frontiers in Microbiology

ORIGINAL RESEARCH
published: 27 July 2017
doi: 10.3389/fmicb.2017.01422

SMRT Sequencing Revealed Mitogenome Characteristics and Mitogenome-Wide DNA Modification Pattern in *Ophiocordyceps sinensis*

*Xincong Kang[1,2,3†], Liqin Hu[2,3†], Pengyuan Shen[2,3], Rui Li[4] and Dongbo Liu[1,2,3,5]**

[1] Hunan Provincial Key Laboratory of Crop Germplasm Innovation and Utilization, Hunan Agricultural University, Changsha, China, [2] Horticulture and Landscape College, Hunan Agricultural University, Changsha, China, [3] State Key Laboratory of Subhealth Intervention Technology, Changsha, China, [4] Nextomics Biosciences, Wuhan, China, [5] Hunan Co-Innovation Center for Utilization of Botanical Functional Ingredients, Changsha, China

Single molecule, real-time (SMRT) sequencing was used to characterize mitochondrial (mt) genome of *Ophiocordyceps sinensis* and to analyze the mt genome-wide pattern of epigenetic DNA modification. The complete mt genome of *O. sinensis*, with a size of 157,539 bp, is the fourth largest Ascomycota mt genome sequenced to date. It contained 14 conserved protein-coding genes (PCGs), 1 intronic protein *rps3*, 27 tRNAs and 2 rRNA subunits, which are common characteristics of the known mt genomes in Hypocreales. A phylogenetic tree inferred from 14 PCGs in Pezizomycotina fungi supports *O. sinensis* as most closely related to *Hirsutella rhossiliensis* in Ophiocordycipitaceae. A total of 36 sequence sites in *rps3* were under positive selection, with dN/dS >1 in the 20 compared fungi. Among them, 16 sites were statistically significant. In addition, the mt genome-wide base modification pattern of *O. sinensis* was determined in this study, especially DNA methylation. The methylations were located in coding and uncoding regions of mt PCGs in *O. sinensis*, and might be closely related to the expression of PCGs or the binding affinity of transcription factor A to mtDNA. Consequently, these methylations may affect the enzymatic activity of oxidative phosphorylation and then the mt respiratory rate; or they may influence mt biogenesis. Therefore, methylations in the mitogenome of *O. sinensis* might be a genetic feature to adapt to the cold and low PO_2 environment at high altitude, where *O. sinensis* is endemic. This is the first report on epigenetic modifications in a fungal mt genome.

Keywords: ***Ophiocordyceps sinensis*, mitochondrial genome, characteristics, mitochondrial epigenetics, methylation, *rps3***

OPEN ACCESS

Edited by:
Orazio Romeo, University of Messina, Italy

Reviewed by:
Balázs Brankovics, Westerdijk Fungal Biodiversity Institute, Netherlands
Yongjie Zhang, Shanxi University, China

***Correspondence:**
Dongbo Liu
chinasaga@163.com

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Fungi and Their Interactions, a section of the journal Frontiers in Microbiology

Received: *31 March 2017*
Accepted: *13 July 2017*
Published: *27 July 2017*

Citation:
Kang X, Hu L, Shen P, Li R and Liu D (2017) SMRT Sequencing Revealed Mitogenome Characteristics and Mitogenome-Wide DNA Modification Pattern in Ophiocordyceps sinensis. Front. Microbiol. 8:1422. doi: 10.3389/fmicb.2017.01422

INTRODUCTION

Ophiocordyceps sinensis (syn. *Cordyceps sinensis*) is an entomopathogenic fungus that infects larvae of Hepialidae ghost moths to form a parasitic complex called "DongChongXiaCao" in Chinese (Commission, 2015). "DongChongXiaCao" is a Traditional Chinese Medicine treatment that has been used for 1000s of years for respiratory, renal, liver and cardiovascular diseases, hyposexuality

Abbreviations: BI, Bayesian inference; dN, non-synonymous substitution; dS, synonymous substitution; HE, homing endonuclease; ML, maximum likelihood; modQV, modification quality values; NJ, neighbor-joining; OXPHOS, oxidative phosphorylation; PCGs, protein-coding genes; RT, reverse transcriptase; SMRT, single molecule, real-time; SSRs, simple sequence repeats.

and hyperlipidemia (Zhu et al., 1998; Yue et al., 2013). *O. sinensis* is endemic to alpine regions on the Tibetan Plateau, with 3000 m as the lowest altitude for the distribution (Li et al., 2011). This fungus is rare in natural resources because of its strict host-specificity, limited geographical distribution and over-exploitation in recent decades (Li et al., 2011).

The mitochondrion is a cellular organelle that is required for respiratory metabolism and ATP production. In the mitochondrion, ATP synthesis is catalyzed by five mitochondrial inner membrane-bound enzyme complexes (Complexes I–V) (Saraste, 1999; Chandel and Schumacker, 2000). Complexes I–V are NADH-ubiquinol oxidoreductase, succinate-ubiquinol oxidoreductase, ubiquinol-cytochrome C oxidoreductase, cytochrome C oxidase, and ATP synthase, respectively. The mt genome encodes for seven subunits of Complex I (NAD1-6, NAD4L), one subunit of Complex III (COB), three subunits of Complex IV (COX1-3), and three subunits of Complex V (ATP6, ATP8, and ATP9) (Saraste, 1999). In addition to the genes encoding OXPHOS proteins, two rRNA subunits and a set of tRNA genes are included in the fungal mt genome. Gene content is highly conserved in mitochondria of various fungi, but some characteristics, such as gene order, tRNA gene clusters, intergenic regions, mobile elements and introns, can be highly variable (Ghikas et al., 2006; Aguileta et al., 2014; Losada et al., 2014; Mardanov et al., 2014). As of December 2016, there were 179 Ascomycota mt genomes submitted to GenBank, and *Sclerotinia borealis* (NC_025200, 203,051 bp) had the largest mt genome (Mardanov et al., 2014)[1]. Due to its small size, fast evolution, high copy number, and relatively conserved gene content, the mt genome has been successfully used in evolutionary biology and phylogeny (Joardar et al., 2012).

The present understanding of the mt epigenome has gone through a series of evolutions, and conflicting data about epigenomes existed in the 1970s and early 1980s due to the uneven distribution of methylation patterns and low sensitivity of detection methods (Castegna et al., 2015). A few studies have shown mt methylation while others have reported the absence of any methylation in the mitochondrion (Cummings et al., 1974; Dawid, 1974; Groot and Kroon, 1979; Pollack et al., 1984). The advent of next generation sequencing technologies has provided opportunities for characterizing mt genome-wide cytosine methylation. Ghosh et al. (2014) reported the first genome-wide map of human mitochondrial methylation. Ghosh et al. (2016) revealed that it was dynamic in nature of the hydroxymethly cytosine marking in the human mitochondrial genome. However, little trace of DNA modification on mtDNA is detected in fungi. The current methodologies for epigenomics, based on bisulfite, enrichment, or sequencing and PCR, impede the processing of the epigenetic study due to the limits in the resolution, modification type, locus and cost (Mensaert et al., 2014).

Single molecule, real-time sequencing has created significant progress in read length and sequencing of base modification sites (Flusberg et al., 2010; Feng et al., 2013). SMRT sequencing can produce read lengths of up to 40 kb and direct sequencing of base modification. The most important characteristics of SMRT sequencing are that SMRT sequencing (1) could be free from amplification biases and (2) easily sequenced the complex regions containing secondary structure in low coverage (Koren et al., 2013). DNA modifications across the genome could be directly determined by the kinetics of DNA synthesis, because SMRT sequencing monitors the processing of single DNA molecules by DNA polymerase (Eid et al., 2009; Feng et al., 2013). The biggest concern in SMRT sequencing is the relatively high error rate, which can be effectively decreased through multiple sequencing passes. Long reads and a circular consensus sequencing strategy make SMRT convenient for: *de novo* assembly of mitochondria, chloroplasts, and microbial genomes, especially for complex genomes; characterizing genomic structural variations; and analyzing targeted sequencing regions or modification sites (Carneiro et al., 2012; Grad et al., 2012; Chin et al., 2013; Feng et al., 2013; Koren et al., 2013).

In the current study, we sequenced the mt genome of wild fungus *O. sinensis* using SMRT sequencing technology on a PacBio RS II sequencing platform. We described the gene content and genomic organization of this high-altitude fungus, and performed a comparative analysis with the sequenced Hypocreales mt genomes. The main focus in this study is on the mitogenome-wide epigenetic DNA modification pattern of *O. sinensis*. This is the first report on epigenetic modifications in a fungal mt genome, and it will provide the basis for further research on mt epigenetics of Hypocrealean fungi.

MATERIALS AND METHODS

Sample Collection

Sample of the fungus *O. sinensis* (CCTCC AF 2017003) was from the fruiting body of the wild *O. sinensis*. Fresh specimens were purchased in a local market in Guoluo of Qinghai Province, China (Latitude 34.48°N, Longitude 100.23°E). Governmental permission is not required for *O. sinensis* purchases in local markets, and the collections of *O. sinensis* specimen sold by local farmers fall under the governmental regulations for traditional Chinese herbal products. All fresh *O. sinensis* specimens were washed thoroughly on site in running water with gentle brushing, soaked in 75% ethanol for 10 min for surface sterilization and washed again three times with sterile water. These samples were snap-frozen in liquid nitrogen immediately after sampling, transported to our laboratory by using the solid carbon dioxide, and stored at −80°C until further use. The wild *O. sinensis* was preliminarily identified according to the morphological characteristics and several nuclear loci, including internal transcribed spacer region (ITS), small and large 18S nuclear ribosomal RNA subunits (nrSSU, nrLSU) and translation elongation factor 1-alpha (TEF-1α). These four genes (GenBank accession number: MF403011, MF403012, MF403013, MF425658) all showed the highest homology to *O. sinensis* ().

[1]https://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?opt=organelle&taxid=451864

DNA Extraction and Genome Sequencing

Total DNA was extracted from the stroma of *O. sinensis* by a modified CTAB method as previously described (Kang et al., 2011). The integrity, quality and concentration of total DNA were analyzed by agarose gel electrophoresis, NanoDrop 1000 spectrophotometer and Qubit fluorometer. DNA was randomly sheared to fragments with an average size of 20 kb by using g-TUBE. Sheared DNA was then DNA damage repaired and end repaired. SMRTbell templates were obtained by ligating the blunt hairpin adapters to the ends of the repaired fragments, followed by the addition of exonuclease to remove failed ligation products. Before annealing the sequencing primer and binding polymerase to SMRTbell templates, the quality of library was assessed by an Agilent 2100 Bioanalyzer High Sensitivity Kit. Eight SMRT cells were sequenced using P6-C4 reagents on a PacBio RS II sequencing platform (Pacific Biosciences, Nextomics Biosciences, Co., Ltd, Wuhan).

