the complete mitochondrial genomes of four cordyceps fungi

Can Zhong [1,2], Jian Jin [2], Rongrong Zhou [3], Hao Liu [2], Jing Xie[2], Dan Wan[2], Shengen Xiao[1]*, and Shuihan Zhang [1,2]*

[1] Horticulture and Landscape College, Hunan Agricultural University, Changsha 410128, P. R. China,[2] Institute of Chinese Materia Medica, Hunan Academy of Chinese Medicine, Changsha 410013, P. R. China,[3] Changchun University of Chinese Medicine, Changchun 130117, P. R. China.***** Correspondence: 2395265030@qq.com (S. X.),*zhangshuihan0220@126.com*(S.Z.) Received: date; Accepted: date; Published: date

Abstract: Cordyceps is a large group of entomogenous, medicinally important fungi. In this study, we sequenced, assembled, and annotated the entire mitochondrial genome of *O. xuefengensis* , in addition to comparing it against three other complete cordyceps mitogenomes that were previously published. Comparative analysis indicated that the four complete mitogenomes are all composed of circular DNA molecules, although their sizes significantly differ due to high variability in intron and intergenic region sizes in the *O. sinensis* and *O. xuefengensis* mitogenomes. All mitogenomes contain 14 conserved genes and two ribosomal RNA genes, but varying numbers of tRNA introns. The Ka/Ks ratios for all 14 PCGs and*rps3* were all less than 1, indicating that these genes have been subject to purifying selection. Phylogenetic analysis was

conducted using concatenated amino acid and nucleotide sequences of the 14 PCGs and *rps3* using two different methods (Maximum Likelihood and Bayesian analysis), revealing highly supported relationships between *O. xuefengensis* and other *Ophiocordyceps* species, in addition to a close relationship with *O. sinensis* . Further, the analyses indicated that *cox1* and *rps3* play important roles in population differentiation. These mitogenomes will allow further study of the population genetics, taxonomy, and evolutionary biology of medicinally important cordyceps species.

Keywords: cordyceps, mitochondrial genome, phylogenetic analysis, introns

1. Introduction

Mitochondrial genomes are powerful tools that can be used in phylogenetic and evolutionary studies [1-4], owing to their small size, high copy numbers, conserved orthologous genes, rare instances of recombination, and high evolutionary rates [5]. The rapid development of new genome sequencing technologies in recent years has led to the production of more than 180 fungal mitogenomes. Fungal mitochondria range in length from approximately 11 kbp (*Hanseniaspora uvarum*) to 272.2 kbp (*Morchella importuna*), depending on the species, but usually contain 14 conserved genes [6]. In addition, mitogenome characteristics, including genome size, genome structure, gene content, gene arrangement, tRNA number, intron number, codon usage, and repeat content, can provide useful information to understand the origin, evolution, and systematics of eukaryotes [7-9].

"Cordyceps" comprises four genera including *Cordyceps* ,*Ophiocordyceps* , *Metacordyceps* , and *Elaphocordyceps* . The fungi have been used in traditional Chinese medicine, as dietary supplements, or as tonic edibles with broad pharmacological properties including antibacterial, antiviral, and antitumor activities, in addition to the ability to enhance human immunity [10, 11]. Although *Cordyceps* have considerable economic and medicinal significance, wild cordyceps resources are on the verge of exhaustion, leading to many investigators seeking alternatives. Among these,*O. xuefengensis* is a newly identified cordyceps resource that we previously discovered on Xuefeng mountain in the Hunan province of China [12]. The novel taxa has been widely used as an ethnopharmacological invigorant by the Yao group for a long time [13]. Further, we have successfully cultivated this cordyceps fungi and also characterized some of its chemical metabolites [14]. *Cordyceps* and*Ophiocordyceps* are the two most widely used groups of cordyceps fungi. In particular, *O. sinensis* is popularly used to treat many diseases in China. Consequently, the cultivation, novel metabolite compounds, and pharmacology of cordyceps fungi have been extensively researched, although further molecular phylogenetic and species identification studies are critically needed.

In the present study, the complete mitogenome from *O. xuefengensis* was sequenced, assembled, and compared to three other complete cordyceps mitogenomes that were previously reported. This study thereby represents the first comparative analysis of the genomic structure, base composition, substitution and evolutionary rates among four cordyceps species, in addition to the comprehensive molecular phylogenetic analysis of cordyceps in the Hypocreales order. The present study aimed to expand our knowledge of the diversity of mitochondrial genomes and phylogenetic relationships of cordyceps.

2. Materials and Methods

2.1 Fungal materials and DNA extraction

O. xuefengensis strain HCMA001 was used in this study for mitogenomic analysis. *In vitro*mycelial fermentation in liquid culture was conducted in Erlenmeyer flasks with shaking at 150 rpm in an incubator at 25°C for 10 days. Mycelia were collected by centrifugation and quickly frozen in liquid nitrogen. Total genomic DNA was then extracted using an improved extraction method [40]. Genomic DNA was quantified using an TBS-380 fluorometer (Turner BioSystems Inc., Sunnyvale, CA). High quality DNA samples (OD$_{260/280}$=1.8~2.0, > 6 ug) were used to construct fragment libraries for genomic sequencing.