Assembly of the Mitochondrial Genome

Clean data were obtained by filtering out the sequencing adapters and low-quality sequences (parameters: minimum sub-read length = 500 bp; minimum polymerase read quality = 0.80). The mt sequences were extracted from the filtered reads containing both nuclear and mt genomes, using BLASR which matches each read against 201 published fungal mitochondrial genomes (Chaisson and Tesler, 2012). About 2815 subreads in 25.4 Mb mt sequencing data were obtained, with an average read length of 8993 bp and a longest read length of 39,358 bp, reaching an average depth of 167 X. The mt genome was assembled through Hierarchical Genome Assembly Process (HGAP) workflow, including preassembly, error correction, Celera assembly and polishing with Quiver (Chin et al., 2013). Long reads were selected to be "seed" reads, which the other subreads were blast against to improve accuracy. The corrected reads were retained and fully assembled by using overlap-layout-consensus (OLC) algorithm in the Celera Assembler program (Myers et al., 2000), and then further refined with Quiver (Chin et al., 2013).

Mitochondrial Genome Annotation

Genes in *O. sinensis* mt genome were predicted by MFannot, RNAweasel[2] and BLASTn against NCBI Organelle Genome Resources[3] (Altschul et al., 1990; Lang et al., 2007; Valach et al., 2014). The PCGs and rRNA genes were identified by MFannot and checked by BLASTn against Hypocreales (Altschul et al., 1990), while the tRNA genes were checked using RNA weasel. Intron–exon boundaries of the PCGs were adjusted manually on the basis of BLASTn against multiple Hypocreales mt coding sequence (Altschul et al., 1990). The nucleotide sequences of PCGs were translated to protein sequences using the Mold, Protozoan and Coelenterate Mitochondrial code (transl_table = 4). Open reading frames (ORFs > 100 bp) in the intergenic and intronic regions were predicted by MFannot (Valach et al., 2014). Predicted ORFs were analyzed by InterProScan[4]. The mitochondrial genome map was generated with Circos software (Krzywinski et al., 2009) and CLC Sequence Viewer 7.8.1 (CLC bio Inc., Cambridge, MA, United States), and then modified by Adobe Illustrator CC 2015 (Version 19.0.0.44, Adobe, San Jose, CA, United States).

Analysis of Repetitive Sequences

Local BLASTn search of mtDNA against itself was performed using a cut-off e-value of 10^{-7} (Altschul et al., 1990). Repetitive sequences were analyzed by several programs, including REPuter, Tandem Repeats Finder and MIcroSAtellite (Benson, 1999; Kurtz et al., 2001; Thiel et al., 2003). REPuter was applied to identify the forward, reverse, complement and palindromic sequences. Tandem Repeats Finder was used to find tandem repeats while MIcroSAtellite detected the microsatellite DNA (1–6 bp).

Methylation Modification

Pacific Biosciences' SMRTPortal analysis platform v. 1.3.1 was used to identify modified positions. At each genomic position, modQVs were computed as the −10 log (*P*-value) for representing a modified base position, based on the distributions of the kinetics of interpulse durations (IPD ratios) from all reads covering this position and from in silico kinetic reference values (Feng et al., 2013). A value of 20 is the minimum default threshold and corresponds to a *P*-value of 0.01. DNA methylation on both mtDNA strands was assessed independently and represented by modQV, which comprises base incorporation rates differing from that of the unmodified reference sequences. The RS_Modification_and_Motif_Analysis.1 protocol of the SMRT analysis v2.0 was used to identify methylation and the corresponding motifs of the responsible DNA methylases.

Phylogenetic Analysis

A phylogenetic tree was inferred by the ML method using the nucleotides of 14 concatenated PCGs (*atp6*, *atp8*, *atp9*, *cox1-3*, *nad1-6*, *nad4L*, *cob*) in mt genome. The 20 fungal mt genomes downloaded from NCBI were shown in Supplementary Table S1. Multiple sequence alignment was performed using the MAFFT program (Katoh and Standley, 2013). Poorly aligned positions and gap positions were removed with Gblocks (Castresana, 2000). Three species belonging to Glomerellales (*Colletotrichum acutatum*, *Colletotrichum lupini*) and Eurotiales (*Penicillium polonicum*) were used as outgroup taxa in the phylogenetic analysis. The best model used in the ML phylogenetic tree was determined by using "find best DNA/protein models (ML)" in MEGA 7.0 (Kumar et al., 2016). The "GTR + G" model produced the lowest values for both the Bayesian Information Criterion and the corrected Akaike information criterion, therefore it was chosen for phylogenetic analysis. The phylogenetic tree was constructed using RAxML 8.0.19 with 500 bootstrap replicates (Stamatakis, 2006). BI analyses were processed with 1,000,000 generations and four chains (one cold and three hot chains), with sampling every 500 generations and a burn-in of 25% (Ronquist and Huelsenbeck, 2003). The confidence values of

[2]http://megasun.bch.umontreal.ca/RNAweasel/

[3]https://www.ncbi.nlm.nih.gov/genome/organelle/

[4]http://www.ebi.ac.uk/interpro/

the BI tree were shown as Bayesian posterior probabilities in percentages. The NJ phylogenetic tree based on 14 PCGs or *rps3* was constructed by MEGA 7.0 and the evolutionary distances were computed using the Tajima-Nei method.

For *rps3* analysis, the outgroup *P. polonicum* was changed to *Penicillium nordicum* to avoid a large number of gaps and to get more genetic information. Because *rps3* gene in *P. polonicum* contained a premature termination codon, resulting a shorter aa sequence (Kang et al., 2016). The dN, dS, and the ratio (dN/dS) of each *rps3* sequence with the reference of *P. nordicum* were calculated using CODEML (RateAncestor = 1) in PAML v4.7a (Yang, 1997; Yang and Nielsen, 2000). Sequence alignment was performed for *rps3* genes in Hypocreales (length from 398 to 544 amino acids) using ClustalW version 2.0.12 (Sievers et al., 2011). The positive selection pressure on *rps3* was detected using CODEML implemented in PAML v4.8 (Yang and Nielsen, 2002; Yang et al., 2005; Yang, 2007; Lin et al., 2015). Excluding the ambiguous sequences, the *rps3* gene sequences were represented by 322 consensus aa sequences. The models of M0 (one-ratio), M1a (neutral), M2a (selection), M7 (beta), and M8 (β and ω) were used to calculate the dN/dS values.

RESULTS

Characteristics of the *O. sinensis* mt Genome

The complete mt genome of *O. sinensis* was a circular double-stranded DNA molecule with 157,539 bp in length (**Table 1** and **Figure 1**, GenBank accession number: KY622006), which is similar to the *O. sinensis* mt genome (157,510 bp) reported by Li et al. (2015) and much larger than the other fungi in the same order Hypocreales (**Table 1**). The size variation of mt genome was caused by the length of intergenic regions, and the number, length of introns and accessory genes (**Table 1**) (Deng et al., 2016). Although the mtDNA size is in large variation (25,615–157,539 bp), the gene content slightly differed in Hypocreales (**Table 1**). There were 14 PCGs, 1 *rps3*, 27 tRNAs, and 2 rRNA subunits encoded in the mt genome. All protein and RNA coding genes were located on positive strand and orient clockwise (**Figure 1** and **Table 2**). In addition, there were 73 ORFs coding putative proteins in the *O. sinensis* mt genome (**Table 1** and Supplementary Table S2).

In the *O. sinensis* mt genome, the 14 PCGs included 7 NADH dehydrogenases (*nad1-nad6*, *nad4L*), 3 cytochrome c oxidases (*cox1-cox3*), 3 ATP synthases (*atp6*, *atp8*, *atp9*), and 1 cytochrome b gene (*cob*), for encoding proteins involved in respiratory chain complexes. In addition, a *rps3* gene which encodes 40S ribosomal protein S3 was found in the intron of *rnl* (*rnl*_I8). The majority of the PCGs were split by several introns into multiple short exons (**Figure 1** and Supplementary Table S3). The introns ($n = 54$) covered ca. 67.64% of the whole mt genome, with the highest number in *cox1* gene ($n = 14$). The intergenic regions have a total length of 29,350 bp, accounting for 18.63% of the whole genome length (**Table 1**). The whole A+T content of the mt genome was 69.8% (**Table 1**), consistent with the characteristic of AT-rich in the Hypocreales fungal mt genome, ranging from 69.0 to 73.5% (**Table 1**).

All PCGs began with a canonical start codon ATG or GTG (**Table 2**) and most terminated with the stop codon TAA, except *cox1*, *cox3*, and *rps3*. As revealed in many fungi, the ATG initiation codon of *nad5* followed immediately after the termination codon of *nad4L*, with an overlap of a base "A"; the genes of *nad2* and *nad3* were uninterrupted (Kouvelis et al., 2004; Shen et al., 2015). The codon frequency analysis showed that a total of 63 codons were used for transcription, with the absence of CGC (Supplementary Table S4). The six most frequently used codons (TTA, ATA, TTT, AAT, TAT, and AAA) reflect the biased usage of A/T nucleotides (Supplementary Table S4). The fraction of codons encoding the hydrophobic amino acids (Met, Trp, Phe, Val, Leu, Ile, Pro, Ala, accounting for 42.59%) (Supplementary Table S4) could explain the hydrophobic nature of respiratory membrane complexes.

A total of 27 tRNA genes were identified in the mt genome of *O. sinensis* coding for all 20 amino acids (**Table 3**), alleviating the need for tRNA import into the mitochondrion from the cytoplasm (Kolesnikova et al., 2000). The presence of tRNA-W recognizing the UGA codon indicates that the *O. sinensis* mt genome is translated according to genetic code 4 (Fox, 1987). As a unique characteristic of Hypocreales mtDNAs, most of the tRNA genes in *O. sinensis* were organized into three clusters with minor differences, except for five tRNAs scattered as a single gene across the mt genome (**Figure 1**). The single tRNA genes were suggested to play a role in transcription or recombination events (Saccone et al., 2002).

Phylogenetic Analysis of *O. sinensis*

A ML phylogenetic tree of 20 taxa was inferred using 14 conserved PCGs associated with the OXPHOS system. *O. sinensis* was most closely related to *Hirsutella rhossiliensis*, with a bootstrap value of 100% (**Figure 2**), and these species then formed a sub-cluster with *H. minnesotensis* and *Tolypocladium ophioglossoides*. The cluster of these four fungi is consistent with the traditional classification that they are all in Ophiocordycipitaceae. As Lin et al. (2015) found, all entomopathogenic fungi formed a cluster in this ML phylogenetic tree, including the fungi in Clavicipitaceae (*Metarhizium* and *Metacordyceps*), Cordycipitaceae (*Cordyceps*, *Beauveria,* and *Lecanicillium*), and Ophiocordycipitaceae (*Hirsutella*, *Ophiocordyceps*, and *Tolypocladium*), while three *Fusarium* plant pathogens were clustered together. The topology of the ML tree agreed with that based on BI (**Supplementary Figure S1**), but differed slightly from the topology obtained using the NJ method (**Supplementary Figure S2**). In the NJ tree, *T. ophioglossoides* was clustered with *Metacordyceps chlamydosporia* and *Metarhizium anisopliae*, and the fungi in Cordycipitaceae branched earlier than that in the ML tree. The ML phylogenetic relationship inferred from the mt genome is principally consistent with the current Pezizomycotina taxonomic system (Taxonomy Database in NCBI).