22 Mitochondrial DNA sequencing and genome assembly

About 8 μg of purified DNA was sequenced using a combination of third-generation sequencing technolo-

2

gies including PacBio RS and Illumina sequencing platforms. The Illumina data was used to evaluate the complexity of the genome and correct the PacBio long reads. The raw paired-end reads were trimmed, and quality controlled using Trimmomatic with parameters including SLIDINGWINDOW: 4:15 MINLEN: 75 (version 0.36*http://www.usadellab.org/cms/uploads/supplementary/Trimmomatic*). Clean data obtained from the above quality control procedures were then used for further analyses.

2.3 Annotation and comparative analysis of the O. xuefengensis mitogenome

Ab initio prediction was used to generate gene models for the *O. xuefengensis* mitogenome. Gene models were identified using MFannot (https://megasun.bch.umontreal.ca/cgi-bin/dev_mfa/mfannotInterface.pl). Gene models were then compared against the non-redundant (NR) NCBI database using BLASTp, in addition to comparison against the SwissProt, KEGG, and COG databases to establish functional annotation, also using BLASTp. tRNAs were identified using the tRNAscan-SE program (v1.23, http://lowelab.ucsc.edu/tRNAscan-SE) and rRNAs were identified using the RNAmmer program (v1.2,*http://www.cbs.dtu.dk/services/RNAmmer/*). Circular maps of the four mitochondrial genomes were drawn using OGDRAW (*https://chlorobox.mpimp-golm.mpg.de/OGDraw.htmlonline*) [41].

2.4 Sequence analysis

Nucleotide compositional skew was calculated according to the formulae: AT-skew = (A-T)/(A+T) and GC-skew = (G-C)/(G+C) [42]. The Sequence Manipulation Suite program package (http://www.bioinformatics.org/sms2/codon_usage.html) was used to analyze codon usage based on genetic code "4" [43]. Genetic distances between each pair of the 14 core PCGs (*atp6, atp8, atp9, cob, cox1, cox2, cox3, nad1, nad2, nad3, nad4, nad4L, nad5,* and*nad6*) were calculated with MEGA X [44], using the Kimura-2-parameter (K2P) substitution model. DnaSP v6 [45] was used to calculate the nonsynonymous (Ka) and synonymous (Ks) substitution rates for the 14 core PCGs among the four mitogenomes. Genomic synteny among the four mitogenomes was analyzed using the Mauve v2.4.0 program [46].

2.5 Repetitive element analysis

Tandem repeats (> 10 bp in length) in the four mitogenomes were detected using the Tandem Repeats Finder program [47], with default parameters. In addition, repeat sequences were identified using the REPuter program [48] to identify forward (direct), reverse, complemented, and palindromic (reverse complemented) repeats across the four mitogenomes, using a minimum repeat size of 30 and a hamming distance set to 3.

2.6 Phylogenetic analysis of Hypocrealean species

To determine the phylogenetic relationships of the cordyceps species among the Hypocreales order, 19 mitogenomes were downloaded from the NCBI database and subjected to Maximum Likelihood (ML) and Bayesian phylogenetic analysis using the aforementioned 14 PCGs and *rps3* . *Rhizopogon salebrosus* and *Rhizopogon vinicolor* were used as outgroups for the analyses (Table S1). Single mitochondrial genes were first aligned using MAFFT v7.037 [49] and these alignments were then concatenated using SequenceMatrix v1.7.8 (*http://www.softpedia.com/get/Science-CAD/Sequence-Matrix.shtml*) [50]. The ML tree was then calculated for the combined gene set using RAxML v8.0.0 [51], while Bayesian inference (BI) analysis was performed with MrBayes v3.2.6 [52]. Other analyses and evaluation methods were conducted based on previously described protocols [9].

2.7 Data availability statement

The newly sequenced mitogenome of *O. xuefengensis* strain HCMA001 was submitted to GenBank under the accession number SAMN16236787.

3. Results

3.1. Characterization of the four cordyceps mitogenomes

Comparative analysis indicated that the complete mitogenomes of the four cordyceps all comprised circular DNA molecules. The mitogenome length of*O. xuefengensis* (78,744 bp) was much smaller than that of*O.*

sinensis (157,559 bp), but longer than those of *Cordyceps militaris* (33,277 bp) and *C. brongniartii* (33,926 bp). All four mitogenomes encoded 14 core PCGs involved in energy metabolism, *rps3* involved in translation, and 2 rRNAs. The PCG lengths varied, but the protein coding regions themselves were conserved. The tRNA numbers of the four mitogenomes ranged from 25 (*C. brongniartii*) to 27 (*O. sinensis*). The tRNAs comprised those for all 20 amino acids and ranged in size from 71 to 86 nucleotides (Table S1). The nucleotide composition of the *O. xuefengensis* mitogenome was identical to that of *O. sinensis* , but not to *C. militaris* and *C. brongniartii* . The *O. sinensis* mitogenome had the highest GC content (30.20%), followed by those of *O. xuefengensis*(29.94%), *C. brongniartii* (27.34%), and *C. militaris*(26.79%). The high frequency of A and T usage in codons contributes to the high AT content of the four mitochondrial genomes, with AT content reaching 70.06% in *O. xuefengensis* , 69.80% in *O. sinensis* , 73.21% in *C. militaris* , and 72.66% in *C. brongniartii* . Both GC skew and AT skew were positive for all four mitogenomes (Table 1).