Introns

Notable in the *O. sinensis* mtDNA, it is the high degree of invasion by mobile DNA-elements (group I and group II introns). The number of introns exhibits remarkable variation in fungal mt genomes (Burger et al., 2003). The largest number of mt introns was documented for *O. sinensis* ($n = 54$) in Hypocreales in our analysis, while *M. chlamydosporia* has only one intron (**Table 1**) (Lin et al., 2015). In *O. sinensis*, 45 introns were found in 11 PCGs and 9 in 2 rRNA subunits, accounting for 84.17% (89,725 bp) and 15.83% (14,805 bp in *rnl*, 2025 bp in *rns*) of the whole intron length, respectively. As shown in other fungi, introns were abundant in *cox1*, *rnl*, and *cob* genes (Lang et al., 2007), containing 14, 8, and 6 introns, respectively, accounting for 94.82, 75.11, and 92.32% of each gene (**Table 2**). To these, *cox2*, and *nad5* should also be included since 93.18 and 81.15% of each gene were covered by introns, respectively. In contrast, several genes, such as *nad3*, *nad4*, and *atp8*, were intronless in *O. sinensis*.

BLASTx and RNAweasel results showed that 46 Group I introns and 6 Group II introns (Supplementary Table S3) were included in the *O. sinensis* mt genome, with two short introns failure to classify. Group I introns are reported to be dominant in fungal mt genomes, while group II introns are found more frequently in plant mt genome (Lang et al., 2007). Group I introns encode various HE genes with LAGLIDADG or GIY-YIG domain motifs, while group II introns mostly encode RTs (Lang et al., 2007). The motifs of HE genes and RTs in introns catalyze the transfer and site-specific integration of the introns into various genes (Lang et al., 2007).

Open Reading Frames (ORFs)

In total, 73 ORFs were identified by MFannot in addition to the conserved genes (Supplementary Table S2), which is a little higher than that predicted by Li et al. (2015). The difference in ORF number may be due to different method used. The fungi in Hypocreales exhibited a broad spectrum of predicted ORFs in this study, from the least ($n = 1$) in *M. chlamydosporia* (Lin et al., 2015) to the most ($n = 73$) in *O. sinensis* (**Table 1**). The length of the ORFs in *O. sinensis* ranged from 306 to 2397 bp, with a total length of 59,421 bp accounting for 37.72% of the mt genome of *O. sinensis* (Supplementary Table S2). The variation in the number and length of predicted ORFs could partly explain the variation in genome size.

Among the 73 ORFs, 12 were free-standing, and the rest were located in introns (Supplementary Table S2). Among free-standing ORFs, only ORF_71 was predicted to be a DNA-directed RNA polymerase, while the functions of the others were unknown. In addition to the ORFs that exhibited similarities to GIY-YIG/LAGLIDADG endonuclease or RTs in introns, there were some ORFs encoding domains of intron encoded nuclease repeat and nuclease-associated module, which are possibly involved in DNA-binding (Supplementary Table S2).

Molecular Evolution of *rps3*

The *rps3* gene was identified in the *rnl* intron of Hypocreales fungi, except *Acremonium fuci*. The *rps3* in *A. fuci* is freestanding. The length of *rps3* ranged from 1131 bp in *A. implicatum* to 1584 bp in *O. sinensis*. In *O. sinensis*, *rps3* was located in the IA intron of *rnl* (*rnl*_I8). A BLASTx search against the NCBI database showed that the *rps3* was fused to a novel LAGLIDADG HE gene. Half of *rnl*-I8 was homologous to *rps3* in *H. rhossiliensis* and half was similar to the HE gene from *Ophiostoma novo-ulmi* subsp. *americana* (Gibb and Hausner, 2005).

The phylogenetic relationships among the Hypocreales inferred from *rps3* genes were different from that based on 14 PCGs (**Supplementary Figure S3**). The sub-clusters (Hypocreaceae, Cordycipitaceae, Clavicipitaceae, and Ophiocordycipitaceae) in Hypocreales were similar between these two phylogenetic trees, except Nectriaceae. In the phylogenetic tree based on *rps3*, *Fusarium oxysporum,* belonging to Nectriaceae, was clustered with *M. anisopliae* rather than *Fusarium* spp. Moreover, the Cordycipitaceae cluster was separated from the Ophiocordycipitaceae cluster in the phylogenetic tree inferred from *rps3*. To measure the selective pressure of the *rps3* genes in Hypocreales, dN/dS values were calculated. The values were 0.2400, 0.1075, 0.1075, 0.1023, and 0.0848, when the adopted models were M0 (one-ratio), M1a (neutral), M2a (selection), M7 (beta), and M8 (β and ω),

TABLE 1 | Characteristics of mt genomes of Hypocreales fungi.

Species	Total length (bp)	AT (%)	Length (number) of PCGs (bp)	Length (number) of tRNAs (bp)	Length (number) of rRNAs (bp)	Intergenic region (bp)	Length (number) of introns	ORFs in introns	ORFs in intergenic regions
Fusarium oxysporum	34477	69.0	14484 (15)	2603 (25)	4742 (2)	11252	2738 (2)	2	0
Acremonium chrysogenum	27266	73.5	13305 (14)	1936 (26)	6236 (2)	4653	1330 (2)	0	3
Metacordyceps chlamydosporia	25615	71.7	14385 (15)	1625 (23)	4580 (2)	4691	1652 (1)	1	0
Cordyceps militaris	33277	72.2	14707 (15)	1924 (26)	4666 (2)	3863	9868 (8)	1	0
Hypocrea jecorina	42130	72.7	16149 (15)	1937 (26)	4369 (2)	10385	10933 (9)	1	4
Hypomyces aurantius	71638	71.7	16572 (16)	1870 (25)	4633 (2)	19416	30573 (17)	22	7
Hirsutella minnesotensis	52245	71.6	15171 (15)	1863 (25)	4898 (2)	14637	17123 (13)	12	4
Hirsutella rhossiliensis	62483	71.8	14571 (15)	1922 (26)	7067 (2)	28081	12477 (9)	7	3
Tolypocladium ophioglossoides	35159	72.4	14406 (15)	1869 (25)	4694 (2)	8230	7319 (6)	3	2
Ophiocordyceps sinensis	157539	69.8	14913 (15)	2010 (27)	6558 (2)	29350	106555 (54)	61	12

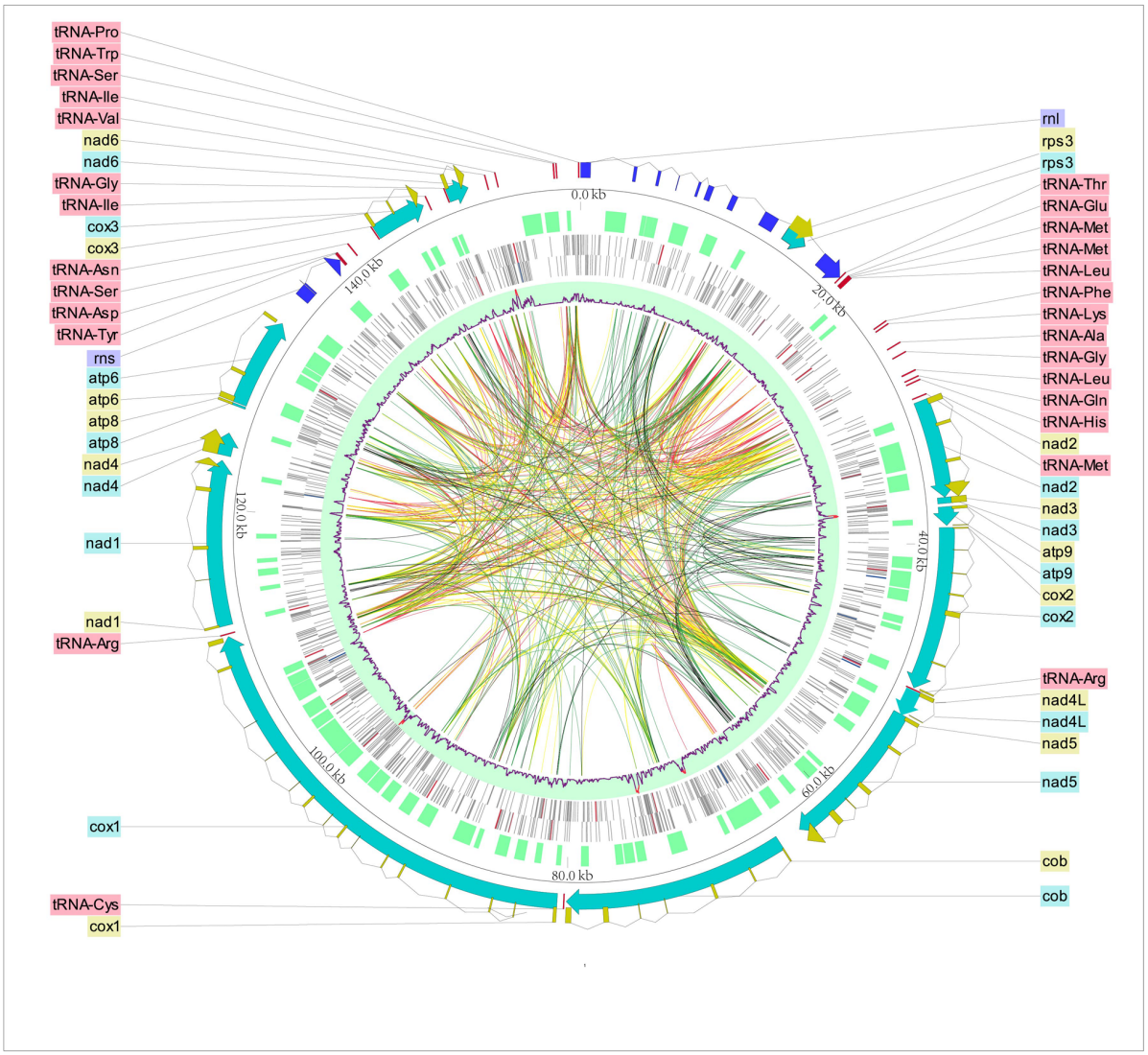

FIGURE 1 | Circular map of the *Ophiocordyceps sinensis* mt genome. The scales indicate location in kb, starting with *rnl*. Using Circos and CLC Sequence Viewer, we integrated the gene annotation, DNA base modification, and some other information. From outermost to innermost: ring 1 with dingy yellow is CDS. Ring 2 includes protein-coding genes (lasureous), rRNAs (blue) and tRNAs (red). ORFs are located on ring 3. Rings 4 and 5 shows DNA base modifications sites in the forward and backward strand, respectively. Red lines indicate 4mC, blue lines show 6mA, while gray lines represent the other base modifications. Ring 6 with light blue as background color is the methylations for both strands in a 200 bp- sliding window. Red lines in the ring 6 mean that the number of the methylation sites in the sliding window is over 10. The ribbons inside the circle connect the repeat sequences with significant (e-value $< 10^{-7}$) similarity (red: identity ≥ 90; yellow: $90 >$ identity ≥ 80; green: $80 >$ identity ≥ 70; gray: identity < 70).

respectively. However, the LRT statistic for comparing M7 [lnL (log likelihood value) $= -10{,}788.91$] and M8 (lnL $= -10{,}702.06$) is 173.70 [$2\Delta = 2 * (10{,}788.91–10{,}702.06) = 173.70$], with a P-value < 0.01 using the Chi-square test (with df $= 2$). To further confirm whether *rps3* genes in Hypocreales are evolving under positive selection or not, the values of dN/dS were calculated in relation to the outgroup *P. nordicum*. The values of dN/dS were above 1.0 in Hypocreales (**Supplementary Figure S3**). Moreover, we explored a total of 36 sites in *rps3* with the values of dN/dS > 1 (**Table 4** and **Figure 3**), and 16 sites were statistically significant ($P \geq 0.95$) (**Table 4** and **Figure 3**).