The *O. sinensis* mitogenome exhibited an unusually enlarged size (157,559 bp). The compositions of regions were thus analyzed for the mitogenomes, indicating that the protein coding regions among the four mitogenomes are basically identical. A significant difference was observed when comparing intronic and intergenic regions, wherein the total number of intronic nucleotides of the *O. sinensis* mitogenome was much higher than for those of the other cordyceps, and the total number of intergenic nucleotides of the *O. xuefengensis* mitogenome is the highest among the four mitogenomes (Fig 2).

The number of introns also significantly differed among the four mitogenomes. Fifty-four introns were present in the mitogenome of *O. sinensis* , 18 in that of *O. xuefengensis* , 8 in that of *C. militaris,* and 6 in that of *C. brongniartii* (Table. 1). Further, the host genes of introns also differ among the four cordyceps. Introns are distributed in the *cob* (no. of introns: 1),*cox1* (1), *cox2* (1), *cox3* (1), and *rnl* (4) genes of the *C. militaris* mitogenome, and in the *cob* (1),*cox1* (2), *nad1* (1), and *rnl* (2) genes of the *C. brongniartii* mitogenome. In *O. xuefengensis* , the introns are distributed in *cob* (2), *cox1* (3), *cox2*(1), *cox3* (3), *nad1* (2), *nad2* (2), *nad5* (1),*rnl* (3), and *rns* (1) genes, while in the *O. sinensis* mitogenome, the introns are distributed in the *cob* (6), *cox1* (14),*cox2* (6), *cox3* (2), *nad1* (3), *nad2* (2),*nad4L* (1), *nad5* (5), *nad6* (1), *atp6* (2),*atp9* (1), *rnl* (8), and *rns* (1) genes (Table S2). The most common intron type among the mitogenomes are group IB introns. Introns also harbor 0–2 homing endonuclease genes, including the LAGLIDADG homing endonuclease and the GIY-YIG endonuclease.

Table 1. Characteristics of cordyceps mitogenomes.

Characteristic	*Ophiocordyceps xuefengensis*	*Ophiocordyceps sinensis*	*Cordyceps militaris*	*Cordyceps brongni*
Size (bp)	78,744	157,559	33,277	33,926
A (%)	36.08	36.70	36.98	36.51
T (%)	33.98	33.10	36.23	36.15
G (%)	16.33	15.70	15.20	15.13
C (%)	13.61	13.70	11.59	12.21
Rate of GC (%)	29.94	30.20	26.79	27.34
Rate of AT (%)	70.06	69.80	73.21	72.66
AT-skew	0.03	0.05	0.01	0.005
GC-skew	0.09	0.09	0.14	0.11
PCGs (n)	14	14	14	14
Introns (n)	18	54	8	6
Intronic ORFs (n)	16	45	8	5
GIY-YIG (n)	4	10	4	4
LAGLIDADG (n)	9	27	3	1
rRNAs (n)	1	1	1	1
nRNAl (n)	1	1	1	1
tRNAs (n)	26	27	26	25
PCG regions (bp)	34,994	104,257	17,735	18,329
Intronic regions (bp)[a]	21,146	107,288	9,812	8,452

4

Characteristic	*Ophiocordyceps xuefengensis*	*Ophiocordyceps sinensis*	*Cordyceps militaris*	*Cordyceps brongni*
Exon regions (bp)[b]	13,848	13,287	12,978	13,098
tRNA genes (bp)	1,934	2,010	1,924	1,867
rRNA genes (bp)[c]	7,225	7,144	4,666	4,801
Intergenic regions (bp)	34,591	27,831	3,797	5,708

[a]: Introns in PCGs and rRNA; [b]: Exons in PCGs; [c]: *rnl* and *rns* including exons.

3.2. Protein-coding genes and codon usage among the four mitogenomes

Most PCGs and *rps3* are initiated with the ATG codon, whereas*cox1* and *nad6* in the *O. xuefengensis* mitogenome are initiated with the TTG codon. Further, *cox3* in the *O. sinensis* mitogenome and *atp9* in both *O. xuefengensis* and*O. sinensis* mitogenomes are initiated with the GTG codon. In addition, most PCGs and *rps3* are terminated with the TAA codon, but *cox1* , *rps3,* and *cox3* in the *O. sinensis*mitogenome, and *cox1* and *cob* of the *O. xuefengensis*mitogenome are terminated with the TAG codon (Table S3). Comparative analysis indicated that all the PCGs and *rps3* encode 20 amino acids, with similar amino acid compositions among the cordyceps mitogenomes. The most overused amino acid is Leu (ranging in prevalence from 628 to 652), whereas the least used amino acid is Cys (32 – 48 residues). The usage of Asn, Lys, Asp, and Glu in the *O. xuefengensis*mitogenome is much higher than in those of the other three cordyceps (Fig. 3, Table S4). The relative synonymous codon usage (RSCU) of the 14 PCGs and *rps3* in the four cordyceps mitogenomes was also analyzed. A total of 4,616, 4,429, 4,326 and 4,366 codons were observed among the PCGs and *rps3* of the *O. xuefengensis, O. sinensis, C. militaris,* and *C. brongniartii* mitogenomes, respectively. The most used codon among all four mitogenomes is UUA, but the most frequently used codon is AGA in the *O. xuefengensis* and*O. sinensis* mitogenomes (Fig. 4, Table S5).