Repetitive Sequences in the mt Genome of *O. sinensis*

Repetitive genes are considered as putative elements for recombination or regulation (Ghikas et al., 2006). A local self BLASTn of the *O. sinensis* mt genome against itself revealed 710 repetitive sequences (e-value $< 10^{-7}$) with a total length of 77,583 bp, accounting for 49.25% of the whole mt genome (**Figure 1**). As shown in the mt genome of *Phlebia radiata* (Salavirta et al., 2014), the abundant repeat sequences were almost exclusively localized into intronic and intergenic region in *O. sinensis*, in particular between positions 110 to 30 kb

(clockwise, **Figure 1**). Reputer identified a total of 3490 bp (2.22%) repeats in the *O. sinensis* mt genome, including 30 forward repeats (61–128 bp, in total 2446 bp), 13 palindromic repeats (61–90 bp, in total 968 bp), and 1 reverse repeat (76 bp) (**Table 5**). Tandem Repeats Finder found 44 tandem repeats (2–123 bp in copy size), with an average length of 63 bp, and accounting for 1.76% (2772 bp) of the mt genome (**Table 5**).

Simple sequence repeats (also known as microsatellites) comprise tandemly repeated genetic loci of 1–6 bp (Tautz and Renz, 1984). SSRs were found in PCGs (*nad5*, *cox1*), rRNA (*rnl*), and non-coding regions (intronic and intergenic regions, Supplementary Table S5), with being more abundant in non-coding regions than in exons, as previously found in nuclear genome (Karaoglu et al., 2005). It assumed that SSRs might play an active role in genome evolution by creating and maintaining genetic variation (Tautz et al., 1986), and serve a functional role in gene expression regulation by influencing transcriptional activity or protein–protein interactions (Gerber et al., 1994; Kashi et al., 1997). Among these SSRs, most (36/43) were consisted of mono-nucleotide repeats, while the di-, tetra-, penta-, and poly-nucleotide repeats were found with much lower frequency (Supplementary Table S5). The reason may be that longer repeats have higher mutation rates and less stability (Wierdl et al., 1997). In agreement with the previous studies in fungi (Karaoglu et al., 2005), a majority of sequences rich in A/T were observed. All mono-nucleotide repeats were consisted of A/T repeats and the di-nucleotide repeats were AT/TA, with G/C only found in the less identified tetra-, penta-, and poly-nucleotide repeats.

TABLE 2 | Annotation of conserved PCGs and rRNA genes of *O. sinensis* mt genome.

Gene	Start	End	Length (bp)	Number of Introns	Intron length (bp)	Exon length (bp)	Coding sequence density	Start/stop codons	Strand
rnl	1	19712	19712	8	14805	4907	24.89%	-	+
rps3	14881	16464	1584	0	0	1584	100.00%	ATG,TAG	+
nad2	29906	36733	6828	3	5097	1731	25.35%	ATG,TAA	+
nad3	36734	37147	414	0	0	414	100.00%	ATG,TAA	+
atp9	37382	38676	1295	1	1070	225	17.37%	GTG,TAA	+
cox2	38810	50128	11319	6	10536	783	6.92%	ATG,TAA	+
nad4L	50363	52138	1776	1	1506	270	15.20%	ATG,TAA	+
nad5	52138	62723	10586	5	8591	1995	18.85%	ATG,TAA	+
cob	64342	79765	15424	6	14239	1185	7.68%	ATG,TAA	+
cox1	80387	111177	30791	14	29195	1596	5.18%	ATG,TAG	+
nad1	111941	123385	11445	4	10335	1110	10.04%	ATG,TAA	+
nad4	123751	125322	1572	0	0	1572	100.00%	ATG,TAA	+
atp8	127221	127367	147	0	0	147	100.00%	ATG,TAA	+
atp6	127453	133700	6248	2	5468	780	12.48%	ATG,TAA	+
rns	135867	139587	3721	1	2025	1696	45.58%	-	+
cox3	142715	146434	3720	2	2910	810	21.77%	GTG,TAG	+
nad6	148244	149731	1488	1	777	711	44.89%	ATG,TAA	+

TABLE 3 | Transfer RNA genes in the mt genome of *O. sinensis*.

tRNA	Codon usage	Start	End	Length(bp)	Strand	tRNA	Codon usage	Start	End	Length(bp)	Strand
tRNA-Thr(T)	ACA	19715	19786	72	+	tRNA-Cys(C)	UGC	79892	79963	72	+
tRNA-Glu(E)	GAA	20017	20089	73	+	tRNA-Arg(R)	AGA	111323	111393	71	+
tRNA-Met(M)	AUG	20093	20163	71	+	tRNA-Tyr(Y)	UAC	139615	139698	84	+
tRNA-Met(M)	AUG	20166	20238	73	+	tRNA-Asp(D)	GAC	139721	139793	73	+
tRNA-Leu(L)	UUA	20239	20321	83	+	tRNA-Ser(S)	AGC	140661	140741	81	+
tRNA-Phe(F)	UUC	23951	24023	73	+	tRNA-Asn(N)	AAC	142587	142658	72	+
tRNA-Lys(K)	AAA	24188	24260	73	+	tRNA-Ile(I)	AUA	146772	146843	72	+
tRNA-Ala(A)	GCA	25669	25741	73	+	tRNA-Gly(G)	GGA	148122	148192	71	+
tRNA-Gly(G)	GGA	26401	26471	71	+	tRNA-Val(V)	GUA	151005	151077	73	+
tRNA-Leu(L)	CUA	27782	27865	84	+	tRNA-Ile(I)	AUC	151708	151779	72	+
tRNA-Gln(E)	CAA	28238	28311	74	+	tRNA-Ser(S)	UCA	155672	155757	86	+
tRNA-His(H)	CAC	28500	28575	76	+	tRNA-Trp(W)	UGA	155865	155936	72	+
tRNA-Met(M)	AUG	29557	29627	71	+	tRNA-Pro(P)	CCA	157390	157462	73	+
tRNA-Arg(R)	CGU	50205	50275	71	+						

"+" means located on the positive strand.

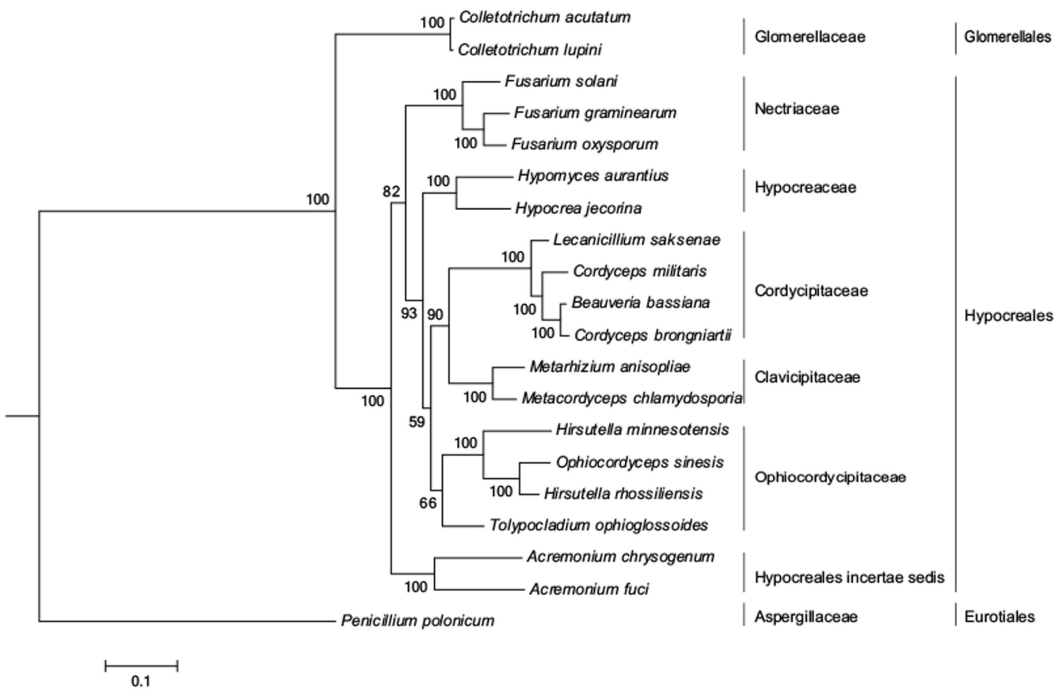

FIGURE 2 | ML phylogeny of *O. sinensis* based on 14 PCGs in mt genome. The phylogenetic tree was inferred from a concatenated alignment of 14 PCGs (*atp6*, *atp8*, *atp9*, *nad1-nad6*, *nad4L*, *cob*, and *cox1-cox3*) using ML analysis. Numbers above branches specify bootstrap percentages (500 bootstrap replicates). MEGA 7.0 was used to determine the best evolutionary model (GTR+R) and RAxML 8.0 was used to infer the phylogenetic tree.

DNA Modification Analysis

DNA modifications were determined in the mt genome of *O. sinensis* in parallel with the acquisition of primary sequence data by SMRT sequencing, based on analysis of the kinetics of DNA synthesis reactions. DNA modifications differ by various modification types, including 5-methylcytosine (5mC), 6-methyladenine (6mA), 4-methylcytosine (4mC) and 5-hydroxymethylcytosine (5-hmC), which can be identified by a series of methyltransferases at specific motifs. A genomic position is covered by several sequenced DNA fragments, and the modQV score shows the consistency by which a specific modification is observed. In the *O. sinensis* mt genome, a total of 1604 sites were determined with an average modQV score of 24.68 at an average coverage of approximately 96x (**Figure 4** and Supplementary Table S6). There were 783 modification sites located on the forward strand versus 821 located on the backward strand (**Figure 4** and Supplementary Table S6). The plot of modification scores against sequencing coverage displayed a dominant signal for modified adenosines and thymine bases (pink and burgundy dots in **Figure 4**). Twenty-eight 4mC (0.13%) and 10 6mA (0.017%) modification sites were identified in the *O. sinensis* mt genome (**Figure 1** and **Table 6**). The 6mA levels were lower than previously reported for nucleotide methylation (0.048–0.21%) in eukaryotic nuclear genomes (Mondo et al., 2017). Most 6mA and 4mC were distributed in intergenic regions (between tRNA and *nad4L*/*nad6*/*cox2*, or between tRNA and tRNA) or in intron regions of different genes (e.g., *nad1-2*, *cox1-2*, *cob*, *rnl*), with only three located in the genes. Two 4mC were methylated on the backward strand in the regions of *nad2* and *nad5* genes, and one 6mA was methylated in the *nad4L* region (**Table 6**). The hypermethylation in promoters and introns, and hypomethylation in exons, are also found in the nuclear genome of eukaryotes (Volpe, 2005). It is reported that methylation is strand-specific within the mt genome (Bellizzi et al., 2013). However, methyl modifications in the *O. sinensis* mt genome were observed on both strands, including 4mC and 6mA modifications (Supplementary Table S6 and **Table 6**).

Because DNA methylation is diverse, widespread, and often deposited by a diverse set of methyltransferases at specific target sequences (motif), a comprehensive understanding of the distribution and diversity of methylation motifs will benefit comprehension of methylation function and evolutionary history of the mt genome. Eight modification motifs "AANNN^{m4}CAGCANNANNNNA," five of which were recognized by 4mC methyltransferases, were detected in the *O. sinensis* genome by SMRT sequencing with a mean QV of 44.4 at a mean coverage of 90.4x (Supplementary Table S6). Among the five methylation motifs, four were in the introns of PCGs (*cob*, *cox1*, and *rnl*), one was in the an intergenic region (tRNA_I_tRNA_G).

DISCUSSION

Single molecule, real-time sequencing, which was recently developed by Pacific Biosciences, can achieve extraordinarily long

TABLE 4 | Log-likelihood values and parameter estimates for *rps3* genes in Hypocreales.

Model	ℓ	Estimates of parameters	Positively selected sites
M1a (neutral)	−10810.63	$\widehat{p}_0 = 0.94987$, $\widehat{p}_1 = 0.05013$	Not allowed
M2a (selection)	−10810.63	$\widehat{p}_0 = 0.94987$, $\widehat{p}_1 = 0.00563$, $\widehat{p}_2 = 0.04450$ $\widehat{\omega}_0 = 0.06045$, $\widehat{\omega}_1 = 1.00000$, $\widehat{\omega}_2 = 1.00000$	6V, 80D, 81E, **84T**, **85L**, 87N, 92A, **94S**, 98E, **101A**, 119E, 192Q, 204V, 219Q, 232S, 235E, **236A**, **237K**, **238S**, 253G, **273V**, 275K, 305S, **321S**
M7 (beta)	−10788.91	$\widehat{p} = 0.06958$, $\widehat{q} = 0.60258$	Not allowed
M8 (β and ω)	−10702.06	$\widehat{p}_0 = 0.98400$, $\widehat{p}_1 = 0.01600$ $p = 0.15774$, $\widehat{q} = 2.18092$, $\widehat{\omega}_s = 1.46813$	6V, 30S, 80D, 81E, **84T**, **85L**, 87N, 92A, **94S**, **98E**, 99L, 100L, **101A**, 102N, 119E, 135E, 145N, 192Q, **204V**, 205I, 207V, 209D, **219Q**, 220V, 231L, **232S**, **235E**, **236A**, **237K**, **238S**, 242I, **253G**, **273V**, **275K**, 305S, **321S**

p is the number of parameters in the ω distribution. Positive selection sites with P ≥ 95% were shown in bold.

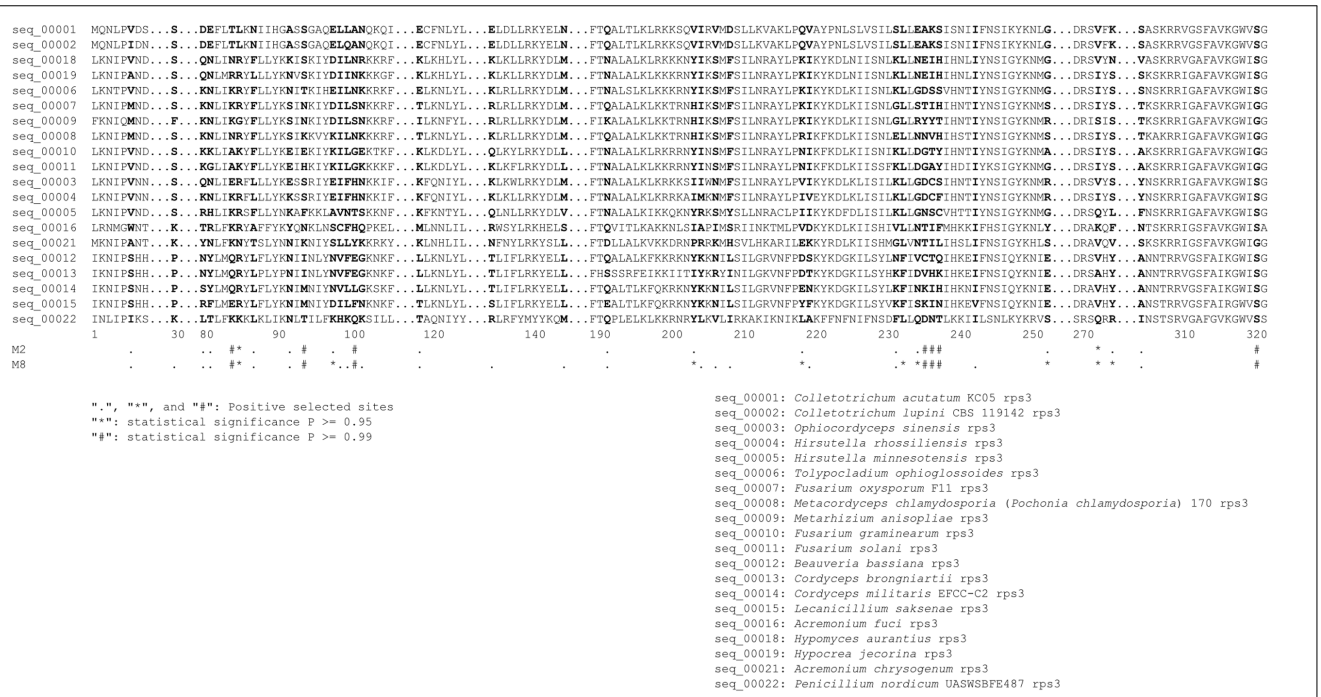

FIGURE 3 | Positive selection sites across *rps3* in Hypocreales. Positive selection sites were identified by Bayes Empirical Bayes dN/dS values and labeled with symbols of ".", "*" (*P* ≥ 0.95) and "#" (*P* ≥ 0.99).

reads (up to 40 kb) with no GC bias, so that it is suitable for sequencing the genomes with low GC and high repeat content. Moreover, SMRT sequencing enables direct genome-wide detection of diverse base modifications by monitoring the kinetic variations of single bases. Thus, SMRT sequencing is convenient for sequencing the mt genome of *O. sinensis* and revealing the base modification pattern of the mt genome. It has long been debated whether modified bases exist in mt genomes.

Previously, *O. sinensis* has been called *Cordyceps sinensis* or *Cordyceps* sp. due to its morphology. However, the colony characteristics of *O. sinensis* cultures are significant different from other *Cordyceps* spp. Most *Cordyceps* species produce bright color and fleshy stromata, while the stromata of *O. sinensis* are often darkly pigmented, tough, fibrous to pliant, and have aperithecial apices, which are the characteristics of *Ophiocordyceps* (Sung et al., 2007). Based on these diagnostic characters and several loci sequence analyses (*nrSSU*, *nrLSU*, *tef1*, *rpb1*, *rpb2*, *tub*, and *atp6*), *C. sinensis* was classified into Ophiocordycipitaceae and renamed "*O. sinensis*" in 2007 (Sung et al., 2007). Our phylogenetic tree (**Figure 2**) based on the 14 PCGs confirms the phylogenetic position of *O. sinensis* in Ophiocordycipitaceae.

TABLE 5 | Repetitive sequences in the mt genome of *O. sinensis*.

Type	Number	Repeat size (bp)	Total length (bp)	Percentage of genome (%)
Forward repeats	30	61–128	2446	1.55
Palindromic repeats	13	61–90	968	0.61
Reverse repeats	1	76	76	0.05
Tandem repeats	44	1–123	2772	1.76
Microsatellite DNA	43	1–6	729	0.46

The mt genome intron number dynamics in the compared fungi (from 1 to 54) are most likely caused by the gain or

loss of introns (Haugen et al., 2005; Cuenca et al., 2016). The introns are probably acquired from ancestors or gained through horizontal gene transfer (Haugen et al., 2005; Lang et al., 2007). HE domains in group I introns help to splice the transcribed intronic RNA, and are then removed from the transcribed pre-mRNA resulting in a contiguous RNA transcript in the process of gene transcription (Taylor and Stoddard, 2012). However, not all group I introns possess auto-catalytic splicing activity, especially for the mini-introns (e.g., *rnl*_I5 and *nad6*_I1), which lack the typical features (Schafer et al., 1991; Lang et al., 2007). It is proposed that the loss of intron splicing activity may be compensated by some other assistant proteins (Kreike et al., 1987; Lambowitz and Perlman, 1990; Lang et al., 2007). If the numerous HE motifs in the *O. sinensis* mtDNA are active, the HE domains might regulate the transcription of their target genes or modify their target genes, as seen in bacterial viruses and the animal-pathogenic *Cryptococcus* spp. (Ma et al., 2009; Stoddard, 2011). In addition, the existence of HEs within fungal mtDNA genes could promote intron mobility, genetic diversity and adaptive responses for mt genomes, when the allelic recombination events may be impossible or rare in mtDNA due to maternal inheritance (Barr et al., 2005; Basse, 2010).

Within the fungi, *rps3* is extremely diverse in location and organization: some are lost, some are free-standing, some are incorporated into group I intron, and others have been invaded by HEs (Sethuraman et al., 2009). Among the Ascomycete fungi, group I intron-encoded *rps3* seems to have a rather complex evolutionary history (Sethuraman et al., 2009). The phylogeny of *rps3* genes inferred in this study displays an evolutionary pattern that is different from that based on 14 conserved PCGs in Hypocrealean fungi, similar to the report of Lin et al. (2015). The *rps3* in *A. fuci* was a free standing gene and was grouped with the intron encoded *rps3*, implying that the *rps3* gene in *A. fuci* was somehow relocated and the *rnl* intron was lost. *rps3* is prone to insertion and deletion, is a refuge for HE genes, and is subject to recombination, making it problematic to resolve taxonomic relationships using *rps3* (Sethuraman et al., 2009).