3.3. Variation, genetic distance, and evolutionary rates of common genes

The lengths and/or GC contents of the 14 core PCGs and rps3 are not all consistent across the four mitogenomes, with the exception of atp8, atp9, and nad4L PCGs. In particular, the length of nad5 in O. xuefengensis is about 1,000 bp longer than in the other three cordyceps mitogenomes. Further, the GC content of all 14 PCGs and rps3 differ, indicating that the core PCGs and rps3 have been modified within different cordyceps. Across all PCGs in the mitogenomes, GC content is highest in atp9 within all four cordyceps mitogenomes, and lowest in atp8 within the O. xuefengnensis and O. sinensis mitogenomes, but lowest in rps3 within the C. militaris and C. brongniartii mitogenomes. AT skew is negative in most PCGs except rps3, while GC skew is present in most PCGs except atp8. In particular, the AT skew of atp6 in the C. militaris mitogenome and nad6 in the O. xuefengnensis mitogenome are negative, while the GC skew of only rps3 in the C. brongniartii mitogenome is negative (Fig. 5).

Across all 14 PCGs and rps3 genes that were evaluated, rps3 has the greatest K2P genetic distance among the mitogenomes, followed by cox1. atp8 has the least genetic distance among the genes investigated, indicating a high degree of conservation. Ka analysis indicated that rps3 and cox1 have relatively high nonsynonymous substitution rates. Further, the Ks of nad1 is the highest across all four mitogenomes, while that of atp8 is the lowest. The Ka/Ks values for all 14 PCGs and rps3 are < 1, suggesting that these genes are subject to purifying selection (Fig. 6, Table S6).

3.4. Repetitive elements

A total of 18 repeat sequences were identified in the mitogenome of O. xuefengensis, 250 in that of O. sinensis, 6 in that of C. militaris and 23 in that of C. brongniartii (Table S7). The length of repeat sequences ranges from 38 bp to 372 bp, with pairwise nuleotide similarities ranging from 74.73% to 100%. The largest repeat region was observed in the O. sinensis mitogenome, within ORF322 and ORF308. The largest repeat region in the O. xuefengensis mitogenome is 220 bp long and located in the intronic region of cox2 and ORF301, while the largest repeat region in the C. militaris mitogenome is 259 bp long and located in the intronic

5

region of rnl. Lastly, the largest repeat region in the C. brongniartii mitogenome is 191 bp long and located between cox1 and trnR.

Seven tandem repeats were detected in both the O. xuefengensis and C. brongniartii mitogenomes, 43 in the O. sinensis mitogenome, and 5 in the C. militaris mitogenome (Table S8). The longest tandem sequence is found in the O. sinensis mitogenome, comprises 123 bp, and is located in the in the ORF138 coding region. Of all tandem repeats in the four mitogenomes, most exist in one or two copies. REPuter identified forward (F) and palindromic (P) repeats in the mitogenome of each cordyceps (Table S9), including 69 F and 34 P repeates in the mitogenome of O. xuefengensis that account for 10.16% of the total mitogenome, along with 16 F and 1 P repeats in the C. militaris mitogenome that account for 4.95% of its total mitogenome. In addition, 38 F and 19 P repeats were identified in the C. brongniartii mitogenome that account for 14.54% of its total mitogenome, and a total of 754 F, 538 P, 31 C, and 60 R repeats were identified in the O. sinensis mitogenome that account for 65.55% of its total mitogenome.

3.5. Gene rearrangements and phylogenetic analysis

The relative positions of the mitochondrial genes (including rnl, rns, rps3, the 14 core PCGs, and the tRNAs) are highly conserved in the four cordycep mitogenomes. The O. xuefengensis mitogenome only lacks trnI, consistent with that of O. sinensis (Fig. 7). Genome synteny analysis also identified several instances of gene rearrangement in the mitogenomes of the four cordyceps. All four cordyceps mitogenomes could be divided into two homologous regions, with the sizes and relative positions of these homologous regions substantially differing among the four species. Mitogenomic rearrangements were not observed among the four species (Fig. 8).

A complete mitogenome was produced for O. xuefengensis in this study and used with 19 other complete mitogenome sequences from six families for phylogenetic reconstruction to further investigate the phylogenetic position of O. xuefengensis within the Hypocreales order. Species and NCBI accession numbers used in these phylogenetic analyses are listed in Supplementary Table S10. Bayesian inference (BI) based on mitochondrial gene datasets using the GTR + I + G nucleotide substitution model yielded an identical and well-supported topology in which all major clades are well supported. The 20 fungal species comprised six major clusters, corresponding to the major groups of the orders Hypocreales and Agaricomycetes. The four cordyceps species were divided into two groups, wherein the close relationship between O. xuefengensis and O. sinensis was highly supported (Fig. 9).

4. Discussion

Cordyceps are fungal species used in traditional medicine that are represented by the well-known fungal resource, O. sinensis.However, natural populations of O. sinensis have been over harvested to the extent that it is an endangered species. It is consequently necessary to explore alternative cordyceps resources, because artificial cultivation of O. sinensis is difficult to achieve [15]. O. xuefengensis is a new potential cordyceps resource that was found on Xuefeng mountain in Hunan. Investigation of its artificial cultivation and biometabolite biosynthesis yielded some positive results [14]. Several previous studies have investigated the phylogenetic relationships of cordyceps, although detailed phylogenetic relationships for the species remain poorly understood, especially since increasing numbers of related species have been reported [16-18]. It is difficult to accurately classify fungal species based on limited morphological characters and overlapping morphological features [19]. Consequently, mitogenomes have been widely used in the phylogenetic analysis of eukaryotes due to many advantages, including uniparental inheritance, rapid evolutionary rates, and the presence of several molecular markers [20]. Here, we report a new mitogenome sequence for O. xuefengensis that could provide new insights into the evolutionary histories of cordyceps fungal mitogenomes and the taxonomic relationships among cordyceps.