To evaluate the balance between purifying selection, neutral evolution and positive selection acting on *rps3*, the dN/dS ratio, an indicator of evolutionary pressure on a gene, was examined in Hypocreales fungi. A dN/dS value < 1.0 in Hypocreales indicates that *rps3* has evolved under functional constraints. However, the significance of LRT showed a clear signal of positive selection in *rps3*. In addition, the dN/dS values that were > 1 in Hypocreales in relation to the outgroup *P. nordicum* and the 36 explored positive selection sites suggest that *rps3* is under positive selection in Hypocreales fungi, which is in accordance with the results of Lin et al. (2015). *rps3* is involved in DNA repair and potentially has endonuclease activities in *Schizosaccharomyces pombe* and nuclear versions of *rps3* in human and *Drosophila melanogaster* (Neu et al., 1998; Lyamouri et al., 2002; Jang et al., 2004). Except for these model organisms, no functional studies confirm that mtDNA-encoded *rps3* genes are actually functional (Kim et al., 2009). Further studies are needed to determine if the *rps3* genes in the mt genomic introns produce functional proteins and how these intron-encoded *rps3* genes are expressed.

Base modifications are indispensable parts of comprehending biological processes such as host-pathogen interactions, DNA damage and DNA repair (Trygve, 2010). Among base modifications, DNA methylation has been one of the most studied modifications. It has emerged as a significant phenotypic determinant for disease susceptibility and pathogenesis in eukaryotes, and involved in the Restriction- Modification system in prokaryotes (Bernstein et al., 2010; Ershova et al., 2015). However, whether mtDNA can be the site of epigenetic modifications has long been mired in controversy because the

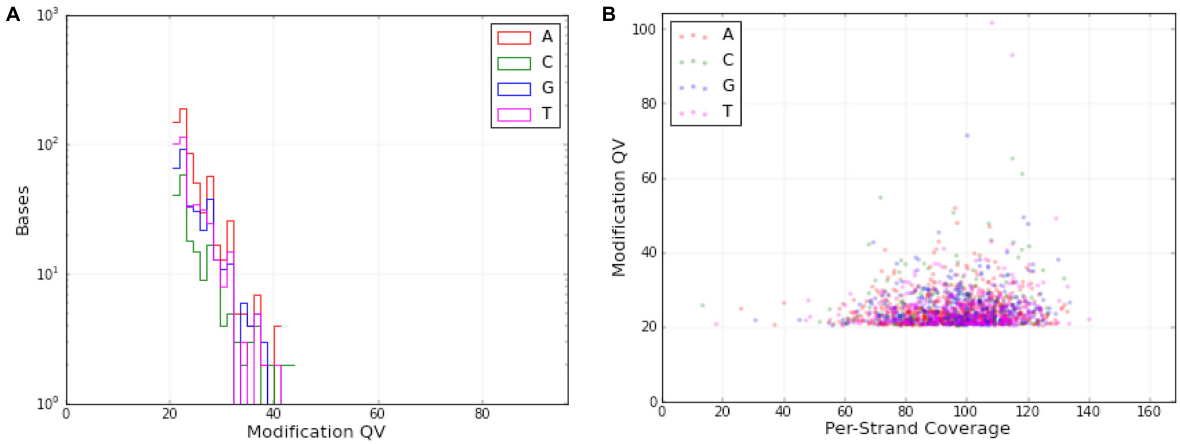

FIGURE 4 | DNA modifications in the *O. sinensis* mt genome. **(A)** Distribution of the number of DNA modification bases with different modification quality value (QV). **(B)** Scatter plot of the DNA modifications with different per-strand coverage. At each genomic position, modification QV were computed as the −10 log (*P*-value) for representing a modified base position, based on the distributions of the kinetics of interpulse durations (IPD ratios) from all reads covering this position and from *in silico* kinetic reference values. Each dot represents a position on either strand with a modQV larger than 20. The color specified the nucleotide base, on which the modification was detected. Adenosines were colored in red, guanosines were colored in green, cytimidines were colored in blue, and thymines were colored in purple.

TABLE 6 | 6mA and 4mC in the mt genome of *O. sinensis*.

Type	Site	Location	modQV	Strand	Type	Site	Location	modQV	Strand
6mA	33641	*nad2*_intron2	30	+	4mC	92477	*cox1*_intron6	41	–
4mC	33651	*nad2*_intron2	25	+	4mC	99214	*cox1*_intron10	33	–
4mC	36698	*nad2*	25	–	4mC	103645	*cox1*_intron12	29	–
4mC	42050	*cox2*_intron3	27	–	6mA	106958	*cox1*_intron13	43	+
6mA	42687	*cox2*_intron3	34	–	4mC	107482	*cox1*_intron13	33	–
6mA	46658	*cox2*_intron5	37	+	4mC	108067	*cox1*_intron13	38	–
6mA	50156	tRNA_R_*nad4L*	27	–	4mC	112158	*nad1*_intron1	30	–
4mC	50187	tRNA_R_*nad4L*	39	–	6mA	122026	*nad1*_intron4	48	+
6mA	50384	*nad4L*	39	–	4mC	130958	*atp6*_intron2	32	–
4mC	62448	*nad5*	39	–	4mC	147647	tRNA_I_tRNA_G	50	+
6mA	65521	*cob*_intron1	52	+	4mC	151590	tRNA_V_tRNA_I	29	–
4mC	72293	*cob*_intron3	61	+	6mA	151635	tRNA_V_tRNA_I	25	+
4mC	72810	*cob*_intron3	42	–	4mC	6591	*rnl*_intron3	32	–
4mC	72821	*cob*_intron3	43	–	6mA	17260	*rnl*_intron8	30	+
4mC	77747	*cob*_intron6	43	+	4mC	17266	*rnl*_intron8	55	+
4mC	81629	*cox1*_intron1	34	–	4mC	20529	tRNA_L_tRNA_F	20	+
4mC	82628	*cox1*_intron1	65	+	4mC	21252	tRNA_L_tRNA_F	42	+
4mC	85113	*cox1*_intron3	23	–	4mC	23245	tRNA_L_tRNA_F	20	+
4mC	86263	*cox1*_intron3	35	–	4mC	26565	tRNA_G_tRNA_L	25	+

mt genome is multi-copy, lacks canonical CpG islands and histones (D'Aquila et al., 2017). SMRT sequencing is a more innovative and sensitive technology by comparing to the first and second generation DNA sequencing. Its advent makes the discovery of mitochondrial epigenetics, including various DNA modifications, more convenient. In this study, we revealed a mitogenome-wide epigenetic DNA modification pattern of *O. sinensis* by SMRT sequencing, resolving the issue of whether or not epigenetic modifications exist in the mt genome of fungi.

Among the DNA modifications in the mt genome of *O. sinensis*, several 6mA and 4mC methylations have been identified, which are always found in prokaryotes and have the function of protecting against restriction enzymes, regulating virulence and controlling DNA replication, repair, and expression (Ratel et al., 2006). Bisulfite sequencing has enabled genome-wide surveys of 5mC methylation (Masser et al., 2016), a well-established epigenomic mark in eukaryotes, but the historic absence of tools for studying 6mA and 4mC modifications has precluded more comprehensive studies of 6mA and 4mC methylation. Recently, Mondo et al. (2017) firstly analyzed 6mA in fungi by SMRT, and revealed its role as a gene-expression-associated epigenomic mark. In this study, SMRT sequencing revealed that not only 6mA but also 4mC were presented in the fungal mt genome. Nucleotide modifications are one of the most evolutionarily conserved properties of RNA, and the sites of modification are under strong selective pressure (Li and Mason, 2014). Mitochondria are largely thought to originate from endosymbiotic α-proteobacteria species (Gray, 2012). Thus, the 6mA methylations, which are often found in α-proteobacteria, in the fungal mitogenome might be remnants of the methylation system from α-proteobacteria. However, the function of these epigenetic modifications in the mt genome of *O. sinensis* remains unknown.

Mitochondria are crucial for responses to hypobaria, hypothermia, and hypoxia due to their central role in energy production and consumption (Chandel and Schumacker, 2000; Chitra and Boopathy, 2013). Cold temperature and low oxygen pressure are the two most remarkable characters of high-altitude environments, where *O. sinensis* is endemic (Li et al., 2011; Yu et al., 2011). Current associations between the mt genome and high-altitude adaption is focused on the mtDNA content and polymorphisms of mt genes, including NADH dehydrogenases, cytochrome c oxidases, ATP synthases and cytochrome b (Luo et al., 2013). However, no studies had been conducted on mt epigenetics. In this study, we found only three methylation affecting coding regions (in *nad2*, *nad4L,* and *nad5*), while the rest is spread between intergenic regions and introns (Supplementary Table S6). Complex I in the mitochondria, including NAD1-6, NAD4L, are involved in collecting electrons from various donors and passing them to coenzyme Q, which then passes the electrons to Complexes III (including COB) and IV (including COX1-3), and finally to the final electron acceptor O_2 to complete the electron transport chain. At the end of this chain, Complex V (including ATP6, ATP8, and ATP9) catalyzes the reaction between ADP and inorganic phosphorus to form ATP (Saraste, 1999). The DNA methylations in eukaryotes are closely related with DNA transcription and translation (Volpe, 2005). Therefore, we supposed that the methylations across the whole *O. sinensis* mt genome, covering protein-coding regions, introns, and intergenic regions, might be closely related to the expression of the electron transport subunits (*nad1-2*, *nad4L*, *nad6* in Complex I, *cob* in Complex III, c*ox 1-2* Complex IV, and *atp6* in Complex V), and thereby modulate the enzymatic activity of OXPHOS and the mt respiratory rate, and then change the ability to capture O_2 and produce energy. Moreover, methylation of mtDNA is supposed to play an important role

in mitochondrial biogenesis by affecting the binding affinity of transcription factor A to mtDNA, impacting the relative activity of promoters (Van der Wijst and Rots, 2015). As the mentioned above, the methylation of mtDNA in *O. sinensis* might be a genetic feature for adaptation to the cold and low PO_2 environment at high altitudes (Volpe, 2005; Luo et al., 2013).

CONCLUSION

Single molecule, real-time sequencing was applied to characterize the *O. sinensis* mt genome in order to mitigate problems with assembly due to high AT and repeat content of the mt genome. The phylogenetic tree inferred from 14 PCGs supports the phylogenetic position of *O. sinensis* in Ophiocordycipitaceae. A total of 36 sequence sites were explored with the values of dN/dS > 1, suggesting that positive selection acts on *rps3* in Hypocreales fungi. Furthermore, we have analyzed the DNA modification pattern of the mitogenome directly. This is the first report of methylation in a fungal mitochondrion, and we propose that methylations in the *O. sinensis* mt genome might closely relate with the environmental responses for adapting to the cold and low PO_2 environments at high altitude sites where *O. sinensis* is endemic.