Mitogenome sizes significantly differ among fungi, ranging from approximately 11.09 kbp (Hanseniaspora uvarum) [21] to 272.2 kbp (Morchella importuna) [22] in length. The variation in size of mitogenomes among different fungal species primarily results from the presence and extent of intergenic regions and introns [23]. The Ascomycota cordyceps fungus O. sinensis contained a mitogenome with introns accounting

6

for 68.09% of the whole mitogenome, while introns in the mitogenomes of *O. xuefengensis*, *C. militaris*, and *C. brongniartii* only accounted for 26.85%, 29.49% and 24.91% of their mitogenomes, respectively. Further, intergenic regions within the *O. xuefengensis* mitogenome accounted for 43.93% of the whole mitogenome, but only 11.41%-17.66% of the mitogenomes of the other three cordyceps species. Thus, considerable variation in the size of introns and intergenic regions are the major reasons underlying the significant differences in the mitogenome sizes of *O. sinensis* and *O. xuefengensis*. In contrast, *C. militaris* and *C. brongniartii* exhibited similarly sized mitogenomes, due to similar numbers and sizes of intergenic regions and introns.

However, the large difference in introns did not affect the taxonomically close relationship between *O. sinensis* and *O. xuefengensis*. Maximum Likelihood and Bayesian phylogenetic analyses were constructed using 14 PCGs and *rps3* from 20 fungi to reveal the taxonomic relationships of these fungi. The phylogenetic analysis highly supported the close relationship between *O. xuefengensis* and *O. sinensis*. This apparent domestication of ancestral introns may indicate an adaptation to their host genome that may be due to their decreased mobility [24], or alternatively, to their potential important role in stabilizing the gene that hosts the intron [25, 26]. Other introns may be acquired late in the divergence of taxa, either through HGT events or through active transposition [27, 28]. The transposition of some introns to other genomic regions with less sequence similarity can occur more frequently under stress-induced conditions [29, 30]. A recent study demonstrated that the number and Pcl of introns are highly variable between two *Rhizopogon* species [31]. Further, the introns in nuclear genes of *S. cerevisiae* play crucial roles in the survival of the organism under starvation conditions [32]. Consequently, the abundance of introns in the mitochondrial genome of *O. sinensis* relative to other cordyceps species may be a consequence of the particular environments they inhabit.

The accumulation of repeated evolutionary events in fungal mitogenomes could lead to over-dispersal of repeat sequences and the introduction of new genes through HGT, thereby contributing to dynamic changes in genome structure and gene order [33]. Moreover, the accumulation of repeat sequences could also be highly related to gene recombination and gene loss [34]. A variety of models have been proposed to investigate these mitochondrial gene rearrangements [35]. Here, the mitogenome of *O. sinensis* harbored a larger proportion of repeat sequences than the mitogenomes of *O. xuefengensis*, *C. militaris* or *C. brongniartii*. Interestingly, collinearity analysis revealed that the mitogenomic gene order in these four fungi was not highly variable. In addition, positive selection of core genes was not observed by Ka/Ks analysis. This differs from observations in plants, wherein mitochondrial gene order is highly variable due to high rates of recombination [33]. These observations collectively indicate that selection upon the genomes by environmental factors is relatively weak among the four cordyceps fungi evaluated here, resulting in highly conserved phylogenetic relationships. Thus, we could also infer that differentiation of the cordyceps genomes remain in early stages, consistent with the molecular phylogenetic analyses. One reasonable explanation for these contradictory results is that the larger proportion of repeat sequences in the genome of *O. sinensis* exists in intronic regions, but not in intergenic regions, leading to little influence on their evolution.

Mitochondrial genomes are obtained from a common ancestor and have been widely used in population genetics and evolutionary studies [36, 37]. Different barcode genes, including ITS, LSU, RPB2, and EF1α, have been used for phylogenetic analysis [9, 38]. However, the large number of available molecular markers in mitogenomes and their independent evolutionary histories make them attractive tools for reconstructing phylogenetic relationships [39]. Here, we reconstructed a well-resolved phylogeny based on the combined alignment of 14 PCGs and the *rps3* gene that separated 20 fungal species into major clades. In addition, K2P and Ka/Ks analysis indicated that *cox1* and *rps3* divergence play important roles in population differentiation. Given that reliable genetic molecular markers are critical for further understanding the phylogenetics and classification of species, these two genes could be further evaluated as barcodes for fungal species identification.

5. Conclusions

This study extends the known mitogenomes of cordyceps fungi and identified variation in PCGs, tRNA genes, and rRNA genes among medicinal cordyceps species, including *O. xuefengensis*. Comparative mitochondrial

7

genome analysis of four cordyceps fungi were conducted, revealing that their mitochondrial genomes are conserved and that differentiation of their mitogenomes is still in the early stages. Nevertheless, results indicated that the *cox1* and *rps3* genes likely play important roles in population differentiation among the taxa. Mitogenome analyses, like those reported here, will allow further study of the population genetics, taxonomic relationships, and evolutionary biology of medicinally important cordyceps species.