AUTHOR CONTRIBUTIONS

DL conceived this study. XK analyzed the data and drafted the manuscript. LH participated in the data analysis and prepared figures. PS participated in the data analysis. RL sequenced and analyzed the mt genome. All authors have read and approved the final manuscript.

FUNDING

This work was supported by the Program of International Science & Technology Cooperation of Ministry of Science and Technology (2013DFG32060).

ACKNOWLEDGMENTS

We thank Runmao Lin at Beijing Normal University for his help in exploring the positively selected signals in *rps3* genes, and a referee for thoughtful comments on the topic of mitochondrial DNA methylations.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.01422/full#supplementary-material

FIGURE S1 | Bayesian phylogeny of *O. sinensis* based on 14 PCGs in mt genome. MrBayes 3.2.6 was applied to construct the phylogenetic tree. Numbers at the nodes are Bayesian posterior probabilities.

FIGURE S2 | Neighbor-joining phylogeny of *O. sinensis* based on 14 PCGs in mt genome. Numbers above branches specify bootstrap percentages (500 bootstraps). The evolutionary distances were computed using the Tajima-Nei method. MEGA 7.0 was used to construct the phylogenetic tree.

FIGURE S3 | Evolutionary characteristics of *rps3* for 20 different taxa. Numbers above branches specify bootstrap percentages (500 bootstraps). The evolutionary history was inferred using the Neighbor-Joining method and the evolutionary distances were computed using the Tajima-Nei method. Evolutionary analyses were conducted in MEGA7.0. The right of the species names in the phylogenetic tree are the dN/dS values, which are calculated based on the comparative analysis with the *rps3* sequence of *P. nordicum*, an outgroup for the phylogenetic analysis.

REFERENCES

Aguileta, 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 Biol. Evol.* 6, 451–465. doi: 10.1093/gbe/evu028

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Barr, C. M., Neiman, M., and Taylor, D. R. (2005). Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. *New Phytol.* 168, 39–50. doi: 10.1111/j.1469-8137.2005.01492.x

Basse, C. W. (2010). Mitochondrial inheritance in fungi. *Curr. Opin. Microbiol.* 13, 712–719. doi: 10.1016/j.mib.2010.09.003

Bellizzi, D., D'aquila, P., Scafone, T., Giordano, M., Riso, V., Riccio, A., et al. (2013). The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern. *DNA Res.* 20, 537–547. doi: 10.1093/dnares/dst029

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573

Bernstein, B. E., Stamatoyannopoulos, J. A., Costello, J. F., Ren, B., Milosavljevic, A., Meissner, A., et al. (2010). The NIH roadmap epigenomics mapping consortium. *Nat. Biotechnol.* 28, 1045–1048. doi: 10.1038/nbt1010-1045

Burger, G., Gray, M. W., and Lang, B. F. (2003). Mitochondrial genomes: anything goes. *Trends Genet.* 19, 709–716. doi: 10.1016/j.tig.2003.10.012

Carneiro, M. O., Russ, C., Ross, M. G., Gabriel, S. B., Nusbaum, C., and Depristo, M. A. (2012). Pacific biosciences sequencing technology for genotyping and variation discovery in human data. *BMC Genomics* 13:375. doi: 10.1186/1471-2164-13-375

Castegna, A., Iacobazzi, V., and Infantino, V. (2015). The mitochondrial side of epigenetics. *Physiol. Genomics* 47, 299–307. doi: 10.1152/physiolgenomics.00096.2014

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334

Chaisson, M. J., and And Tesler, G. (2012). Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* 13:238. doi: 10.1186/1471-2105-13-238

Chandel, N. S., and Schumacker, P. T. (2000). Cellular oxygen sensing by mitochondria: old questions, new insight. *J. Appl. Physiol.* 88, 1880–1889.

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. *Nat. Methods* 10, 563–569. doi: 10.1038/nmeth.2474

Chitra, L., and Boopathy, R. (2013). Adaptability to hypobaric hypoxia is facilitated through mitochondrial bioenergetics: an in vivo study. *Br. J. Pharmacol.* 169, 1035–1047. doi: 10.1111/bph.12179

Commission, C. P. (2015). *Pharmacopoeia of the People's Republic of China*. Beijing: China Medical Science Press.

Cuenca, A., Ross, T. G., Graham, S. W., Barrett, C. F., Davis, J. I., Seberg, O., et al. (2016). Localized retroprocessing as a model of intron loss in the plant mitochondrial genome. *Genome Biol. Evol.* 8, 2176–2189. doi: 10.1093/gbe/evw148

Cummings, D. J., Tait, A., and Goddard, J. M. (1974). Methylated bases in DNA from *Paramecium aurelia*. *Biochim. Biophys. Acta* 374, 1–11. doi: 10.1016/0005-2787(74)90194-4

D'Aquila, P., Montesanto, A., Guarasci, F., Passarino, G., and Bellizzi, D. (2017). Mitochondrial genome and epigenome: two sides of the same coin. *Front. Biosci. (Landmark Ed)* 22, 888–908.

Dawid, I. B. (1974). 5-methylcytidylic acid: absence from mitochondrial DNA of frogs and HeLa cells. *Science* 184, 80–81. doi: 10.1126/science.184.4132.80

Deng, Y., Zhang, Q., Ming, R., Lin, L., Lin, X., Lin, Y., et al. (2016). Analysis of the mitochondrial genome in *Hypomyces aurantius* reveals a novel twintron complex in fungi. *Int. J. Mol. Sci.* 17, E1049. doi: 10.3390/ijms17071049

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009). Real-time DNA sequencing from single polymerase molecules. *Science* 323, 133–138. doi: 10.1126/science.1162986

Ershova, A. S., Rusinov, I. S., Spirin, S. A., Karyagina, A. S., and Alexeevski, A. V. (2015). Role of restriction-modification systems in prokaryotic evolution and ecology. *Biochemistry (Mosc)* 80, 1373–1386. doi: 10.1134/S0006297915100193

Feng, Z., Fang, G., Korlach, J., Clark, T., Luong, K., Zhang, X., et al. (2013). Detecting DNA modifications from SMRT sequencing data by modeling sequence context dependence of polymerase kinetic. *PLoS Comput. Biol.* 9:e1002935. doi: 10.1371/journal.pcbi.1002935

Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., et al. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* 7, 461–465. doi: 10.1038/nmeth.1459

Fox, T. D. (1987). Natural variation in the genetic code. *Annu. Rev. Genet.* 21, 67–91. doi: 10.1146/annurev.ge.21.120187.000435

Gerber, H. P., Seipel, K., Georgiev, O., Hofferer, M., Hug, M., Rusconi, S., et al. (1994). Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science* 263, 808–811. doi: 10.1126/science.8303297

Ghikas, D. V., Kouvelis, V. N., and Typas, M. A. (2006). The complete mitochondrial genome of the entomopathogenic fungus *Metarhizium anisopliae* var. anisopliae: gene order and trn gene clusters reveal a common evolutionary course for all Sordariomycetes, while intergenic regions show variation. *Arch. Microbiol.* 185, 393–401. doi: 10.1007/s00203-006-0104-x

Ghosh, S., Sengupta, S., and Scaria, V. (2014). Comparative analysis of human mitochondrial methylomes shows distinct patterns of epigenetic regulation in mitochondria. *Mitochondrion* 18, 58–62. doi: 10.1016/j.mito.2014.07.007

Ghosh, S., Sengupta, S., and Scaria, V. (2016). Hydroxymethyl cytosine marks in the human mitochondrial genome are dynamic in nature. *Mitochondrion* 27, 25–31. doi: 10.1016/j.mito.2016.01.003

Gibb, E. A., and Hausner, G. (2005). Optional mitochondrial introns and evidence for a homing-endonuclease gene in the mtDNA rnl gene in *Ophiostoma ulmi* s. lat. *Mycol. Res.* 109, 1112–1126. doi: 10.1017/S095375620500376X

Grad, Y. H., Lipsitch, M., Feldgarden, M., Arachchi, H. M., Cerqueira, G. C., Fitzgerald, M., et al. (2012). Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3065–3070. doi: 10.1073/pnas.1121491109

Gray, M. W. (2012). Mitochondrial evolution. *Cold Spring Harb. Perspect. Biol.* 4:a011403. doi: 10.1101/cshperspect.a011403

Groot, G. S., and Kroon, A. M. (1979). Mitochondrial DNA from various organisms does not contain internally methylated cytosine in -CCGG- sequences. *Biochim. Biophys. Acta* 564, 355–357. doi: 10.1016/0005-2787(79)90233-8

Haugen, P., Simon, D. M., and Bhattacharya, D. (2005). The natural history of group I introns. *Trends Genet.* 21, 111–119. doi: 10.1016/j.tig.2004.12.007

Jang, C. Y., Lee, J. Y., and Kim, J. (2004). RpS3, a DNA repair endonuclease and ribosomal protein, is involved in apoptosis. *FEBS Lett.* 560, 81–85. doi: 10.1016/S0014-5793(04)00074-2

Joardar, V., Abrams, N. F., Hostetler, J., Paukstelis, P. J., Pakala, S., Pakala, S. B., et al. (2012). Sequencing of mitochondrial genomes of nine *Aspergillus* and Penicillium species identifies mobile introns and accessory genes as main sources of genome size variability. *BMC Genomics* 13:698. doi: 10.1186/1471-2164-13-698

Kang, X. C., Liu, C. C., Liu, D. B., Zeng, L., Shi, Q. Q., Qian, K., et al. (2016). The complete mitochondrial genome of huperzine A-producing endophytic fungus *Penicillium polonicum*. *Mitochondrial DNA Part B Resour.* 1, 202–203. doi: 10.1080/23802359.2016.1155086

Kang, X. C., Liu, D. B., Xia, Z. L., and Chen, F. (2011). Comparison of genomic DNA extraction from *Cordyceps militaris*. *J. Hunan Agric. Univ. (Nat. Sci.)* 37, 210.