Supplementary Materials: The following materials are available online. Table S1: General tRNA features in the four cordyceps mitogenomes. Table S2: The number of introns and their host genes. Table S3: Complete mitochondrial genome characteristics for the four cordyceps analyzed here. Table S4: Amino acid composition of mitogenomes among the four cordyceps species. Table S5: The RSCU of 14 PCGs and *rps3* in the four cordyceps mitochondrial genomes. Table S6: Genetic analysis of 14 core PCGs and *rps3* among the four cordyceps genomes. Table S7: Local BLAST analysis of the four cordyceps mitogenomes against themselves. Table S8: Tandem repeats detected in the mitogenomes of cordyceps using the Tandem Repeats Finder program. Table S9: Distribution of repeat loci in the four cordyceps mitogenomes identified by REPuter. Table S10: The GenBank accessions for genomes used in the phylogenetic analysis.

Author Contributions: Conceptualization: C.Z., S.X., and. S.Z. Methodology: J.J., C.Z., R.Z., and H.L. Software analysis: C.Z., J.J., and J.X. Investigation: R.Z. Resources: C.Z. and H.L. Data curation: C.Z., J.J., and J.X. Writing (original draft preparation): C.Z. Writing (review and editing): S.Z., J.J., and S.X. Visualization: J.X. Project administration: D.W. Funding acquisition: S.Z. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Acknowledgment

Fig. 1. Circular maps of the mitochondrial genomes of four cordyceps species. Genes are represented by different colored blocks. Colored blocks outside each ring indicate that the genes are on the direct strand, while colored blocks within the ring indicate that the genes are located on the reverse strand.

Fig. 2. The abundances of protein-coding, intronic, intergenic, and RNA gene regions (rRNAs and tRNAs) of the entire mitochondrial genomes from four cordyceps species.

Fig. 3. Codon usage in the mitochondrial genomes of the four cordyceps species. Codon numbers are plotted on the y-axis for A, Ophiocordyceps xuefengensis; B, Ophiocordyceps sinensis; C, Cordyceps militaris; D, Cordyceps brongniartii.

Fig. 4. RSCU values for 14 PCGs and *rps3* in the mitochondrial genomes of four cordyceps.

Fig. 5. Variation in the length and base composition of each of the 14 PCGs and rps3 among four cordyceps mitogenomes. A, PCG length variation; B, GC content across PCGs; C, AT skew; D, GC skew.

Fig. 6. Genetic analysis of 14 PCGs and *rps3* across four cordyceps mitogenomes. K2P, the overall mean Kimura-2-Parameter distance; Ka, the mean number of nonsynonymous substitutions per nonsynonymous site; Ks, the mean number of synonymous substitutions per synonymous site.

Fig. 7. Gene order comparison among four cordyceps mitogenomes.

Fig. 8. Colinearity analysis of four cordyceps mitogenomes, generated with Mauve 2.4.0. Two homologous regions were detected across the four mitochondrial genomes overall.

Fig. 9. Molecular phylogeny of 20 fungal species based on Bayesian inference (BI) and Maximum likelihood (ML) analysis of 14 PCGs and *rps3* genes. Support values are Bayesian posterior probabilities (before slash) and bootstrap (BS) values (after slash). Species and NCBI accession numbers for genomes used in the phylogenetic analyses are provided in Supplementary Table S10.

1. Cameron SL, Lambkin CL, Barker SC, Whiting MF, Entomology JS:**A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision** . *Systematic Entomology* 2007,**32** (1):40 - 59.

2. Song F, Li H, Liu G-H, Wang W, James P, Colwell DD, Tran A, Gong S, Cai W, Biology RSJS: **Mitochondrial genome fragmentation unites the parasitic lice of eutherian mammals** . *Systematic Biology*2019, **68** (3):430-440.

3. Saccone C, Gissi C, Lanave C, Larizza A, Reyes AJG: **Evolution of the mitochondrial genetic system: an overview** . *Genes* 2000,**261** (1):153-159.

4. Liu Y, Song F, Jiang P, Wilson JJ, Cai W, Li HJMP, Evolution:**Compositional heterogeneity in true bug mitochondrial phylogenomics** . *Molecular Phylogenetics Evolution* 2018,**118** :135-144.

5. Curole JP, Kocher TDJTiE, Evolution: **Mitogenomics: digging deeper with complete mitochondrial genomes** . *Trends in Ecology Evolution* 1999, **14** (10):394-398.

6. Pramateftaki PV, Kouvelis VN, Panagiotis L, Typas MAJFYR: **The mitochondrial genome of the wine yeast Hanseniaspora uvarum: a unique genome organization among yeast/fungal counterparts** . *FEMS Yeast Research* 2006, **6** (1):77-90.

7. Sankoff D, Leduc G, Antoine N, Paquin B, Cedergren LR: **Gene order comparisons for phylogenetic inference: Evolution of the mitochondrial genome** . *Proceedings of the National Academy of Sciences of the United States of America* 1992,**89** (14):6575-6579.

8. Poliseno A, Feregrino C, Sartoretto S, Aurelle D, Wrheide G, Mcfadden CS, Vargas SJMP, Evolution: **Comparative mitogenomics, phylogeny and evolutionary history of Leptogorgia (Gorgoniidae)** . *Molecular Phylogenetics Evolution* 2017, **115** :181-189.

9. Qiang L, Min L, Mei Y, Chuan X, Xin J, Zuqin C, Wenli HJIJoBM:**Characterization of the mitochondrial genomes of three species in the ectomycorrhizal genus Cantharellus and phylogeny of Agaricomycetes** . *International Journal of Biological Macromolecules* 2018, **118** :756-769.

10. Zhong C, Jin J, Liu H, Cai Y, Qin Y, Xie J, Wang W, Qin Y, Huang H, Zhang S: **Research status quo, problems and prospects of***Ophiocordyceps xuefengensis* . *Journal of Microbiology*2019, **39** (4):107-114.

11. Yue K, Ye M, Lin X, Zhou ZJIJoMM: **The artificial cultivation of medicinal Caterpillar Fungus, Ophiocordyceps sinensis (Ascomycetes): a review** . *International Journal of Medicinal Mushrooms* 2013,**15** (5):425-434.

12. Wen TC, Zhu RC, Kang JC, Huang MH, Tan DB, Ariyawansha H, Kevin D, Liu H: *Ophiocordyceps xuefengensis* **sp. nov. from larvae of** *Phassus nodus* **(Hepialidae) in Hunan Province, southern China** .*Phytotaxa* 2013, **123** (1):41-50.

13. Jin J, Qin Y, Zhong C, Zhou R, Xie J, Liu H, Xiao J, Cai P, Zhang S, Qin Y: **Differential gene expression and associated metabolite accumulation in fungus** *Ophiocordyceps xuefengensis* **cultivated under breathable and airtight conditions** . *Mycoscience* 2019,**60** (5):281-286.

14. Jin J, Zhong C, Qin Y, Cai Y, Zhen L, Shen B, Chen L, Wan D, Qin Y, Zhang S: **A new cordycepin-producing caterpillar fungus***Ophiocordyceps xuefengensis* **with artificial infection to the host,**

cultivation of mycelia and stromata . *FEMS Microbiology Letters* 2017, **364** (20):1-8.

15. Zhang J, Wang P, Wei X, Li L, Cheng H, Wu Y, Zeng W, Yu H, Chen YJFRI: **A metabolomics approach for authentication of Ophiocordyceps sinensis by liquid chromatography coupled with quadrupole time-of-flight mass spectrometry** . *Food Research International*2015, **76** (OCT.PT.3):489-497.

16. Zhang Y, Zhang S, Zhang G, Liu X, Wang C, Xu J: **Comparison of mitochondrial genomes provides insights into intron dynamics and evolution in the caterpillar fungus Cordyceps militaris %J Fungal Genetics & Biology Fg & B** . *Fungal Genetics and Biology* 2015, **77** :95-107.

17. Ghikas DV, Kouvelis VN, Typas MAJBM: **Phylogenetic and biogeographic implications inferred by mitochondrial intergenic region analyses and ITS1-5.8S-ITS2 of the entomopathogenic fungi Beauveria bassiana and B. brongniartii** . *BMC Microbiology* 2010, **10** (1):174-188.

18. Li Y, Hu XD, Yang RH, Hsiang T, Wang K, Liang DQ, Liang F, Cao DM, Zhou F, Wen GJeR: **Complete mitochondrial genome of the medicinal fungus Ophiocordyceps sinensis** . *J entific Reports* 2015, **5** :13892-13900.

19. Li Q, Wang Q, Jin X, Chen Z, Xiong C, Li P, Zhao J, Huang WJIJoBM:**Characterization and comparison of the mitochondrial genomes from two Lyophyllum fungal species and insights into phylogeny of Agaricomycetes** . *International Journal of Biological Macromolecules* 2019, **121** :364-372.

20. Boore JLJNAR: **Animal mitochondrial genomes** . *Nucleic Acids Research* 1999, **27** (8):1767-1780.

21. Pramateftaki PV, Kouvelis VN, Panagiotis L, Typas MA: **The mitochondrial genome of the wine yeast *Hanseniaspora uvarum*: a unique genome organization among yeast/fungal counterparts** . *FEMS Yeast Research* 2006, **6** (1):77-90.

22. Liu W, Cai Y, Zhang Q, Chen L, Shu F, Ma X, Bian Y: **The mitochondrial genome of *Morchella importuna* (272.2kb) is the largest among fungi and contains numerous introns, mitochondrial non-conserved open reading frames and repetitive sequences** .*International Journal of Biological Macromolecules* 2020, **143** :373-381.

23. Lang BF, Laforest M-Je, Burger G: **Mitochondrial introns: a critical view** . *TRENDS in Genetics* 2007, **23** (3):119-125.

24. Novikova O, Belfort M: **Mobile Group II Introns as Ancestral Eukaryotic Elements** . *Trends in genetics : TIG* 2017, **33** (11):773-783.

25. Korovesi AG, Ntertilis M, Kouvelis VNJMP, Evolution: **Mt-rps3 is an ancient gene which provides insight into the evolution of fungal mitochondrial genomes** . *Molecular Phylogenetics Evolution* 2018, **127** :74-86.

26. Genetics BSJC: **Genetic conservation versus variability in mitochondria: the architecture of the mitochondrial genome in the petite-negative yeast Schizosaccharomyces pombe** . *Current Genetics* 2003, **43** (5):311-326.

27. Baojun, Wu, Adnan, Buljic, Weilong, biology HJM, evolution:**Extensive Horizontal Transfer and Homologous Recombination Generate Highly Chimeric Mitochondrial Genomes in Yeast** .*Molecular biology evolution* 2015, **32** :2559–2570.