Karaoglu, H., Lee, C. M., and Meyer, W. (2005). Survey of simple sequence repeats in completed fungal genomes. *Mol. Biol. Evol.* 22, 639–649. doi: 10.1093/molbev/msi057

Kashi, Y., King, D., and Soller, M. (1997). Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.* 13, 74–78. doi: 10.1016/S0168-9525(97)01008-1

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kim, T. S., Kim, H. D., and Kim, J. (2009). PKCdelta-dependent functional switch of rpS3 between translation and DNA repair. *Biochim. Biophys. Acta* 1793, 395–405. doi: 10.1016/j.bbamcr.2008.10.017

Kolesnikova, O. A., Entelis, N. S., Mireau, H., Fox, T. D., Martin, R. P., and Tarassov, I. A. (2000). Suppression of mutations in mitochondrial DNA by tRNAs imported from the cytoplasm. *Science* 289, 1931–1933. doi: 10.1126/science.289.5486.1931

Koren, S., Harhay, G. P., Smith, T. P., Bono, J. L., Harhay, D. M., Mcvey, S. D., et al. (2013). Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol.* 14:R101. doi: 10.1186/gb-2013-14-9-r101

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 Genet. Biol.* 41, 930–940. doi: 10.1016/j.fgb.2004.07.003

Kreike, J., Schulze, M., Ahne, F., and Lang, B. F. (1987). A yeast nuclear gene, *MRS1*, involved in mitochondrial RNA splicing: nucleotide sequence and mutational analysis of two overlapping open reading frames on opposite strands. *EMBO J.* 6, 2123–2129.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633

Lambowitz, A. M., and Perlman, P. S. (1990). Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. *Trends Biochem. Sci.* 15, 440–444. doi: 10.1016/0968-0004(90)90283-H

Lang, B. F., Laforest, M. J., and Burger, G. (2007). Mitochondrial introns: a critical view. *Trends Genet.* 23, 119–125. doi: 10.1016/j.tig.2007.01.006

Li, S., and Mason, C. E. (2014). The pivotal regulatory landscape of RNA modifications. *Annu. Rev. Genomics Hum. Genet.* 15, 127–150. doi: 10.1146/annurev-genom-090413-025405

Li, Y., Hu, X. D., Yang, R. H., Hsiang, T., Wang, K., Liang, D. Q., et al. (2015). Complete mitochondrial genome of the medicinal fungus *Ophiocordyceps sinensis*. *Sci. Rep.* 5, 13892. doi: 10.1038/srep13892

Li, Y., Wang, X. L., Jiao, L., Jiang, Y., Li, H., Jiang, S. P., et al. (2011). A survey of the geographic distribution of *Ophiocordyceps sinensis*. *J. Microbiol.* 49, 913–919. doi: 10.1007/s12275-011-1193-z

Lin, R., Liu, C., Shen, B., Bai, M., Ling, J., Chen, G., et al. (2015). Analysis of the complete mitochondrial genome of *Pochonia chlamydosporia* suggests a close relationship to the invertebrate-pathogenic fungi in Hypocreales. *BMC Microbiol.* 15, 5. doi: 10.1186/s12866-015-0341-8

Losada, L., Pakala, S. B., Fedorova, N. D., Joardar, V., Shabalina, S. A., Hostetler, J., et al. (2014). Mobile elements and mitochondrial genome expansion in the soil fungus and potato pathogen *Rhizoctonia solani* AG-3. *FEMS Microbiol. Lett.* 352, 165–173. doi: 10.1111/1574-6968.12387

Luo, Y., Yang, X., and Gao, Y. (2013). Mitochondrial DNA response to high altitude: a new perspective on high-altitude adaptation. *Mitochondrial DNA* 24, 313–319. doi: 10.3109/19401736.2012.760558

Lyamouri, M., Enerly, E., and Lambertsson, A. (2002). Organization, sequence, and phylogenetic analysis of the ribosomal protein S3 gene from *Drosophila virilis*. *Gene* 294, 147–156. doi: 10.1016/S0378-1119(02)00763-1

Ma, H., Hagen, F., Stekel, D. J., Johnston, S. A., Sionov, E., Falk, R., et al. (2009). The fatal fungal outbreak on Vancouver Island is characterized by enhanced intracellular parasitism driven by mitochondrial regulation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12980–12985. doi: 10.1073/pnas.0902963106

Mardanov, A. V., Beletsky, A. V., Kadnikov, V. V., Ignatov, A. N., and Ravin, N. V. (2014). The 203 kbp mitochondrial genome of the phytopathogenic fungus *Sclerotinia borealis* reveals multiple invasions of introns and genomic duplications. *PLoS ONE* 9:e107536. doi: 10.1371/journal.pone.0107536

Masser, D. R., Stanford, D. R., Hadad, N., Giles, C. B., Wren, J. D., Sonntag, W. E., et al. (2016). Bisulfite oligonucleotide-capture sequencing for targeted base- and strand-specific absolute 5-methylcytosine quantitation. *Age (Dordr.)* 38:49. doi: 10.1007/s11357-016-9914-1

Mensaert, K., Denil, S., Trooskens, G., Van Criekinge, W., Thas, O., and De Meyer, T. (2014). Next-generation technologies and data analytical approaches for epigenomics. *Environ. Mol. Mutagen.* 55, 155–170. doi: 10.1002/em.21841

Mondo, S. J., Dannebaum, R. O., Kuo, R. C., Louie, K. B., Bewick, A. J., Labutti, K., et al. (2017). Widespread adenine N6-methylation of active genes in fungi. *Nat. Genet.* 49, 964–968. doi: 10.1038/ng.3859

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. doi: 10.1126/science.287.5461.2196

Neu, R., Goffart, S., Wolf, K., and Schafer, B. (1998). Relocation of urf a from the mitochondrion to the nucleus cures the mitochondrial mutator phenotype in the fission yeast *Schizosaccharomyces pombe*. *Mol. Gen. Genet.* 258, 389–396. doi: 10.1007/s004380050746

Pollack, Y., Kasir, J., Shemer, R., Metzger, S., and Szyf, M. (1984). Methylation pattern of mouse mitochondrial DNA. *Nucleic Acids Res.* 12, 4811–4824. doi: 10.1093/nar/12.12.4811

Ratel, D., Ravanat, J. L., Berger, F., and Wion, D. (2006). N6-methyladenine: the other methylated base of DNA. *Bioessays* 28, 309–315. doi: 10.1002/bies.20342

Ronquist, F., and Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574. doi: 10.1093/bioinformatics/btg180

Saccone, C., Gissi, C., Reyes, A., Larizza, A., Sbisa, E., and Pesole, G. (2002). Mitochondrial DNA in metazoa: degree of freedom in a frozen event. *Gene* 286, 3–12. doi: 10.1016/S0378-1119(01)00807-1

Salavirta, H., Oksanen, I., Kuuskeri, J., Mäkelä, M., Laine, P., Paulin, L., et al. (2014). Mitochondrial genome of *Phlebia radiata* is the second largest (156 kbp) among fungi and features signs of genome flexibility and recent recombination events. *PLoS ONE* 9:e97141. doi: 10.1371/journal.pone.0097141

Saraste, M. (1999). Oxidative phosphorylation at the fin de siecle. *Science* 283, 1488–1493. doi: 10.1126/science.283.5407.1488

Schafer, B., Merlos-Lange, A. M., Anderl, C., Welser, F., Zimmer, M., and Wolf, K. (1991). The mitochondrial genome of fission yeast: inability of all introns to splice autocatalytically, and construction and characterization of an intronless genome. *Mol. Gen. Genet.* 225, 158–167. doi: 10.1007/BF00282654

Sethuraman, J., Majer, A., Iranpour, M., and Hausner, G. (2009). Molecular evolution of the mtDNA encoded rps3 gene among filamentous ascomycetes fungi with an emphasis on the Ophiostomatoid fungi. *J. Mol. Evol.* 69, 372–385. doi: 10.1007/s00239-009-9291-9

Shen, X. Y., Li, T., Chen, S., Fan, L., Gao, J., and Hou, C. L. (2015). Characterization and phylogenetic analysis of the mitochondrial genome of *Shiraia bambusicola* reveals special features in the order of pleosporales. *PLoS ONE* 10:e0116466. doi: 10.1371/journal.pone.0116466

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. doi: 10.1038/msb.2011.75

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446

Stoddard, B. L. (2011). Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure* 19, 7–15. doi: 10.1016/j.str.2010.12.003

Sung, G. H., Hywel-Jones, N. L., Sung, J. M., Luangsa-Ard, J. J., Shrestha, B., and Spatafora, J. W. (2007). Phylogenetic classification of Cordyceps and the clavicipitaceous fungi. *Stud. Mycol.* 57, 5–59. doi: 10.3114/sim.2007.57.01

Tautz, D., and Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* 12, 4127–4138. doi: 10.1093/nar/12.10.4127

Tautz, D., Trick, M., and Dover, G. A. (1986). Cryptic simplicity in DNA is a major source of genetic variation. *Nature* 322, 652–656. doi: 10.1038/322652a0

Taylor, G. K., and Stoddard, B. L. (2012). Structural, functional and evolutionary relationships between homing endonucleases and proteins from their host organisms. *Nucleic Acids Res.* 40, 5189–5200. doi: 10.1093/nar/gks226

Thiel, T., Michalek, W., Varshney, R. K., and Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 106, 411–422. doi: 10.1007/s00122-002-1031-0

Trygve, T. (2010). *Handbook of Epigenetics: The New Molecular and Medical Genetics*. Cambridge, MA: Academic Press.

Valach, M., Burger, G., Gray, M. W., and Lang, B. F. (2014). Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. *Nucleic Acids Res.* 42, 13764–13777. doi: 10.1093/nar/gku1266

Van der Wijst, M. G., and Rots, M. G. (2015). Mitochondrial epigenetics: an overlooked layer of regulation? *Trends Genet.* 31, 353–356. doi: 10.1016/j.tig.2015.03.009

Volpe, P. (2005). The language of methylation in genomics of eukaryotes. *Biochemistry (Mosc)* 70, 584–595. doi: 10.1007/s10541-005-0152-x

Wierdl, M., Dominska, M., and Petes, T. D. (1997). Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics* 146, 769–779.

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* 13, 555–556. doi: 10.1093/bioinformatics/13.5.555

Yang, Z., and Nielsen, R. (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* 17, 32–43. doi: 10.1093/oxfordjournals.molbev.a026236

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088

Yang, Z., and Nielsen, R. (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol. Biol. Evol.* 19, 908–917. doi: 10.1093/oxfordjournals.molbev.a004148

Yang, Z., Wong, W. S., and Nielsen, R. (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22, 1107–1118. doi: 10.1093/molbev/msi097

Yu, L., Wang, X., Ting, N., and Zhang, Y. (2011). Mitogenomic analysis of Chinese snub-nosed monkeys: evidence of positive selection in NADH dehydrogenase genes in high-altitude adaptation. *Mitochondrion* 11, 497–503. doi: 10.1016/j.mito.2011.01.004

Yue, K., Ye, M., Zhou, Z., Sun, W., and Lin, X. (2013). The genus Cordyceps: a chemical and pharmacological review. *J. Pharm. Pharmacol.* 65, 474–493. doi: 10.1111/j.2042-7158.2012.01601.x

Zhu, J. S., Halpern, G. M., and Jones, K. (1998). The scientific rediscovery of an ancient Chinese herbal medicine: *Cordyceps sinensis*: part I. *J. Altern. Complement. Med.* 4, 289–303. doi: 10.1089/acm.1998.4.3-289

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Kang, Hu, Shen, Li and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.