28. Mardanov AV, Beletsky AV, Kadnikov VV, Ignatov AN, Ravin NV, Palazzo AFJPO: **The 203 kbp Mitochondrial Genome of the Phytopathogenic Fungus Sclerotinia borealis Reveals Multiple Invasions of Introns and Genomic Duplications** . *Plos One* 2014, **9** (9):e107536.

29. Robbins JB, Smith D, Belfort M: **Redox-responsive zinc finger fidelity switch in homing endonuclease and intron promiscuity in oxidative stress** . *Curr Biol* 2011, **21** :243-248.

30. Coros CJ, Piazza CL, Chalamcharla VR, Smith D, Belfort M:**Global Regulators Orchestrate Group II Intron Retromobility** .*Molecular cell* 2009, **34** (2):250-256.

31. Li Q, Ren Y, Shi X, Peng L, Zhao J, Song Y, Zhao GJIJoMS:**Comparative Mitochondrial Genome Analysis of Two Ectomycorrhizal Fungi (Rhizopogon) Reveals Dynamic Changes of Intron and Phylogenetic Relationships of the Subphylum Agaricomycotina** . *International Journal of Molecular Sciences* 2019, **20** (20):5167.

32. Parenteau J, Maignon L, Berthoumieux M, Catala M, Gagnon V, Abou Elela SJN: **Introns are mediators of cell response to starvation** . *Nature* 2019, **565** (7741):612-617.

33. Aguileta G, de Vienne DM, Ross ON, Hood ME, Giraud T, Petit E, Gabaldón T: **High Variability of Mitochondrial Gene Order among Fungi** . *Genome Biol Evol* 2014, **6** (2):451-465.

34. Zou H, Jakovlić I, Chen R, Zhang D, Zhang J, Li W-X, Wang G-T:**The complete mitochondrial genome of parasitic nematode Camallanus cotti: extreme discontinuity in the rate of mitogenomic architecture evolution within the Chromadorea class** . *BMC Genomics* 2017, **18** (1):840.

35. Xia Y, Zheng Y, Murphy RW, Zeng X: **Intraspecific rearrangement of mitochondrial genome suggests the prevalence of the tandem duplication-random loss (TDLR) mechanism in Quasipaa boulengeri** .*BMC Genomics* 2016, **17** (1):965.

36. Al-Reedy RM, Malireddy R, Dillman CB, Kennell JC:**Comparative analysis of Fusarium mitochondrial genomes reveals a highly variable region that encodes an exceptionally large open reading frame** . *Fungal Genetics* 2012, **49** (1):2-14.

37. Grosemans T, Morris K, Thomas WK, Rigaux A, Moens T, Derycke S:**Mitogenomics reveals high synteny and long evolutionary histories of sympatric cryptic nematode species** . *Ecology Evolution* 2016, **6** (6):1854-1870.

38. Buyck B, Kauff F, Eyssartier G, Couloux A, Hofstetter V: **A multilocus phylogeny for worldwide Cantharellus (Cantharellales, Agaricomycetidae)** . *Fungal Diversity* 2014,**64** (1):101-121.

39. Jiang L, Zhao L, Cheng D, Zhu L, Zhang M, Ruan Q, Chen W:**The complete mitochondrial genome sequence of the Sichuan Digging Frog, Kaloula rugifera (Anura: Microhylidae) and its phylogenetic implications** . *Gene* 2017, **626** :367-375.

40. Chen J, Guan R, Chang S, Du T, Zhang H, Xing HJPO:**Substoichiometrically Different Mitotypes Coexist in Mitochondrial Genomes of Brassica napus L** . *PLoS ONE* 2011,**6** (3):e17662.

41. Conant GC, Wolfe KH: **GenomeVx: simple web-based creation of editable circular chromosome maps** . *Bioinformatics* 2008,**24** (6):861-862.

42. Perna NT, Kocher TD: **Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes** .*Journal of Molecular Evolution* 1995, **41** (3):353-358.

43. Stothard: **The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences** .*BioTechniques* 2000, **28** (6):1102-1104.

44. Joseph C: **MEGA evolutionary software Re-engineered to handle today's big data demands** . *Molecular Biology & Evolution* 2016,**33** (7):1887-1888.

45. Rozas J, Ferrer-Mata A, Carlos J, Sara S-D, Pablo G-R: **DnaSP 6: DNA sequence polymorphism analysis of large datasets** .*Molecular Biology Evolution* 2017(34):3299-3302.

46. Darling AC, Mau B, Blattner FR, Perna NT: **Mauve: multiple alignment of conserved genomic sequence with rearrangements** .*Genome research* 2004, **14** (7):1394-1403.

47. Benson G: **Tandem repeats finder: a program to analyze DNA sequences** . *Nucleic Acids Res* 1999, **27** (2):573-580.

48. Stefan K, Choudhuri JV, Enno O, Chris S, Jens S, Robert G:**REPuter: the manifold applications of repeat analysis on a genomic scale** . *Nucleic Acids Research* 2001(22):4633-4642.

49. Katoh K, Rozewicki J, Yamada K: **MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization** . *Briefings in Bioinformatics* 2019,**20** (4):1160-1166.

50. Vaidya G, Lohman DJ, Meier R: **SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information** . *Cladistics-the International Journal of the Willi Hennig Society* 2011, **27** (2):171–180.

51. Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies** . *Bioinformatics*2014, **30** (9):1312-1313.

52. Huelsenbeck JP: **MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space** . *Systematic Biology* 2012, **61** (3):539-542.